Check for updates

Examining Multilevel Poverty-Causing Factors in Poor Villages: a Hierarchical Spatial Regression Model

Yanhui Wang[1,2,3] · Yuewen Jiang[1,2,3] · Duoduo Yin[1,2,3] · Chenxia Liang[1,2,3] · Fuzhou Duan[1,2,3]

Abstract

The examination of poverty-causing factors and their mechanisms of action in poverty-stricken villages is an important topic associated with poverty reduction issues. Although the individual or background effects of multilevel influencing factors have been considered in some previous studies, the spatial effects of these factors are rarely involved. By considering nested geographic and administrative features and integrating the detection of individual, background, and spatial effects, a bilevel hierarchical spatial linear model (HSLM) is established in this study to identify the multilevel significant factors that cause poverty in poor villages, as well as the mechanisms through which these factors contribute to poverty at both the village and county levels. An experimental test in the region of the Wuling Mountains in central China revealed the following findings. (1) There were significant background and spatial effects in the study area. Moreover, 48.28% of the overall difference in poverty incidence in poor villages resulted from individual effects at the village level. Additionally, 51.72% of the overall difference resulted from background effects at the county level. (2) Poverty-causing factors were observed at different levels, and these factors featured different action mechanisms. Village-level factors accounted for 14.29% of the overall difference in poverty incidence, and there were five significant village-level factors. (3) The hierarchical spatial regression model was found to be superior to the hierarchical linear model in terms of goodness of fit. This study offers technical support and policy guidance for village-level regional development.

Keywords Hierarchical spatial linear model (HSLM) · Multilevel · Spatial effect · Background effects · Poverty factors · Development difference

✉ Fuzhou Duan
  duanfuzhou@cnu.edu.cn

Extended author information available on the last page of the article

🍂 Springer

Introduction

Regional poverty is one of the most severe challenges faced by the international community. Regional poverty has been widely studied by scholars across the globe, and China has made a significant contribution to global poverty reduction through its precision poverty alleviation policy (Dunford et al., 2019; Wang & Chen, 2017). However, poverty is still regional, relative, and dynamic (Wang et al., 2019). It is still an arduous task to alleviate relative poverty and to maintain stable poverty alleviation after 2020 despite the current rapid development of social poverty reduction actions. Ambitious programmes are required, such as the United Nations 2030 Sustainable Development Goals policy or the comprehensive rural rejuvenation programme proposed by the Chinese government. In recent years, poverty in rural China, especially in contiguous poverty-stricken areas, has changed from a singular economic problem to a complex and diverse problem because of great changes in rural society and rapid growth in economics (Jiang et al., 2020). Academics are required to seriously consider poverty issues and to propose new coping models and methods for improving the relevance and practicability of research. A key role in the accurate implementation of anti-poverty policies is to correctly evaluate the development of anti-poverty and to accurately detect the factors related to its causes. Accurately detecting the causes of poverty and their mechanisms has become the focus of current research in the field of poverty alleviation to provide guidance and support for the precise determination of why poverty occurs and approaches for reducing the impacts of poverty-causing factors.

Many experts and scholars have attempted to explore the causes of poverty from different views using different methods (Carneiro et al., 2016; Wang & Qian, 2017; Aristondo, 2018; Skare et al., 2018; Boemi & Papadopoulos, 2019). These existing studies have been mostly based on traditional single-level linear or spatial statistical regression models, such as orthogonal least squares (OLS) regression, least squares regression, multiple linear regression, and spatial regression models (Behruz et al., 2014; Guo et al., 2018; Peirovedin et al., 2016; Thongdar et al., 2012). These studies examined candidate poverty-contributing factors on a single level and only considered the individual factors at the dependents' levels that could be used to identify the self-development characteristics of the poor populations or poor areas; however, these studies ignored the influences of those factors identifying their surrounding environment in which they live on their sustainable development (Kim, 2019; Kwadzo, 2015; Odhiambo, 2019).

Furthermore, recent studies have been directed towards investigating multidimensional and multilevel poverty-contributing factors. For example, Kim et al. (2016) examined the geographical background factors of poverty from three levels: family, village, and state. Ward (2016) analysed the transference of rural economic poverty using balanced panel data. These studies have proven that poverty-causing factors are no longer limited to individual characteristics. There may be multilevel factors; that is, the poverty status of the poor may be affected not only by multidimensional factors such as family characteristics and economic

development at the individual level but also by multidimensional background factors poverty-stricken people depend on at higher levels, such as economic development, social development, the ecological environment, and so on (Alkire, 2019; Boemi & Papadopoulos, 2019; Ibrahim et al., 2016; Jiang et al., 2020; Liang et al., 2019; Mowat, 2019). This indicates that the occurrence of poverty is driven by both the individual effects (i.e., variations caused by individual characteristics) and background effects (i.e., group effects or pond effects, which refer to variations caused by an individual's environment) of poverty-causing factors at different levels (Jiang et al., 2020; Ma et al., 2018; Park & Nam, 2018; Wang et al., 2019).

Therefore, some other studies have used the hierarchical linear model (HLM) to study both individual and background effects. For example, Chen et al. (2015) used this model to explore the poverty-related factors of families in Taiwan at both the household level and the regional level. Ren et al. (2017) used this model to study the factors causing poverty in counties in contiguous destitute areas of China at both the county and district levels. Wang et al. (2019) used this model to explore village-level regional development differences at both the village level and the county level. However, related studies have shown that there may be spatial effects in the evolution of poverty; the occurrence of poverty may be affected by the development of other spatially adjacent individuals in the region, resulting in spatial aggregation (Wang et al., 2019). This effect can cause heterogeneity with different degrees of poverty alleviation based on spatial proximities among poverty-stricken villages and may also have an influence on the detection of poverty-related factors (Liang et al., 2019; Su et al., 2019). In this case, the objectivity and reliability of the results of current studies need to be further examined due to their neglect of spatial effects (i.e., spatial dependence and spatial heterogeneity). In addition, among the existing quantitative studies of poverty-related factors, most studies in China and globally have taken provincial, municipal, county, village and other geographic unit scales as the identifying objects (Alkire, 2019; Carneiro et al., 2016; Kim et al, 2016; Michalek & Madajova, 2019; Wang & Chen, 2017; Ward, 2016), and there are few studies that have used administrative villages as research units due to limitations in data availability. These administrative villages are the smallest regional administrative unit for which regional poverty alleviation measures can be implemented in China. Furthermore, these units can have more effective impacts on regional poverty alleviation than other higher-level units, such as counties and provinces.

From the new, combined perspectives of the multilevel comprehensive detection of poverty-contributing factors, this study selects poverty-stricken villages as the basic measurement units (microlevel) and selects counties as the background level (meso-administrative level), considers the nested geographic and administrative features between county units and village units, and attempts to improve the HLM and to construct a hierarchical spatial linear model (HSLM) that considers the combined influences of the individual, background, and spatial effects of poverty-related factors on poor villages. Furthermore, this study attempts to answer the following three questions by using an empirical study from the region of the Wuling Mountains in China. (1) Do background effects and spatial effects exist? If so, how can the impacts of spatial effects on the detection of poverty-causing factors be mitigated

while background effects are taken into account? (2) What are the significant poverty-causing factors at each level in the study area, and how do these poverty-causing factors interact among different levels? (3) Does the HSLM perform better than the HLM in the detection of poverty-causing factors? By doing so, this study can not only reveal the multilevel effects of poverty-causing factors but can also further highlight the scientificity and accuracy of the new model in the detection of poverty-causing factors and provide more scientific and reasonable support for the implementation of precise policy measures for poverty alleviation.

Study Area and Data Source

Study Area

This paper selects the contiguous poverty-stricken areas of the Wuling Mountains as the research area, as shown in Fig. 1. As one of the 14 contiguous poverty-stricken areas designated by the Chinese government and one of the main battlefields for rural poverty alleviation in the new stage of sustainable development and overall revitalization, this region is known as a "hard bone" in current targeted poverty alleviation measures. This area is located at the junction of Hubei, Hunan, Chongqing, and Guizhou. The study area features complex terrain, a large interface with other provinces and cities, many ethnic minorities, and a wide distribution of impoverished residents. The study area includes 71 counties, including 11 counties in Hubei Province, 37 counties in Hunan Province, 7 counties in Chongqing city, and 16
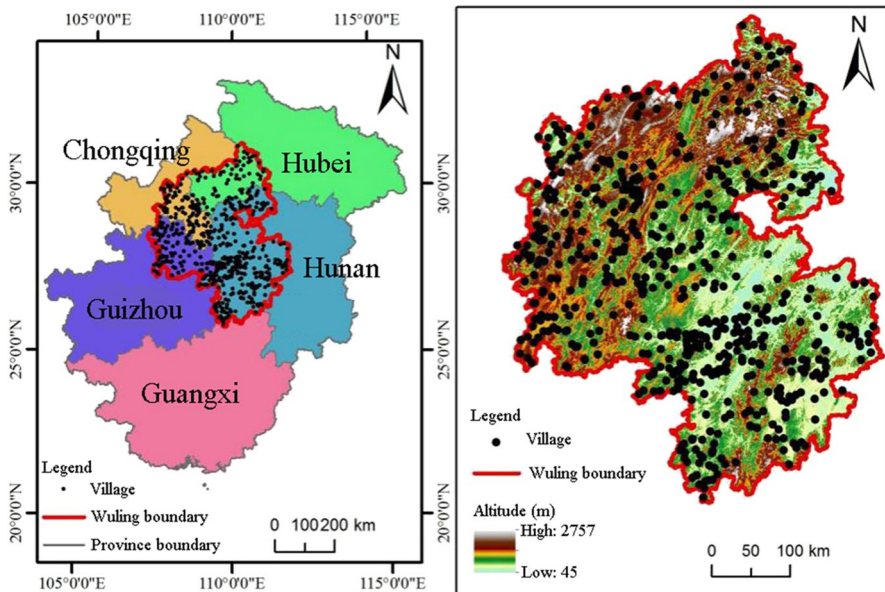

Fig. 1 Study area

counties in Guizhou Province. A total of 11,303 impoverished villages in the study area have been involved in the "village promotion" list. Poverty has been historically prevalent in the study area. The area has a low level of economic development, weak infrastructure, a fragile ecological environment, and uneven regional development.

Data Sources

The individual sample data for the poverty-stricken villages used in this study come from the study area's "Entire-Village Advancement" archived village dataset, which was issued in 2013 by China's State Council Leading Group Office of Poverty Alleviation and Development (COPAD). Given that COPAD data are not open to the outside world due to China's population data privacy, a stratified sampling method was used to extract 1,482 poverty-stricken villages from the data source as sample objects. This was necessary to ensure the representativeness and full coverage of the samples. Other socioeconomic statistics were obtained from the statistical yearbooks of the study area in 2013. All these data identified the socioeconomic development information of each village and county in the study area, e.g., the production and living conditions, social security, medical and health facilities, road infrastructure, economic income, education and school, and land resources of each village and county.

The geographic data included a 1:250,000 national geographic dataset and 90-m-resolution digital elevation models. The ecological data at the county level were obtained using a comprehensive calculation method (Cao et al., 2016). The meteorological data, temperature data, and other thematic data came from the Earth System Science Data Sharing Platform of China (www.geodata.cn). Moreover, we obtained the vector locations of the poverty-stricken villages from the Baidu Map API by using the geographic address interpretation method.

We pre-processed all of the above geographic and socioeconomic data by joint adoption of georeferencing, vectorization, stitching, clipping, and other methods in ArcGIS 10.2. The original data were standardized by the adoption of range normalization.

Methods

Basic Idea

If multiple poor villages belong to the same county administratively, then these villages tend to share certain socioeconomic development and geographical environment characteristics. Observations from the individual villages are not fully independent; indeed, these neighbouring villages tend to be more similar than randomly derived villages sampled from the entire county are, and are thus not fully independent (Wang et al., 2019).

However, most analytic techniques (e.g., OLS) require the independence of observations as a primary assumption for regression analysis, and this assumption that observations should be independent is violated in the presence of

hierarchical data (Graves, 2011). Furthermore, this fact may be ignored and may severely inhibit the validity of a study's results when some form of multilevel modelling is not adopted. Due to possible spatial or background effects, the variations between individuals may be mutated (Wang et al., 2019). In fact, only by concentrating on individual effects while ignoring background and spatial effects would individual villages be seen as independent of each other. In this situation, against the assumption of independence, the poverty similarity degree of intragroup individuals with similar backgrounds would be higher than that of intergroup individuals with different backgrounds; these results would cause the correlation coefficient between the independent and dependent variables to represent an incorrect result, increasing the mistake probability because the observed effects would include both individual and group effects. On the other hand, if a model only pays attention to group effects but ignores individual effects, allowing the independent variables to work only at the second level, it thus loses important individual information.

In this context, regression models with random effects or spatial effects are becoming increasingly popular in the statistical analysis of data. HLM is an effective statistical method for processing data with nested structures. HLMs, also referred to as mixed-effect models, random-effect models or multilevel linear models, separate the variance in the explanatory (independent) variables into individual variations and intergroup defences and analyses the action modes and intensities of different levels of independent variable effects on the response (dependent) variables. Meanwhile, it allows the intercepts (means) and slopes (relationships between independent variables and dependent variables) to vary between higher level units (Graves, 2011). This variability can then be modelled by treating the group intercepts and slopes as dependent variables in the next level of analysis.

Although the traditional HLM can effectively deal with background effects and can analyse the degree and difference in the independent variable's effect on the dependent variable at different levels (Wang et al., 2019), it cannot reduce the spatial effects among independent variables. A so-called spatial effect refers to the phenomenon of material, energy and information redistribution and transmission complexity in geographical systems caused by differences in surface structure and the transformation of spatial patterns. Spatial effects indicate the surface gathering and dispersing organization of economic and social activities and their processes, and these spatial processes are actually processes of migration. In short, a spatial effect comprises the different responses of geographical locations to various social and economic affairs that result in various effects on the properties of objects, reaction mechanisms, and reaction speeds. Understanding the influences on these reactions is sometimes helpful, which is called the spatial assisted effect; sometimes these reactions are hindered, which is called the spatial blocking effect. Spatial effects can be expressed different aspects., e.g., the redistribution of material and energy by surface morphology; the geographical flow caused by the imbalances of the land surface causes different performances with changes in the distance from the centre; the unequal effects of product distribution and social consumption due to space; the formation, expression and influence of the spatial geographical field; the laws of energy, matter and movement in non-equilibrium space, and so on.

Indeed, spatial effects refer to spatial dependence in empirical data, including spatial autocorrelation and spatial heterogeneity. Spatial dependence can arise because of an omitted variable that is correlated with the spatial locations of sample data, for example, uncaptured county characteristics. Spatial dependence can take the form of spatial heterogeneity of observations, as some farms are located in primarily rural areas, while other farms are located in suburban areas (Seo, 2016). In this study, if the explanatory variables of the impoverished villages are related and not independent, it may lead to the existence of spatial effects on both the overall spatial dependence and local spatial heterogeneity of the poverty incidence. Ignoring these spatial effects will cause errors or deviations in the model estimation results (Buonanno et al., 2012). Therefore, the first prerequisite step before HLM modelling is to detect the significance of the spatial effects before determining the poverty-causing factors. If the spatial effects are significant, then we need to reduce the impacts of the spatial effects on the detection accuracy of the poverty-causing factors in the model.

Therefore, in terms of the disadvantage that most previous research on poverty was conducted on a large regional scale and the detection of poverty-causing factors ignored the background effects or spatial effects of the studied groups, this study attempts to establish a bilevel spatial hierarchical linear model (HSLM) to explore the combined influences of the surrounding county-level socioeconomic and environmental factors at the neighbourhood level and the socioeconomic variables at the regional village level on the poverty incidence in villages. Specifically, poverty incidence was used as the dependent variable. By using an empirical test, this study first examines the spatial effects and background effects that may exist in the distribution of poverty-stricken villages, and an improved hierarchical spatial linear model (i.e., HSLM) is then built to attempt to detect the background effects while weakening the possible impacts of the spatial effects to more accurately reveal the significant poverty-contributing factors and their action mechanisms between the village level and the county level.

Prerequisite Test of HSLM Modelling: Detection of Spatial Effects

For both poverty reduction and the development of poverty-stricken villages, there is a certain degree of mutual influence between geographically adjacent villages. This mutual influence is related to the spatial distance between villages. The distance between villages reflects a certain spatial effect and exhibits a certain degree of spatial dependence and spatial heterogeneity (Goodchild et al., 2000; Su et al., 2019). In the detection of poverty-related factors, significant spatial effects such as spatial dependence and spatial heterogeneity are likely to affect the detection process and lead to inaccurate detection results. To respond, this study uses spatial autocorrelation, i.e., global spatial autocorrelation statistics (*Moran's I*) and local spatial autocorrelation statistics (*Getis-Ord* index, *Gi**), in GIS to detect the significance levels of the spatial dependence and spatial heterogeneity of poverty distribution in the study area, respectively. The spatial autocorrelation definition in GIS measures the degree to which objects can be compared to proximate objects and helps clarify the degree to which a given object is similar to other nearby objects. The global Moran's

index (Moran's I) and local index of spatial association (*Gi**) are two such spatio-statistical measures that have been proven useful for examining spatial aggregation distributions (Goodchild et al, 2000, Wang et al., 2018).

Moran's I can reflect the spatial dependency characteristics of spatial data in the entire study area to an extent. *Gi** can reflect the local clustering characteristics of spatial data (or the spatial heterogeneity characteristics) in the study area (Das & Ghosh, 2017; Liu et al, 2019). The calculation formula for *Moran's I* is shown in formula (1) (Getis & Ord, 1992).

$$Moran's I = \frac{N}{\sum_{ij} w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \tag{1}$$

In the formula, i and j represent village numbers, and i is not equal to j; N is the number of villages; x is the poverty incidence of the village; and w is the spatial weight matrix. Moran's I is generally between -1 and 1. A Moran's I value closer to 1 suggests that there is a closer relationship and more significant spatial dependence between the villages.

The calculation formula for G_i^* is shown in formula (2) (Getis & Ord, 1992).

$$G_i^* = \frac{\sum_{j \neq i}^n W_{ij} X_j}{\sum_{j \neq i}^n X_j} \tag{2}$$

In the Formula, *i* and *j* Represent Village Numbers, And *I* is Not Equal to *j*; *n* is the Number of Villages; *x* is the Poverty Incidence of the Village; and *w* is the Spatial Weight Matrix. *Gi** is Normalized to Acquire *Z*, Which is Shown in Formula (3).

$$Z = \frac{G_i^* - E(G_i^*)}{\sqrt{Var(G_i^*)}} \tag{3}$$

$E(G_i^*)$ represents the expectation of G_i and *Var* (G_i^*) represents the variance of G_i^*. In spatial autocorrelation, the null hypothesis is rejected at the 5% confidence level when the absolute value of *Z* is greater than 1.96. This suggests that there is a spatial autocorrelation of poverty incidence. Moreover, a large absolute value for *Z* suggests that there is significant spatial heterogeneity.

HSLM Modelling

To take into account individual villages' poverty degrees, adjusted for group variations, as well as predictions of group villages' poverty degrees, adjusted for individual variations within groups, a new model, the HSLM, was proposed based on the HLM to systematically detect significant poverty-causing factors and their mechanisms at different levels. The HSLM was constructed to analyse data with nested geographic and administrative structural features, in which lower-level village analysis units are nested within higher-level county analysis units, based on the candidate set of indicators of poverty-related factors at the village and county levels.

Although the HSLM can be used on data with many levels, we will only build 2-level models to detect the poverty-contributing factors in the villages in this study due to the limited data availability. The lowest level of analysis is level 1 ($L1$, village level), and the second lowest level is level 2 ($L2$, county level). The village level corresponds to the individual level, and the county level corresponds to the background level. By adopting HLM7.0, we will systematically examine the multilevel factors that affect the sustainable development of poor villages at both the village and county levels, as well as the action mechanisms of these factors as they affect the development of poverty-stricken villages.

Candidate Indicator System

The regional poverty degree of a village is affected not only by its own individual factors, such as the economic development, social development and ecological environment at the individual village level, but also possibly by the surrounding environment of the village, such as the economic, social and environmental factors at a higher county level. Moreover, the "Entire-Village Advancement" strategy in rural China was a key targeted measure for poverty alleviation, aiming to eliminate spatial poverty traps and ensure the eradication of poverty by 2020. This strategy clearly required that poverty alleviation measures in administrative villages should be comprehensively promoted from the combined perspectives of improving living standards and improving the ecological environment and social and economic development. Therefore, the candidate indicator system of poverty-causing factors at different scales in this study should be considered from the above combined perspectives to ensure the sustainable development of the studied villages. During selection of the candidate indicator system, we also adopted principles such as typicality, representativeness, and availability, and a candidate set of indicators was constructed that included dimensions based on the village level (micro level) and county level (meso level) as well as the actual characteristics of the study area at different levels.

As shown in Table 1, we proposed that there were five dimensions at the village level, including the geographical environment, demographics, production and living conditions, labour status, and medical facilities and social security. The reasons for the selection of these five dimensions are as follows: the geographical environment reflects the living environment in the countryside, and a better living environment is associated with improved village development. The demographic characteristics can reflect the distribution of resources, and a larger village population results in fewer resources per capita and lower village development. The production and living conditions indicate the ability of villagers to meet their own needs, and a higher production capacity can lead to higher incomes. A village's living conditions reflect the development of infrastructure in the village. The labour status of a village can indicate the villagers' abilities to obtain income and support themselves. Medical facilities and social security can reflect the status of social welfare, which can help to alleviate health conditions and reduce mortality rates. Moreover, a higher rate of coverage leads to less pressure on villagers.

In the same way, based on the principles of the typicality, representativeness and accessibility of the selected indicators combined with the actual characteristics of

Table 1 Candidate indicators and screening at the village and county levels

Level	Dimension	Index	Paraphrase	Descriptive statistics				C.V (retain? yes or no)	Collinearity detection (retain? yes or no)
				Min	Max	Mean	S.D		
	Geographical environment (X1)	X_{11}	Distance from the nearest town's bazaar*	The distance from the village to the nearest town's bazaar (km)	0.10	41.20	9.64	15.35	yes	yes
		X_{12}	Terrain type*	Terrain type (namely, plain, hilly, plateau, basin)	1	3	2.69	0.46	yes	yes
		X_{13}	Frequency of natural disasters*	The frequency of natural disasters	0	2	1.37	0.88	yes	yes
	Demographics (X2)	X_{21}	Population density*	Population density (person/km²)	11	425	353.04	337.08	yes	yes
		X_{31}	Per capita cultivated land area	Per capita cultivated area (km²)	0.56	3.68	1.11	1.45	yes	yes
		X_{32}	Road access ratio	Villages with motorways (%)	10	100	74.24	30.01	yes	yes
Village level	Production and living conditions (X3)	X_{33}	Electricity access ratio	Households with electricity (%)	95	100	97.49	7.24	no	—
		X_{34}	Phone access ratio	Households with a telephone (%)	8	100	73.44	23.29	yes	yes
		X_{35}	Safe drinking water access ratio	Households with safe drinking water (%)	20	100	54.67	23.09	yes	yes
	Labour status (X4)	X_{41}	Labour force ratio	Total labour force (%)	21.57	79.94	53.23	11.20	yes	yes
		X_{42}	Migrant labour force ratio	Migrant labour force (%)	13	67	51.96	19.36	yes	yes

Table 1 (continued)

Level	Dimension	Index	Paraphrase	Descriptive statistics					C.V (retain? yes or no)	Collinearity detection (retain? yes or no)
				Min	Max	Mean	S.D			
		X_{43}	Ratio of illiterate labour forces	Total labour force with junior high school education or above (%)	5	89	63.11	26.17	yes	yes
	Medical facilities and social security (X5)	X_{51}	Ratio of population enrolled in the new rural cooperative medical insurance	People participating in the new rural cooperative medical insurance of China (%)	85	100	92.59	12.58	no	—
		X_{52}	Ratio of population enrolled in the new rural basic pension insurance	People enrolled in the new rural basic pension insurance (%)	9	100	67.51	28.87	yes	yes
		Z_{11}	Per capita GDP (yuan)	County GDP/the number of people in the county	7,799	27,142	14,725	4,880	yes	yes
		Z_{12}	Per capita income (yuan)	(net income of farmers + disposable income of urban residents)/the number of people in the county	817	18,240	7,426	2,803	yes	yes
	Economy ($Y1$)	Z_{21}	Gross enrolment rate in the first three years	Children aged 3 to 6 entering kindergarten (%)	15	100	75.54	17.42	yes	yes

Table 1 (continued)

Level	Dimension	Index	Paraphrase	Descriptive statistics				C.V (retain? yes or no)	Collinearity detection (retain? yes or no)
				Min	Max	Mean	S.D		
		Z_{22} Secondary gross enrolment ratio	High school students to total age-appropriate population (%)	30	97.8	77.85	15.87	yes	yes
		Z_{23} Number of hospital beds per ten thousand people	The number of county beds/total county population	16.38	62.52	31.39	13.24	yes	yes
County level	Society (Y2)	Z_{24} Number of health workers per ten thousand people	The number of county health workers/total county population	14.06	50.57	32.43	13.62	yes	yes
		Z_{25} Ratio of poverty-stricken villages with cement or asphalt roads	Poor villages with open cement or asphalt roads (%)	23.74	100	71.55	21.46	yes	yes
		Z_{26} Ratio of poverty-stricken villages with passenger buses	Poor villages with passenger shuttle buses (%)	13.50	100	58.31	23.47	yes	yes
	Ecology (Y3)	Z_{31} Altitude (m)*	Mean height difference between the county and sea levels	312.47	1,174.61	697.94	249.73	yes	yes
		Z_{32} Vegetation coverage	Vegetation area/total land area (%)	25.36	79.22	56.48	20.06	yes	yes
		Z_{33} Terrain relief (m)*	The highest altitude—the lowest altitude	304.82	1,306.40	733.38	233.72	yes	yes

Table 1 (continued)

Level	Dimension	Index	Paraphrase	Descriptive statistics				C.V (retain? yes or no)	Collinearity detection (retain? yes or no)
				Min	Max	Mean	S.D		
		Z_{34}	Multi-year average Rainfall (mm)	Average of the quotient of the sum of annual rainfall and years	1,020	1,583.29	1,293.67	138.34	no	—
		Z_{35}	Ecological fragility*	Ecosystem sensitivity and resilience	0.51	0.71	0.58	0.05	no	—

An indicator with * indicates a positive factor. The "terrain type" is divided into three types, plains, hills, and mountains, which are assigned values of 1, 2, and 3, respectively

the study area at different scales, the candidate poverty-causing factors at the county scale include three dimensions: economy, society and ecology. The indicators of the economic dimension can directly reflect the overall economic development level and poverty status of each county. The indicators of the social dimension serve to indicate the farmers' living burdens and living conditions. The indicators of the ecological dimension indirectly reflect poverty by influencing sustainable production and the living conditions of farmers.

For the independence of the candidate indicators, this study uses the coefficient of variation method and collinearity detection to screen the village-level and county-level indicators to ensure that the retained indicators are independent. As a result, 12 indicators were retained at the village level, and 11 indicators were retained at the county level (Table 1).

Constructing the HSLM for the Detection of Poverty-Causing Factors

The poverty incidence of a village is affected by both village-level and county-level factors since each village is nested within a county. Therefore, it is necessary to construct a multilevel regression model to detect the significant poverty-related factors and their mechanisms at different levels. Additionally, there may be spatial effects in the study area that may impact the detection of poverty-related factors to a certain degree. Therefore, this study uses the poverty incidence values of poverty-stricken villages as the dependent variable, considers the factors at the village and county levels to be the explanatory variables, and attempts to improve the traditional hierarchical linear model to reduce the impacts of spatial effects and more accurately detect the factors of poverty. This study introduces a spatial weight matrix (W) that measures the distance between villages as the coefficient of the village-level explanatory variable (X) (Solmaria & Paul, 2015). A new explanatory variable (WX) is then obtained that includes the relationship between villages. WX is included in the traditional hierarchical linear model to reduce the interference of spatial effects, thereby obtaining the hierarchical spatial linear model (HSLM).

Referring to the classic HLM modelling process (Dunifon, 2005; Graves, 2011; Jiang et al., 2020), the HSLM modelling process used in this study is divided into a series of steps. First, the null model (Model I), that is, a hierarchical regression model that does not contain any explanatory variables, is constructed to explore whether there is a background effect on the poverty incidence in poverty-stricken villages. Second, if a significant background effect on poverty incidence in the study area is determined to be present, a random effect regression model (Model II) that considers spatial effects is constructed. The village-level explanatory variable (WX) is added to the first level to explore the significant factors at the village level and their action mechanisms. Finally, the full model (Model III) is constructed, in which the explanatory variables at the village level and county level are added to the first level and second level of the model, respectively, to explore all significant factors at the county level and the interaction mechanisms between the two levels.

(1) **Model I- null model: Examining individual and background effects**

First, the null model is constructed, and variance component analysis is used to judge the individual effects at the village level and the background effects at the county level. The *intraclass correlation coefficient* (*ICC*) is the proportion of variance between groups in the total variance. This measurement is introduced to determine the proportion of the difference between the second level (group) and the overall difference associated with the dependent variables. Moreover, the *ICC* is used to determine the differences among counties in their proportions of poverty incidence to the overall difference. More specifically, after implementing a likelihood ratio test to verify the significance of the *ICC*, the *ICC* could be used to verify whether there is a significant background effect at the county level associated with poverty-related factors in the study area. According to Wang et al. (2019), it is necessary to add county-level factors to the model to explore their impacts on poverty incidence in poverty-stricken villages when the *ICC* is greater than 0.059. A larger *ICC* value is associated with a more significant background effect. The null model and the variance proportion formula are shown in Table 2.

(2) Model II- random effect regression model: Identifying significant poverty-causing factors at the village level

A random effect regression model that considers spatial effects was used to detect significant factors at the village level. Significant spatial effects were found in the study area using spatial effect detection. The significant county-level background effect on poverty incidence was verified through the null model. The village-level explanatory variables were then added to the first level in the random effect regression model, and no explanatory variable was added in the second level (Table 3).

The spatial weight matrix (*W*) defines the adjacent relationships between spatial units and determines the degree to which the characteristics of any spatial unit contribute to adjacent spatial units. A binary symmetric matrix is usually used to express the adjacent relationship between *n* spatial units (Dong et al., 2019).

Table 2 Null model and variance proportion formula

Level	Null model	Variance proportion	
Level 1 (village level)	$Y_{ij} = \beta_{0j} + r_{ij}$	Village level	$\rho = \frac{\sigma_{eo}^2}{\sigma_{eo}^2 + \sigma_{uo}^2}$
Level 2 (county level)	$\beta_{0j} = \gamma_{00} + \mu_{0j}$	County level	$ICC = \frac{\sigma_{uo}^2}{\sigma_{eo}^2 + \sigma_{uo}^2}$

1) *i* represents the village and *j* represents the county; β_{0j} is the average value for Y_{ij}; γ_{00} is the average value for β_{0j}; β_{0j} and γ_{00} are fixed effects; r_{ij} is the random effect of Y_{ij} and represents variation at the village level; μ_{0j} is the random effect of β_{0j} and represents variation at the county level; Y_{ij} is the poverty incidence in poverty-stricken villages; X_{ij} is the explanatory variable at the village level; and Z_{1j} is the explanatory variable at the county level. 2) The term represents the differences between villages within a county; the term σ_{uo}^2 represents the differences between counties; and ρ and *ICC* represent the proportion of the variance at the village level and county level, respectively. A larger *ICC* value is associated with a more significant county-level background effect on poverty incidence

Table 3 Random effect regression model

Level	Random effect regression model
Level 1 (village level)	$Y_{ij} = \beta_{0j} + \beta_{1j} WX_{ij} + r_{ij}$
Level 2 (county level)	$\beta_{0j} = \gamma_{00} + \mu_{0j}$ $\beta_{1j} = \gamma_{10} + \mu_{1j}$

1) i and j represent the village and county, respectively; β_{0j} is the average value of Y_{ij}; γ_{00} and γ_{10} are the average values of β_{0j} and β_{1j}, respectively; β_{1j} is the fixed effect that represents the regression slope of WX_{ij}; r_{ij} is the random effect of Y_{ij} that represents the variation at the village level; μ_{0j} and μ_{1j} are the random effects of β_{0j} and β_{1j}, respectively, and represent the variation at the county level; Y_{ij} is the poverty incidence in poverty-stricken villages; and X_{ij} is the explanatory variable at the village level

2) W is the inverse distance spatial weight matrix for poverty-stricken villages (this paper assumes that there is at least one neighbouring village around each poverty-stricken village)

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} & W_{14} \\ W_{21} & W_{22} & W_{23} & W_{24} \\ \dots & \dots & \dots & \dots \\ W_{n1} & W_{n2} & W_{n3} & W_{nn} \end{bmatrix}$$

According to Tobler's first law, the correlation between features is inversely proportional to the distance between them. Therefore, the inverse of the Euclidean distance (W_{ij}) was chosen to reflect the neighbourhood relationships between poverty-stricken villages. The inverse distance spatial weight method means that objects close to each other are more similar than objects far away from each other. In other words, the closer two objects are, the more similar their properties are. In contrast, the farther two objects are from each other, the less similar they are. In this case, the closer neighbouring villages of a target village will have a greater influence on the target village than neighbouring villages that are farther away. W_{ij} is calculated using formula (3):

$$W_{ij} = \frac{1}{d_{ij}^a} \tag{4}$$

where i and j represent two different villages; d is the distance between the two villages; and W_{ij} is equal to W_{ji}.

Intervillage relations have a direct impact on poverty incidence since villages are used as the basic units in this study. The hierarchical model mainly considers the impact of the highest level on the lowest level. Therefore, this study only considered the spatial effects at the village level. The improved random effect regression model is shown in Table 3.

(3) Model III- full model: Revealing comprehensive significant factors of poverty and their interaction mechanisms between the village and county levels

Based on the above random effect regression model (Model II), the county-level explanatory variables were added to the second-level equations of the model to build the full model, as shown in Table 4. The full model can explore

Table 4 Full model

Level	Null model
Level 1 (village level)	$Y_{ij} = \beta_{0j} + \beta_{1j}WX_{ij} + r_{ij}$
Level 2 (county level)	$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \mu_{0j}$ $\beta_{1j} = \gamma_{10} + \gamma_{11}Z_{1j}\mu_{1j}$

Y_{ij} is the poverty incidence of poverty-stricken villages; X_{ij} is the explanatory variable at the village level; Z_{1j} is the explanatory variable at the county level; γ_{01} and γ_{11} are the regression slopes of Z_{1j}; and W is the inverse distance spatial weight matrix of poverty-causing villages. The other parameters are the same as those in Table 3

significant factors at the county level and the interaction mechanisms between the village level and county level.

Additionally, the variance ratio formula can be used to detect the degree of explanation of the poverty incidence at each level. This formula can also explain the influences of county-level factors on the village-level regression coefficients. The calculation of the variance ratio is shown in formula (4):

$$z = \frac{x - y}{x} \tag{5}$$

where z represents the variance ratio; x represents the original variance, which is the variance component present in the random effects before the explanatory variable is added; and y represents the conditional variance, which is the remaining variance component in the random effects after the explanatory variable is added.

Results and Analysis

### Test Results of the Spatial Effects of Poverty-Causing Factors

Table 5 displays the statistics calculated using the global spatial autocorrelation and local autocorrelation methods. *Moran's I* was 0.134, and the p-value was 0.000. These values suggest that there was significant spatial dependence in the study area. *Gi** was 0.001, and the z-value was 4.525. The z-value was much greater than 1.96 and was significant at the 0.001 level. The *Gi** value and the z-value indicate that there was an obvious local spatial aggregation phenomenon present; namely, there was significant spatial heterogeneity in the study area.

The aggregation state of poverty-stricken villages is visualized in Fig. 2. The poverty-stricken villages in the study area were found to present a local aggregation state.

Table 5 Results of the detection of spatial effects

Spatial dependence			Spatial heterogeneity		
Moran's I	*z*-value	*p-value*	*Gi**	*z-value*	*p-value*
0.134	13.486	0.000	0.001	4.525	<0.001

Fig. 2 Clustering effect observed in poverty-stricken villages
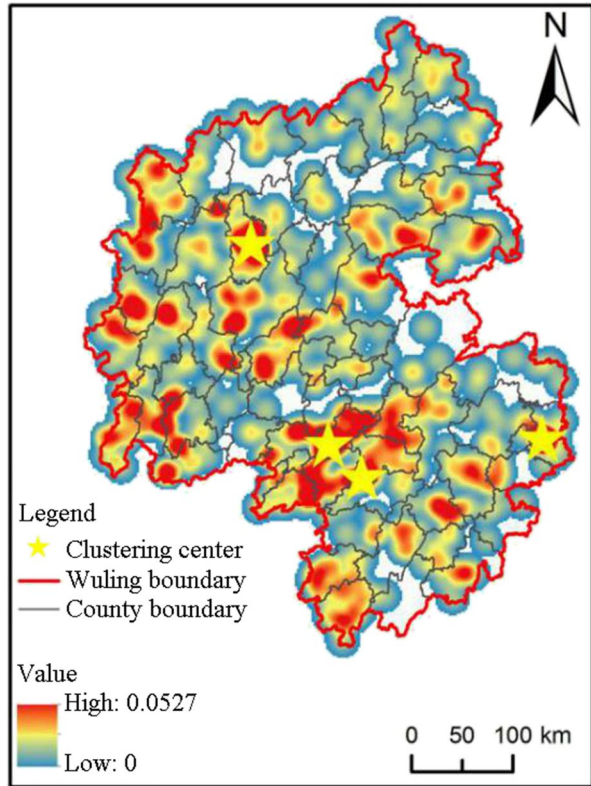
More specifically, there was significant spatial heterogeneity in the study area, and it was identified that spatial factors objectively influenced the distribution of villages with different degrees of economic development. This also means that the impact of each variable on the poor villages in the study area was heterogeneous. Therefore, it is necessary to build a hierarchical spatial linear model to reduce the influences of spatial dependence and spatial heterogeneity on the detection of poverty-related factors.

Multilevel Comprehensive Poverty-Causing Factors in Poverty-Stricken Villages

The poverty-causing factors in the study area were analysed according to the estimation results from the null model, the random effect model, and the full model.

Individual and Background Effects of Poverty-Causing Factors

Before conducting analysis with the null model, we adopted HLM software to implement a likelihood ratio test and verified the significance of the ICC. The degree of freedom (DF) was 46. The chi-square value was 1246.884, and the P-value was 0.000. Therefore, a high ICC value (51.72%) proved the existence of contextual effects. The results estimated by the null model are displayed in Table 6. The

Table 6 Detection results of the null model

Fixed effect		Random effect		Variance ratio	
G_{00}	0.0622***	R	0.0014	Village level (ρ)	48.28%
		U_0	0.0015***	County level (ICC)	51.72%

*p<0.05, **p<0.01,***p<0.001

variation ratio for ρ was 48.28%, which indicates that 48.28% of the overall difference in poverty incidence resulted from differences between villages. In other words, there was a significant individual effect on poverty incidence in the study area. The random effect (U_0) at the county level was found to be significant. The variance ratio for the ICC was 51.72%, which indicates that 51.72% of the overall difference in poverty incidence resulted from differences between counties. There was a significant background effect on poverty incidence in the study area.

From the above, it can be seen that, overall, the poverty incidence of poverty-stricken villages in the study area was affected by factors at the village and county levels. Therefore, it was necessary to construct the hierarchical spatial linear model (HSLM) to detect poverty-related factors due to the significant spatial and background effects associated with the poverty incidence of each village.

Multi-Level Poverty-Causing Factors

(1) Village level

The estimation results of the random effect regression model are shown in Table 7. Among the village-level explanatory variables, five variables had statistical significance. These included the terrain type (X_{12}), per capita cultivated land area (X_{31}), safe drinking water access ratio (X_{35}), labour force ratio (X_{41}), and ratio of

Table 7 Estimated results of village-level variables from the random effect regression model

Explanatory variables	Parameter	Coefficient	Standard error	T-ratio
	$B0$	0.0628	0.0059	10.630***
X_{11} (Distance from the nearest town's bazaar)	$B1$	0.3002	0.1373	2.187*
X_{12} (Terrain type)	$B2$	0.0368	0.0086	4.294***
X_{13} (Frequency of natural disasters)	$B3$	0.0104	0.0153	-0.682
X_{21} (Population density)	$B4$	0.0476	0.0316	1.506
X_{31} (Per capita cultivated land area)	$B5$	-1.2888	0.4352	-2.961**
X_{32} (Road access ratio)	$B6$	-0.0006	0.0119	-0.048
X_{34} (Phone access ratio)	$B7$	-0.0096	0.0058	1.651
X_{35} (Safe drinking water access ratio)	$B8$	-0.0598	0.0181	-3.306**
X_{41} (Labour force ratio)	$B9$	-0.0518	0.0227	-2.279*
X_{42} (Migrant labour force ratio)	$B10$	-0.0048	0.0062	-0.778
X_{43} (Ratio of illiterate labour forces)	$B11$	-0.0059	0.0038	1.560
X_{52} (Ratio of population enrolled in the new rural basic pension insurance)	$B12$	-0.0376	0.0182	-2.071*

population enrolled in the new rural basic pension insurance in the village (X_{52}). The regression coefficients of each variable can be ranked in ascending order as follows: the per capita cultivated land area (-1.2888), safe drinking water access ratio (-0.0598), labour force ratio (-0.0518), ratio of population enrolled in the new rural basic pension insurance in the village (-0.0376), and type of terrain (0.0368).

There was a significantly negative correlation between per capita cultivated land area and poverty incidence. A larger per capita cultivated land area was associated with a lower poverty incidence. This variable reflects the material basis of villager survival, and increasing the per capita cultivated land area is helpful for increasing villager income. The safe drinking water access ratio was found to be negatively correlated with poverty incidence; a higher safe drinking water access ratio is associated with a lower poverty incidence. The proportion of safe drinking water can reflect improvements in social infrastructure. A complete social infrastructure can improve the convenience of access to water resources, which can reduce the poverty incidence. The labour force ratio was found to have a significantly negative correlation with poverty incidence, which suggests that a higher labour force ratio led to a lower poverty incidence. The labour force ratio can reflect the ability of villagers to obtain income. Moreover, a higher labour force ratio for a village is associated with a greater income level in that village. This can help reduce the possibility of villagers falling into poverty, thereby reducing the poverty incidence. There was a significantly negative correlation between the ratio of the population enrolled in the new rural basic pension insurance in each village and the poverty incidence. A higher ratio of the population enrolled in the new rural basic pension insurance was associated with a lower poverty incidence. This reflects the impact of national welfare on the development of poverty-stricken villages. National welfare can help improve quality of life, and higher welfare coverage is related to improved living conditions. There was also a significantly positive correlation between terrain type and poverty incidence, indicating that complex terrain is associated with a higher poverty incidence. Moreover, this type of terrain can affect agricultural output. For example, plateau regions are not conducive to agricultural production and can reduce agricultural output and villager income.

The random effect estimation results from the null model and the random effect regression model are displayed in Table 8. The variance ratio was 14.29% at the village level. This suggests that village-level factors could explain 14.29% of the differences in poverty incidence among villages.

(2) County level

The results of the full model are provided in Table 9. Among the county-level explanatory variables, five variables were found to be statistically significant. These

Table 8 Variances from the null model and random effect regression model at the village level

	Variance component	Standard deviation
Null model	0.0014	0.0375
Random effect regression model	0.0012	0.0345
Variance ratio	14.29%	

included the per capita income (Z_{12}), gross enrolment rate in the first three years (Z_{21}), ratio of poverty-stricken villages with passenger buses (Z_{26}), vegetation coverage (Z_{32}), and terrain relief (Z_{33}). The regression coefficients can be listed in ascending order as follows: the per capita income (-0.2288), gross enrolment rate in the first three years (-0.1398), vegetation coverage (-0.1289), ratio of poverty-stricken villages with passenger buses (-0.0953), and terrain relief (0.1815).

There was a significantly negative correlation between per capita income and poverty incidence, which indicates that higher per capita income was tied to lower poverty incidence. Per capita income can reflect the level of regional economic development. A higher level of economic development can lower the possibility of villagers falling into poverty. There was a significantly positive correlation between terrain relief and poverty incidence, indicating that greater terrain relief was associated with a higher poverty incidence. Terrain relief can reflect the topographic conditions of a landscape. A greater degree of terrain relief is unfavourable to agricultural production and can reduce agricultural output values and villager incomes, thus increasing the possibility of villagers falling into poverty. The gross enrolment rate in the first three years was found to be negatively correlated with poverty incidence, which indicates that a larger gross enrolment rate was associated with a lower poverty incidence. The gross enrolment rate in the first three years can indicate the degree of education in a region. Education should start with young children to ensure higher attendance rates. This is necessary for achieving long-term development in the studied region and for reducing the possibility of children falling into poverty in the future. Vegetation coverage was found to be negatively correlated with poverty incidence, indicating that higher vegetation coverage led to a lower poverty incidence. The amount of vegetation coverage can reflect the ecological environment

Table 9 Estimation results of county-level variables from the full model

Explanatory variables	Parameter	Coefficient	Standard error	T-ratio
	$G00$	0.0634	0.0040	15.916***
Z_{11} (Per capita GDP)	$G01$	-0.0540	0.0593	0.911
Z_{12} (Per capita income)	$G02$	-0.2288	0.0951	-2.407*
Z_{21} (Gross enrolment rate in the first three years)	$G03$	-0.1398	0.0413	3.383**
Z_{22} (Secondary gross enrolment ratio)	$G04$	-0.0545	0.0356	-1.533
Z_{22} (Number of hospital beds per ten thousand people)	$G05$	-0.0031	0.1005	0.031
Z_{24} (Number of health workers per ten thousand people)	$G06$	-0.1018	0.0502	-2.029
Z_{25} (Ratio of poverty-stricken villages with cement or asphalt roads)	$G07$	-0.0162	0.0356	0.455
Z_{26} (Poverty-stricken villages with passenger buses)	$G08$	-0.0953	0.0432	2.208*
Z_{31} (Altitude)	$G09$	0.0799	0.0555	1.441
Z_{32} (Vegetation coverage)	$G10$	-0.1289	0.0327	3.944***
Z_{33} (Terrain relief)	$G11$	0.1815	0.0604	3.006**

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

of a region. High vegetation coverage can reduce wind erosion of soils as well as the frequency of natural disasters and the damage caused by natural disasters. As a result, vegetation coverage can lessen the economic and cultural impacts of natural disasters. There was also a significantly negative correlation found between the ratio of poverty-stricken villages with passenger buses and poverty incidence, indicating that a higher ratio was tied to a lower poverty incidence. This variable reflects the ability of members of a region to communicate with the outside world. Convenient transportation is conducive to regional development. Transportation also promotes local economic development and increases opportunities for villagers to escape from poverty.

The random effect estimation results from the null and full models are displayed in Table 10. The variance ratio at the county level was 40%. This suggests that county-level factors could explain 40% of the differences in poverty incidence among counties.

Interaction Mechanisms of Poverty-Causing Factors Between the Village Level and County Level

The variance estimation results of the random effect regression model are displayed in Table 11. The variances of the regression coefficient of the frequency of natural disasters ($U3$) and the variances of the regression coefficient of the per capita cultivated land area ($U5$) were found to be significant, indicating that the impacts of these two indicators on the poverty incidence varied significantly among different counties. In the second level of the full model, the county-level explanatory variables were added to the regression coefficient ($B3$) equation representing the frequency of natural disasters. The county-level explanatory variables were also added to the regression coefficient ($B5$) equation representing the per capita cultivated land area. These additions were necessary to explain the differences in the impacts of the two factors on the poverty incidence among different counties (Table 12).

As shown in Table 7 B3 (the regression coefficient of X_{13}) was 0.0104, and $B5$ (the regression coefficient of X_{31}) was -1.2888. According to the detection results in Table 12, the regression coefficients for altitude (Z_{31}), vegetation coverage (Z_{32}), and terrain relief (Z_{33}) on B3 at the county level were -0.0047, 0.0162, and -0.2354, respectively. These results suggest that a higher altitude was associated with a weaker impact of the frequency of natural disasters on poverty incidence. Higher vegetation coverage was tied to a stronger impact of the frequency of natural disasters on poverty incidence. A greater degree of terrain relief was associated with a

Table 10 Variances obtained from the null model and the full model at the county level

	Variance component	Standard deviation
Null model	0.0015	0.0384
Full model	0.0009	0.0291
Variance proportion	40%	

Table 11 Random effect estimation results of the random effect regression model and the full model

Random effect regression model				Full model			
Parameter	Variance component	Standard deviation	Chi-square	Parameter	Variance component	Standard deviation	Chi-square
U1	0.2564	0.5064	14.3471	U1	0.1200	0.3464	9.1430
U2	0.0009	0.0307	13.8696	U2	0.0003	0.0180	7.8336
U3	0.0042	0.0647	33.59170**	U3	0.0020	0.0443	20.12638*
U4	0.0114	0.1069	10.2871	U4	0.0082	0.0904	6.1003
U5	4.4268	2.1040	107.97872***	U5	2.2633	1.5044	40.67336***
U6	0.0009	0.0301	10.9398	U6	0.0002	0.0122	7.0654
U7	0.0005	0.0218	16.8680	U7	0.0002	0.0155	11.0760
U8	0.0060	0.0773	7.6443	U8	0.0020	0.0444	3.9468
U9	0.0083	0.0909	19.8051	U9	0.0030	0.0545	12.7871
U10	0.0003	0.0175	5.5430	U10	0.0002	0.0134	3.7605
U11	0.0001	0.0096	5.9641	U11	0.0000	0.0035	4.4054
U12	0.0033	0.0570	19.6020	U12	0.0023	0.0481	12.6552

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 12 Detection results of the interactions between the village level and the county level

Parameter	Coefficient	Standard error	T-ratio	Variance ratio
For X_{13} slope, B3				52.38%
intercept, G30	0.0058	0.0154	0.3790	
Z_{31}, G31	-0.0047	0.2208	-0.0210	
Z_{32}, G32	0.0162	0.0955	0.1690	
Z_{33}, G33	-0.2354	0.2879	-0.8180	
For X_{31} slope, B5				28.07%
intercept, G50	-0.8291	0.3783	-2.1920	
Z_{31}, G51	8.4005	9.6322	0.8720	
Z_{33}, G52	-14.1429	10.5857	-1.3360	

$*p < 0.05$, $**p < 0.01$, $***p < 0.001$

weaker impact of the frequency of natural disasters on the poverty incidence. Both altitude and terrain relief at the county level were observed to weaken the impact of the frequency of natural disasters on poverty incidence at the village level. Moreover, vegetation coverage at the county level increased the impact of the frequency of natural disasters on poverty incidence at the village level. The regression coefficients for altitude (Z_{31}) and terrain relief (Z_{33}) were 8.4005 and -14.1429 at the county level, respectively, indicating that a higher altitude was tied to a weaker impact of the per capita cultivated land area on the poverty incidence. A greater degree of terrain relief resulted in a stronger impact of the per capita cultivated land area on the poverty incidence. Moreover, the altitude at the county level weakened the impact of the per capita cultivated land area on the poverty incidence at the village level. Additionally, terrain relief at the county level increased the impact of the per capita cultivated land area on poverty incidence at the village level.

In addition, the variance estimation results of the random effect regression model and the full model are displayed in Table 12. The altitude (Z_{31}), vegetation coverage (Z_{32}), and terrain relief (Z_{33}) at the county level were found to explain 52.38% of the difference in the impact of the frequency of natural disasters (X_{13}) on the poverty incidence at the village level. Altitude (Z_{31}) and terrain relief (Z_{33}) at the county level could explain 28.07% of the difference in the impact of the per capita cultivated land area (X_{31}) on the poverty incidence at the village level.

Model Comparison

Both the HSLM and HLM are originally driven from an OLS model. That is, all of these models can be used to detect significant poverty-causing factors at either

Table 13 Model comparison

HLM	HSLM
Level 1:$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$	Level 1:$Y_{ij} = \beta_{0j} + \beta_{1j}WX_{ij} + r_{ij}$
Level 2:$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \mu_{0j}$	Level 2:$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \mu_{0j}$
$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_{1j} + \mu_{1j}$	$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_{1j} + \mu_{1j}$

The parameters are the same as those listed in Table 3 and Table 4

Table 14 Comparison of model fitting effects

	Null model	HLM	HSLM
Deviance	-4,296.62 (Reference value)	-4,415.24	-4,468.17
Difference value	—	118.62	171.55

the village level or the county level. However, the two models HSLM and HLM can simultaneously detect multilevel poverty-causing factors and interactions among different levels, while an OLS model can only deal with poverty-causing factors at each level separately. Compared with the HLM, the HSLM considers the impact of inter-village relations on poverty incidence and can reduce the interference of spatial effects. Conversely, the HLM ignores the possible influences of spatial effects on the detection results and modelling accuracy. A qualitative principle comparison of the two models is provided in Table 13.

From the perspective of quantitative comparison, we adopted the deviance value in the model estimation results (Deal et al., 2011) to test the goodness of fit values of the HSLM and HLM. In this study, the deviance of the null model was used as the reference value. The difference between the deviance of the full model and the reference value was then used to compare the HSLM with the HLM (Wang et al., 2019). A larger difference value indicated that the HSLM had a better fit than the HLM. The difference value was calculated using Formula (5):

$$c = |a - b| \qquad (6)$$

where c is the difference value, a is the deviance of the full model, and b is the reference value.

The results of the goodness-of-fit test are shown in Table 14. The difference value between the deviance of the HSLM and the reference value was 171.55, which was greater than the difference value (118.62) between the HLM and the reference value. Overall, the HSLM had the best fit.

Conclusions and Discussions

There are few studies that have considered spatial effects in relation to the multi-level factors that influence poverty. In this study, a hierarchical spatial linear model (HSLM) was designed to reduce the impacts of spatial effects on the multilevel detection of poverty factors. The model revealed whether background and spatial effects are obvious in the study area and examined how to mitigate the impacts of spatial effects on the detection of poverty-related factors while taking background effects into account. It also revealed the significant poverty-related factors at each level in the study area and their interactions between different levels. The performance of the HSLM in detecting poverty-related factors was also evaluated, and the method was found to be effective in detecting multilevel effects.

Based on the nested geographical and administrative features in the data, the proposed model reveals the impacts of multilevel poverty-causing factors on regional village development and investigates the causality of the reported discrepancies. The detection results for the region of the Wuling Mountains revealed the following conclusions. (1) There were significant background and spatial effects in the study area. Moreover, 48.28% of the overall difference in the poverty incidence resulted from individual effects at the village level, and 51.72% resulted from background effects at the county level. There were also significant spatial effects present in the study area. (2) There were different factors leading to poverty at different levels via different mechanisms. The significant factors at the village level included the terrain type, per capita cultivated land area, safe drinking water access ratio, labour force ratio, and ratio of the population enrolled in the new rural basic pension insurance in each village. The significant factors at the county level included the per capita income, gross enrolment rate in the first three years, ratio of poverty-stricken villages with passenger buses, vegetation coverage, and terrain relief. The differences among counties regarding the impacts of the frequency of natural disasters and per capita cultivated land area were affected by factors at the county level. (3) The HSLM was found to produce a better fit than the HLM.

By case testing, we have proven that both individual effects and group effects have significant impacts on the incidence of poverty. Through the random effect regression model and the full model in HSLM modelling, we comprehensively detected different significant factors at both the village level and county level. In addition, we detected that village-level factors could explain 14.29% of the difference in poverty incidence among villages, and county-level factors could explain 40% of the difference in poverty incidence among counties. If we ignored the individual effects, 14.29% of the difference in the poverty incidence among villages would not be explained. Similarly, if we ignored the group effects, 40% of the difference in the poverty incidence among counties would not be explained. Therefore, both the village-level individual effects and the county-level group effects cannot be ignored.

Furthermore, combined with the model results, optimization strategies for improving village-level regional development are proposed for the region of the Wuling Mountains. For example, at the village level, focusing on the geographical environment, production and living conditions, and social security is recommended. (1) Concerning the geographical environment, the focus should be on improving terrain conditions, overcoming terrain disadvantages, and minimizing the impacts of natural disasters. (2) In terms of production and living conditions, cultivated land resources should be protected, and villagers' awareness of arable land protection needs to be strengthened. Additionally, attempts can be made to reclaim wastelands to increase the per capita cultivated land area and agricultural output values. Additionally, infrastructure construction in villages should be strengthened, especially infrastructure related to the provision of safe drinking water. (3) For the development of the labour force, capable personnel should be called to actively participate in work-related projects and strive to change their poverty status through self-reliance. (4) In terms of social security, there is a need to increase the coverage of the new rural basic pension insurance. This can help

to provide security for the elderly and to reduce family burdens caused by inadequate pensions. Additionally, at the county level, it is recommended that economic development, social development, and ecological construction are focused on. (1) Administrative counties should adhere to the concept of sustainable development. Furthermore, measures should be taken to drive local economic development, increase per capita income, and help lead poverty-stricken villages out of poverty. (2) There should be an increase in the enrolment rate of kindergarteners and other school-age children. Children are the future, and strengthening early childhood education is important for the long-term development of society. Additionally, it is necessary to strengthen the construction of transportation infrastructure and to strive for full coverage of passenger buses in administrative villages. There is also a need to construct an "outbound and inbound, village-to-village, bus-to-the-village, safe and convenient" transportation network to increase communication between local communities. (3) There should be a focus on ecological and sustainable development. It is necessary to enhance communities' ability to resist natural disasters. For example, increased vegetation coverage can prevent reductions in cultivated land area caused by soil erosion. In mountainous terrain, it is necessary to either improve terrain conditions or to grow alternate crops such as fruit trees to increase agricultural output as well as improve villager incomes and living standards.

The proposed method offers theoretical and technical support for village-level regional development, and the empirical findings provide scientific policy guidance and technological support for more precise targeting of poor villages and for formulating poverty alleviation measures at both the village level and county level. However, this study does have some limitations. For example, we only detected poverty factors for one year with cross-sectional data and have not yet considered changes in the time series as they may affect the mechanisms of poverty-causing factors. Therefore, the next step is to improve the existing model for the detection of poverty-related factors within a given time series using panel data. Additionally, the HSLM needs to be further improved with regards to its efficiency and scientificity, which will also be considered in our future research.

Declarations

Informed Consent The authors already know and agree to publish the paper.

Conflict of Interest The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Alkire, S., & Fang, Y. F. (2019). Dynamics of multidimensional poverty and uni-dimensional income poverty: An evidence of stability analysis from China. *Social Indicators Research, 142*(1), 25–64.

Aristondo, O. (2018). Poverty decomposition in incidence, intensity and inequality. A review. *Hacienda Publica Espanola-Review of Public Economics, 225,* 109–130.

Behruz, M. Y., Mehdi, C., & Zahra, Y. (2014). Analysis of the factors affecting the spatial distribution of poverty in rural areas, by emphasizing on the economic-social characteristics, case study: Mahmoudabad Village, Shahin Dej Town Ship. *Geography and Territorial Spatial Arrangement, 4*(13), 83–95.

Boemi, S. N., & Papadopoulos, A. M. (2019). Monitoring energy poverty in Northern Greece: the energy poverty phenomenon. *International Journal of Sustainable Energy, 38*(1), 74–88.

Buonanno, P., Pasini, G., & Vanin, P. (2012). Crime and social sanction. *Papers in Regional Science, 91*(1), 193–218.

Cao, S. S., Wang, Y. H., Duan, F. Z., et al. (2016). Coupling between ecological vulnerability and economic poverty in contiguous destitute areas, China: Empirical analysis of 714 poverty-stricken counties in contiguous destitute areas. *Chinese Journal of Applied Ecology, 27*(8), 2614–2622.

Carneiro, D. M., Bagolin, I. P., & Tai, S. H. T. (2016). Poverty determinants in Brazilian Metropolitan Areas from 1995 to 2009. *Nova Economia, 26*(1), 69–96.

Chen, K. M., & Wang, T. M. (2015). Determinants of poverty status in Taiwan: A multilevel approach. *Social Indicators Research, 123*(2), 371–389.

Das, M., & Ghosh, S. K. (2017). Measuring moran's I in a cost-efficient manner to describe a land-cover change pattern in large-scale remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 10*(6), 2631–2639.

Deal, K. H. B., Susanne, M., et al. (2011). Effects of field instructor training on student competencies and the supervisory alliance. *Research on Social Work Practice, 21*(6), 712–726.

Dong, G. P., Levi, W., Alekos, A., & Dani, A. (2019). Inferring neighborhood quality with property transaction records by using a locally adaptive spatial multi-level model. *Computers Environment and Urban Systems, 73*, 118–125.

Dunford, M., Gao, B., Li, W. (2019). Who, where and why? Characterizing China's rural population and residual rural poverty. Area Development & Policy, 1–30.

Dunifon, R. (2005). The labor-market effects of an anti-poverty program: Results from hierarchical linear modeling. *Social Science Research., 34*(1), 1–19.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by distance statistics. *Geographical Analysis, 24,* 189–206.

Goodchild, M. F., Anselin, L., Appelbaum, R. P., et al. (2000). Toward spatially integrated social science. *International Regional Science Review, 23*(2), 139–159.

Graves, S. (2011). Hierarchical linear modeling, Encyclopedia of child behavior and development. Springer US, 744–745.

Guo, Y. Q., Chang, S. S., Sha, F., et al. (2018). Poverty concentration in an affluent city: Geographic variation and correlates of neighborhood poverty rates in Hong Kong. *PLoS ONE, 13*(2), e0190566.

Ibrahim, I., Baiquni, M., Ritohardoyo, S., et al. (2016). Analysis of the factors affecting the poverty in rural areas around gold mine areas in West Sumbawa Regency. *Journal of Degraded and Mining Lands Management, 3*(3), 585–594.

Jiang, Y., Huang, C., Yin, D., et al. (2020). Constructing HLM to examine multi-level poverty-contributing factors of farmer households: Why and how? *PLoS ONE, 15*(1), e0228032.

Kim, H. (2019). Beyond monetary poverty analysis: The dynamics of multidimensional child poverty in developing countries. *Social Indicators Research, 141*(3), 1107–1136.

Kim, R., Mohanty, S. K., & Subramanian, S. V. (2016). Multilevel geographies of poverty in India. *World Development, 87,* 349–359.

Kwadzo, M. (2015). Choosing concepts and measurements of poverty: A comparison of three major poverty approaches. *Journal of Poverty, 19*(4), 409–423.

Liang, C. X., Wang, Y. H., Xu, H. T., et al. (2019). Analyzing spatial distribution of poor villages and their poverty contributing factors: A case study from Wumeng Mountain Area. *Geographical Research, 38*(6), 1389–1402.

Liu, D. D., Zhao, Q., Guo, S. L., et al. (2019). Variability of spatial patterns of autocorrelation and heterogeneity embedded in precipitation. *Hydrology Research, 50*(1), 215–230.

Ma, Z. B., Chen, X. P., & Chen, H. (2018). Multi-scale spatial patterns and influencing factors of rural poverty: A case study in the Liupan Mountain Region, Gansu Province. *China. Chinese Geographical Science, 28*(2), 296–312.

Michalek, A., & Madajova, M. S. (2019). Identifying regional poverty types in Slovakia. *GeoJournal, 84*(1), 85–99.

Mowat, J. G. (2019). Exploring the impact of social inequality and poverty on the mental health and wellbeing and attainment of children and young people in Scotland. *Improving Schools, 22*(3), 203–223.

Odhiambo, F. O. (2019). Assessing the predictors of lived poverty in Kenya: A secondary analysis of the afro barometer survey 2016. *Journal of Asian and African Studies, 54*(3), 452–464.

Park, E. Y., & Nam, S. J. (2018). Influential factors of poverty dynamics among Korean households that include the aged with disability. *Applied Research in Quality of Life, 13*(2), 317–331.

Peirovedin, M. R., Mahdavi, M., & Ziyari, Y. (2016). An analysis of effective factors on spatial distribution of poverty in rural regions of hamedan province. *International Journal of Geography & Geology, 5*(5), 86–96.

Ren, Z. P., Ge, Y., Wang, J. F., et al. (2017). Understanding the inconsistent relationships between socioeconomic factors and poverty incidence across contiguous poverty-stricken regions in China: Multilevel modelling. *Spatial Statistics, 21,* 406–420.

Skare, M., Prziklas Druzeta, R., & Skare, D. (2018). Measuring poverty cycles in the US 1959–2013. *Technological and Economic Development of Economy, 24*(4), 1737–1754.

Seo, S. N. (2016). Microbehavioral econometric methods: Theories, models, and applications for the study of environmental and natural resources. Academic Press, pp. 69–115.

Solmaria, H. V., & Paul, E. J. (2015). *Journal of Regional Science, 55*(3), 339–363.

Su, S. L., Li, L., & Weng, M. (2019). Spatial data analysis, (1st ed., pp 106–115). *Beijing: Science Press*.

Thongdara, R, Samarakoon, L, Shrestha, RP, et al. (2012) Using GIS and Spatial Statistics to Target Poverty and Improve Poverty Alleviation Programs: A Case Study in Northeast Thailand. Applied Spatial Analysis and Policy, 5(2):157–182.

Wang, Y. H., & Chen, Y. F. (2017). A PPI-MVM model for identifying poverty-stricken villages: A Case study from Qianjiang District in Chongqing. *China. Social Indicators Research, 130*(2), 497–522.

Wang, Y. H., Chen, Y. F., Chi, Y., et al. (2018). Village-level multidimensional poverty measurement in China: Where and how. *Journal of Geographical Sciences, 28*(10), 1444–1466.

Wang, Y. H., Liang, C. X., & Li, J. C. (2019). Detecting village-level regional development differences: A GIS and HLM method. *Growth and Change, 50,* 222–246.

Wang, Y.H., & Qian, L. Y. (2017). A PPI-MVM model for identifying poverty-stricken villages: A case study from Qianjiang District in Chongqing, China. *Social Indicators Research, 130*(2), 497–522.

Ward, P. S. (2016). Transient poverty, poverty dynamics, and vulnerability to poverty: An empirical analysis using a balanced panel from Rural China. *World Development., 78,* 541–553.

Authors and Affiliations

Yanhui Wang[1,2,3] · **Yuewen Jiang**[1,2,3] · **Duoduo Yin**[1,2,3] · **Chenxia Liang**[1,2,3] ·
Fuzhou Duan[1,2,3] (iD)

Yanhui Wang
huiwangyan@sohu.com

Yuewen Jiang
rerereewe@qq.com

Duoduo Yin
504141907@qq.com

Chenxia Liang
cnulcx@126.com

[1]   3D Information Collection and Application Key Lab of Education Ministry, Capital Normal
      University, Beijing 100084, China

[2]   Beijing State Key Laboratory Incubation Base of Urban Environmental Processes and Digital
      Simulation, Capital Normal University, Beijing 100048, China

[3]   Engineering Research Center of Sapatial Information Technology, MOE, Capital Normal
      University, Beijing 100048, China