



OPEN ACCESS

EDITED BY

Wei Wang,
Capital Medical University, China

REVIEWED BY

Jianping Wang,
Sichuan University, China
Xiaofeng Zhang,
Capital Medical University, China

*CORRESPONDENCE

Mingyang Yu
✉ 2296991140@qq.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 10 July 2025

ACCEPTED 22 September 2025

PUBLISHED 23 October 2025

CITATION

Hu K, Yang X, Ou S, Li C, Zhang X, Xiao D and Yu M (2025) A multi-method spatial analysis of dysentery incidence determinants across Chinese provinces.
Front. Public Health 13:1663473.
doi: 10.3389/fpubh.2025.1663473

COPYRIGHT

© 2025 Hu, Yang, Ou, Li, Zhang, Xiao and Yu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A multi-method spatial analysis of dysentery incidence determinants across Chinese provinces

Ke Hu^{1†}, Xingjin Yang^{2†}, Shuiping Ou^{3†}, Chaojie Li⁴, Xing Zhang⁵, Di Xiao⁶ and Mingyang Yu^{7*}

¹Department of Scientific Research and Teaching, Xiamen Haicang Hospital, Xiamen, Fujian, China, ²QianDongNanZhou Center for Disease Control and Prevention, QianDongNanZhou, Guizhou, China, ³Honwing Pharma(Guizhou) Company Limited, QianDongNanZhou, Guizhou, China, ⁴Xingtai Center for Disease Control and Prevention, Xingtai, Hebei, China, ⁵Nanjing Lishui Dongping Street Health Center, Nanjing, Jiangsu, China, ⁶Community Health Service Center of Jiuxian Tongliang District, Chongqing, China, ⁷Hypertension Prevention and Control Center, Fuwai Central China Cardiovascular Hospital, Zhengzhou, Henan, China

Introduction: Dysentery remains a significant notifiable Class B infectious disease in China, exhibiting distinct spatial variations in incidence patterns. This persistent geographical heterogeneity necessitates a systematic investigation into the underlying influencing factors to inform targeted prevention and control strategies.

Methods: Our analytical approach incorporated Moran's I index for spatial autocorrelation analysis, multiple linear regression (MLR) for preliminary assessment, and advanced spatial regression models including spatial error model (SEM), geographically weighted regression (GWR), and multiscale geographically weighted regression (MGWR). The analysis incorporated socioeconomic, educational, healthcare, and demographic factors within a unified spatial framework.

Results: The analysis revealed three key findings: (1) Significant spatial clustering of dysentery incidence with identified high-risk concentration in the Beijing-Tianjin region; (2) Superior performance of MGWR modeling in capturing spatial heterogeneity compared to conventional methods; (3) Distinct regional variations in dominant factors, with economic development most influential in western China, education factors predominant in northeastern areas, and healthcare resource availability showing strongest impact in the northeast but minimal effect in southern regions.

Conclusions: The study demonstrates the value of multiscale spatial analysis in understanding geographical disease patterns, revealing that dysentery incidence in China is governed by different factors across regions.

KEYWORDS

dysentery incidence, influencing factors, multiple linear regression, spatial error model, geographically weighted regression, multiscale geographically weighted regression

1 Introduction

As a notifiable Class B infectious disease in China, dysentery shows pronounced spatial aggregation and geographical disparities in incidence rates. Nationwide epidemiological surveillance data reveal pronounced regional aggregation, with persistently high-incidence areas centered on the Beijing-Tianjin-Hebei urban cluster (1, 2), while provinces like

Hunan and Zhejiang in the Yangtze River basin also demonstrate elevated disease risk (3, 4). This spatial heterogeneity manifests more precisely at finer scales, evidenced by significant local hotspots such as the Hangzhou Bay urban belt in Zhejiang (4) and mountainous areas of western and southern Hunan (3). Such complex spatial patterns not only reflect the multidimensional transmission mechanisms but also underscore the necessity for investigating underlying determinants.

Existing studies have revealed multidimensional driving factors influencing the spatial distribution of dysentery. At the socioeconomic level, per capita GDP demonstrates significant explanatory power (5), likely attributable to improved sanitation infrastructure and enhanced healthcare accessibility resulting from economic development. The impact of educational level is particularly prominent. Studies have shown that individuals with lower educational attainment face a significantly higher risk of infection (6, 7). This association is even more pronounced among children (8), reflecting the critical role of health literacy in disease prevention.

The uneven distribution of medical resources also exerts considerable influence (5). However, population density has a dual effect: it can potentially facilitate transmission by increasing the frequency of contact (1, 2), while simultaneously generating protective effects through enhanced public health investments (9). Additionally, larger household size elevates risk by increasing close contact opportunities (10, 11), and sanitation disadvantages in low-income areas (1) coupled with environmental pressures from rapid urbanization (9) have been conclusively linked to disease prevalence. These associations are particularly evident in highly urbanized regions like the Beijing-Tianjin-Hebei area (1).

Methodologically, spatial econometrics offers a suite of innovative analytical techniques. While Multiple Linear Regression (MLR) identifies global risk factors, it fails to account for spatial autocorrelation (12, 13). Spatial lag models (SLM) incorporate spatial weight matrices to capture spillover effects from neighboring areas (14, 15), while spatial error models (SEM) address spatial dependence in the error terms (16, 17). Geographically weighted regression (GWR) advances the field by relaxing spatial stationarity assumptions to reveal local spatial heterogeneity (18, 19). The more recent Multiscale Geographically Weighted Regression (MGWR), which allows variable-specific bandwidths, has demonstrated superior performance in provincial-scale analyses (20, 21), by simultaneously identifying both broad-scale and localized determinants (22, 23).

It is precisely to address the complex and multiscale nature of dysentery's spatial drivers that this study develops a multidimensional approach. Building upon this methodological foundation, we construct an integrated spatial analytical framework that systematically combines MLR, SEM, SLM, GWR, and MGWR. This comprehensive strategy is designed to disentangle the intricate spatial distribution patterns of dysentery incidence across China. By integrating global, local, and multiscale perspectives, the framework enables a more robust examination of the underlying driving mechanisms at the provincial level, thereby offering deeper insights for targeted public health interventions.

TABLE 1 Key factors selected for analysis.

Variable	Factors
Economic development	GDP per capita
Education level	Number of college students per 100,000 persons
Healthcare resources	Number of hospital beds per 1,000 persons
Population structure	Population density
Family structure	Average household size
Consumption level	Household consumption expenditure per capita
Urban-rural structure	Urbanization rate

2 Methods

2.1 Data

Building upon the theoretical framework established in the background section and considering data availability, this study selected seven core influencing factor indicators (detailed in Table 1) covering multiple dimensions including economic development, education level, population structure, family structure, healthcare resources, consumption level, and urban-rural structure. The analysis units comprised 31 provincial-level administrative regions (including provinces, autonomous regions, and municipalities directly under the central government) of mainland China. All data were obtained from authoritative statistical sources: dysentery incidence data in 2022 were extracted from the China Health Statistical Yearbook, while other explanatory variables were sourced from the officially published China Statistical Yearbook.

2.2 Spatial distribution characteristics analysis

The study first employed Geographic Information System spatial visualization techniques to construct a thematic map of dysentery incidence across 31 provincial-level administrative regions in mainland China. By transforming provincial incidence data into intuitive spatial distribution maps, the regional variations in disease incidence intensity were clearly presented.

2.3 Spatial autocorrelation analysis

This research adopted spatial autocorrelation analysis methods to systematically evaluate the spatial distribution characteristics of dysentery incidence from the following two dimensions.

2.3.1 Global spatial autocorrelation analysis

The study employed Global Moran's I index to quantitatively assess the spatial autocorrelation of dysentery

incidence rates. The mathematical expression is as follows (Equation 1) (24):

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

Where n represents the sample size, w_{ij} denotes the elements of the spatial weight matrix, x_i and x_j are observed values, and $\bar{x}$ is the mean value. The spatial weight matrix is represented by the elements w_{ij} .

The Moran's I index ranges from -1 to 1 , where positive values ($I > 0$) indicate positive spatial autocorrelation (clustering of similar values), negative values ($I < 0$) represent negative spatial autocorrelation (dissimilar adjacent values), and values approaching zero ($I \approx 0$) suggest random spatial distribution. All results were verified for statistical significance using Z-test ($p < 0.05$) (25).

2.3.2 Local Moran's index (LISA)

The Local Indicators of Spatial Association (LISA) method was applied to examine spatial correlation patterns between provincial administrative regions and their neighboring areas. Using the same variable definitions as the global Moran's I, The computational formula is expressed as (Equation 2) (26):

$$I_i = \frac{n (x_i - \bar{x}) \sum_{j=1}^n W_{ij} (x_j - \bar{x})}{\sum_{j=1}^n (x_j - \bar{x})^2} \quad (2)$$

The LISA analysis identified four characteristic spatial patterns: high-high clusters (HH, indicating adjacent high-incidence areas), low-low clusters (LL, indicating adjacent low-incidence areas), high-low outliers (HL, high-incidence areas surrounded by low-incidence areas), and low-high outliers (LH, low-incidence areas surrounded by high-incidence areas). All identified spatial patterns were statistically significant ($p < 0.05$) and were visualized through spatial mapping to effectively demonstrate the geographical variations in dysentery incidence (27).

2.4 Multiple linear regression (MLR)

The study employed MLR to analyze influencing factors of dysentery incidence, with the basic model formulation (Equation 3) (28):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (3)$$

Where Y represents the incidence rate of dysentery, X_1 through X_p denote the influencing factors, β_0 is the constant term, β_1 through β_p are the variable coefficients, and ε is the random error term, which follows a normal distribution with a mean of 0 and a variance of σ^2 . Model evaluation included R^2 for explanatory power, Akaike Information Criterion (AIC) as well as the log-likelihood value for model comparison, and Variance Inflation Factors (VIF; threshold = 5) for multicollinearity assessment (29).

2.5 Global spatial regression models

2.5.1 Spatial error model (SEM)

The SEM addresses spatial autocorrelation in regression residuals through its dual-component structure. The model specification consists of two components (Equations 4 and 5):

$$Y = X\beta + u \quad (4)$$

$$u = \lambda W_u + \varepsilon \quad (5)$$

Where λ denotes the spatial error autocorrelation coefficient, and W_u represents the spatial lag effect of the error terms. By incorporating spatial dependence in error terms, SEM effectively resolves estimation bias caused by ignored spatial autocorrelation in conventional regression (30).

2.5.2 Spatial lag model (SLM)

The SLM characterizes spatial interaction effects through a spatially lagged dependent variable: its mathematical formulation is expressed as (Equation 6) (31):

$$Y = \rho W_Y + X\beta + \varepsilon \quad (6)$$

Where ρ quantifies the intensity of spatial dependence, W defines spatial relationships via the weight matrix, W_Y represents the weighted average influence from neighboring areas, X is the matrix of explanatory variables, β denotes the coefficient vectors, and ε is the random error term. This model accounts for spatial spillover effects, overcoming limitations of traditional regression that neglect spatial dependence.

2.6 Local spatial regression models

2.6.1 Geographically weighted regression (GWR)

GWR represents an innovative spatial analysis approach that addresses the limitation of spatially invariant parameters in traditional regression models, effectively revealing spatial heterogeneity in variable relationships. The model incorporates spatial location information into the regression equation, allowing coefficients to vary geographically. Its basic formulation is expressed as (Equation 7) (26):

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^m \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (7)$$

Where (u_i, v_i) denotes the spatial coordinates of sample points and $\beta_k(u_i, v_i)$ represents location-specific regression coefficients. The model employs Gaussian kernel functions and distance-decay methods to construct spatial weight matrices, estimates parameters through weighted least squares, and optimizes bandwidth selection using cross-validation (CV) techniques. Compared with conventional methods, GWR provides more precise characterization of spatial variations in explanatory variable effects, offering refined analytical capabilities for spatial data.

2.6.2 Multiscale geographically weighted regression (MGWR)

MGWR, as an important extension of GWR, introduces variable-specific bandwidth parameters (b_{wk}) to differentially characterize the spatial scales of various explanatory variables. Its basic formulation is Equation 8:

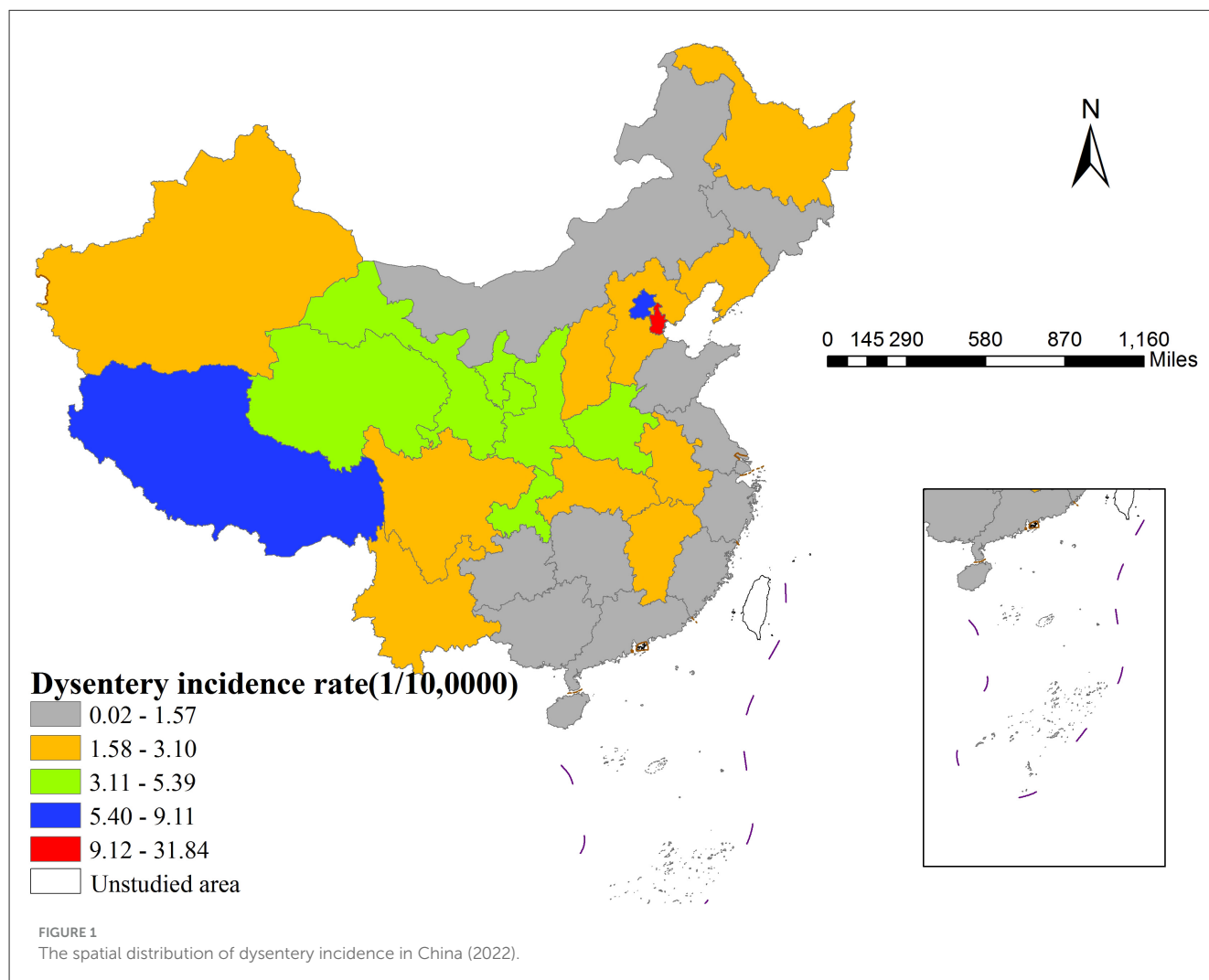
$$Y_i = \beta_0(u_i, v_i, bw_0) + \sum_{k=1}^m \beta_k(u_i, v_i, bw_k) x_{ik} + \epsilon_i \quad (8)$$

Where each explanatory variable x_{ik} corresponds to an independent bandwidth parameter b_{wk} . This design enables more accurate capture of spatial heterogeneity in each variable's influence range. For parameter estimation, MGWR employs an iterative back-fitting algorithm for optimization, determines optimal bandwidth combinations through CV criteria, and constructs spatial weight matrices based on Gaussian kernel functions (32). Compared with conventional GWR, key advantages of MGWR include: (1) effectively addressing parameter estimation bias caused by single bandwidth through variable-specific bandwidth settings; (2) precisely identifying differential spatial influence ranges among explanatory variables; and (3) significantly improving model accuracy in analyzing complex spatial dependence relationships.

These characteristics establish MGWR as a more reliable and precise methodological tool for spatial heterogeneity analysis, particularly suitable for research scenarios where explanatory variables exhibit varying spatial scales of influence.

2.7 Software

This study adopted a multi-platform collaborative approach for spatial statistical analysis. Spatial data processing and visualization were conducted using ArcGIS 10.2 software, encompassing analytical procedures such as spatial autocorrelation testing, and multicollinearity diagnostics. The construction and parameter estimation of conventional regression and spatial econometric models (SLM/SEM) were performed in the GeoDa 1.22 environment, while the analysis of GWR and its multiscale extension (MGWR) was implemented using the specialized MGWR2.2 software. The foundational geographic data were sourced from the National Platform for Common Geospatial Information Services [Map Review No.: GS (2024)0650], with all statistical tests employing two-tailed testing at a significance level of $\alpha = 0.05$.



3 Results

3.1 Descriptive and spatial autocorrelation analysis results

As shown in Figure 1, there was significant spatial heterogeneity in the incidence of dysentery across Chinese provinces in 2022. Geographically, Tianjin recorded the highest incidence rate nationwide at 31.84 per 100,000, while Shanghai had the lowest rate (0.02 per 100,000). Moran’s I index analysis ($I = 0.203$, $p < 0.001$) indicated a significant spatial clustering pattern in

dysentery incidence. LISA cluster analysis (Figure 2) identified specific local patterns, most notably a high-high cluster centered in the Beijing-Tianjin region, demonstrating clear spatial aggregation and geographical disparities in disease distribution.

3.2 Results of MLR

Results In the preliminary analysis phase, this study employed VIF to diagnose multicollinearity among explanatory variables. The

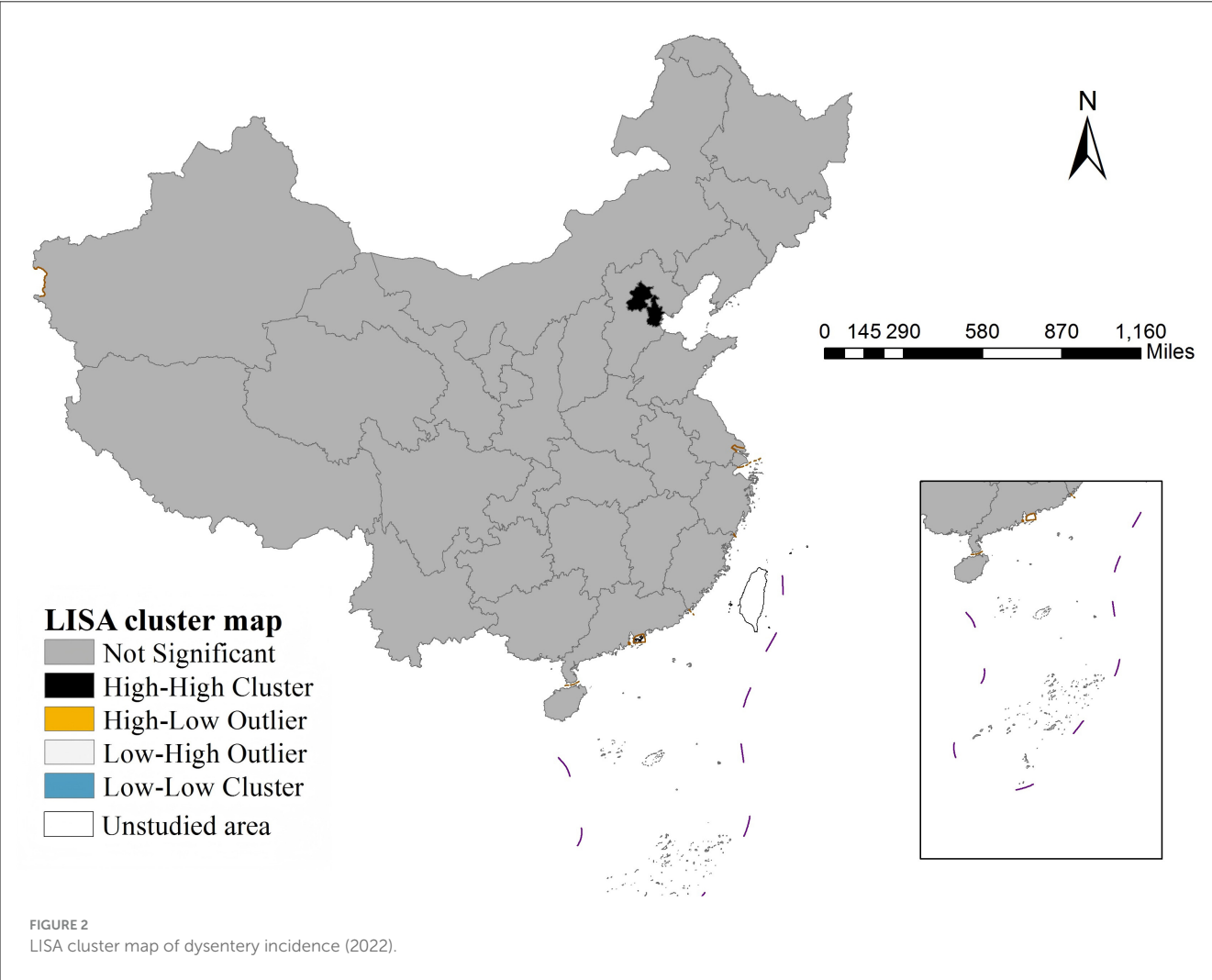


TABLE 2 Results of the SEM and MLR.

Model fit	SEM		MLR	
	Coefficient	p-Value	Coefficient	p-Value
Intercept	−0.0387	0.8691	−0.0000	1.0000
GDP per capita	−0.3917	0.0961	−0.2593	0.3999
Number of college students per 100,000 persons	0.4164	0.0185	0.4584	0.0324
Number of hospital beds per 1,000 persons	−0.5395	0.0022	−0.4742	0.0288
Population density	0.0780	0.6888	0.0262	0.9183
Average household size	−0.2117	0.3138	−0.0558	0.8030

results revealed significant multicollinearity between household consumption expenditure per capita, urbanization rate, and other variables ($VIF > 9$). After excluding these two variables, the VIF values of the remaining variables all fell within an acceptable range ($VIF < 4$).

Using the optimized variable set, MLR analysis identified only two statistically significant predictors: the number of college students per 100,000 persons and the number of hospital beds per 1,000 persons (Table 2). Notably, spatial autocorrelation testing of the model residuals reached a significant level ($p = 0.01$), providing a theoretical basis

for subsequent adoption of spatial regression models (Table 3).

3.3 Results of global spatial regression models

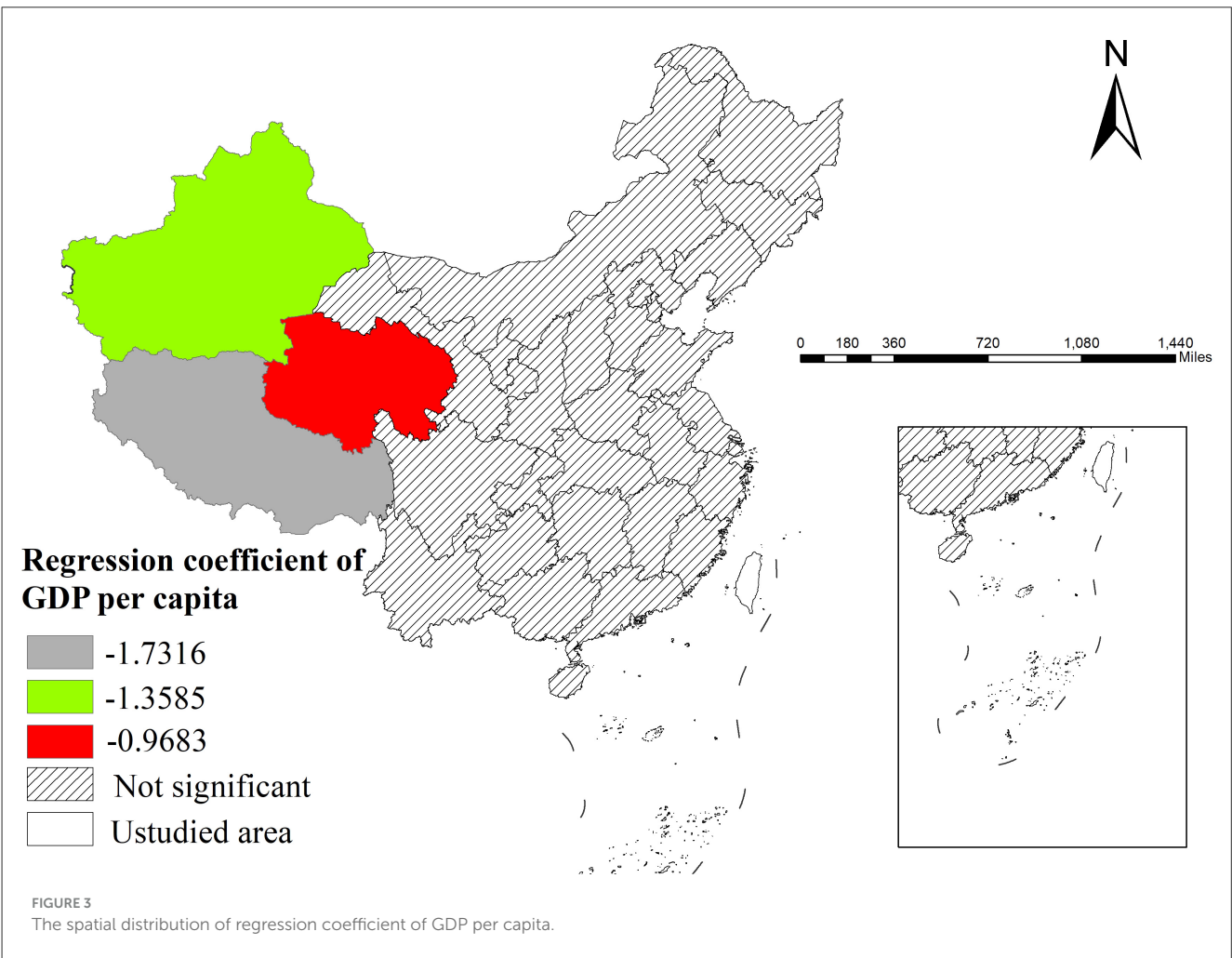
This study constructed SLM and SEM for analysis. The results showed that the spatial error coefficient ($\lambda = 0.428, p = 0.035$) was more significant than the spatial lag coefficient ($W_Y = 0.34, p = 0.099$), indicating that the SEM was more suitable for this study.

After controlling for spatial effects, both the number of college students per 100,000 persons and the number of hospital beds per 1,000 persons remained statistically significant ($p < 0.05$), with their significance further enhanced compared to MLR (Table 2). Model comparisons revealed that the SEM exhibited improved R^2 , higher log-likelihood values, and lower AIC, confirming its superiority (Table 3).

Residual tests indicated that the SEM effectively eliminated spatial autocorrelation, addressing the spatial dependence bias inherent in traditional regression models and ensuring the reliability of the research findings (Table 3).

TABLE 3 Goodness-of-fit of the four models.

Model fit	MGWR	GWR	SEM	MLR
Moran's I	−0.072 ($p = 0.705$)	0.189 ($p = 0.035$)	−0.024 ($p = 0.927$)	0.190 ($p = 0.01$)
R^2	0.584	0.318	0.392	0.305
AIC	79.874	90.317	84.959	87.682
Log likelihood	−30.391	−38.066	−36.480	−37.841



3.4 Results of local spatial regression models

This study systematically compared model fitting performance by constructing both GWR and MGWR. The results demonstrated that the MGWR model exhibited superior explanatory power regarding the relationship between dysentery incidence and influencing factors, with significantly better fitting performance than both GWR and SEM. Notably, the GWR model not only underperformed SEM in fitting effectiveness but also showed significant residual spatial autocorrelation (Table 3). Consequently, this study ultimately adopted the MGWR model with its multiscale analytical advantages for in-depth analysis.

The MGWR model results revealed that among the five core explanatory variables, three demonstrated statistically significant impacts on dysentery incidence ($p < 0.05$) with distinct spatial heterogeneity characteristics.

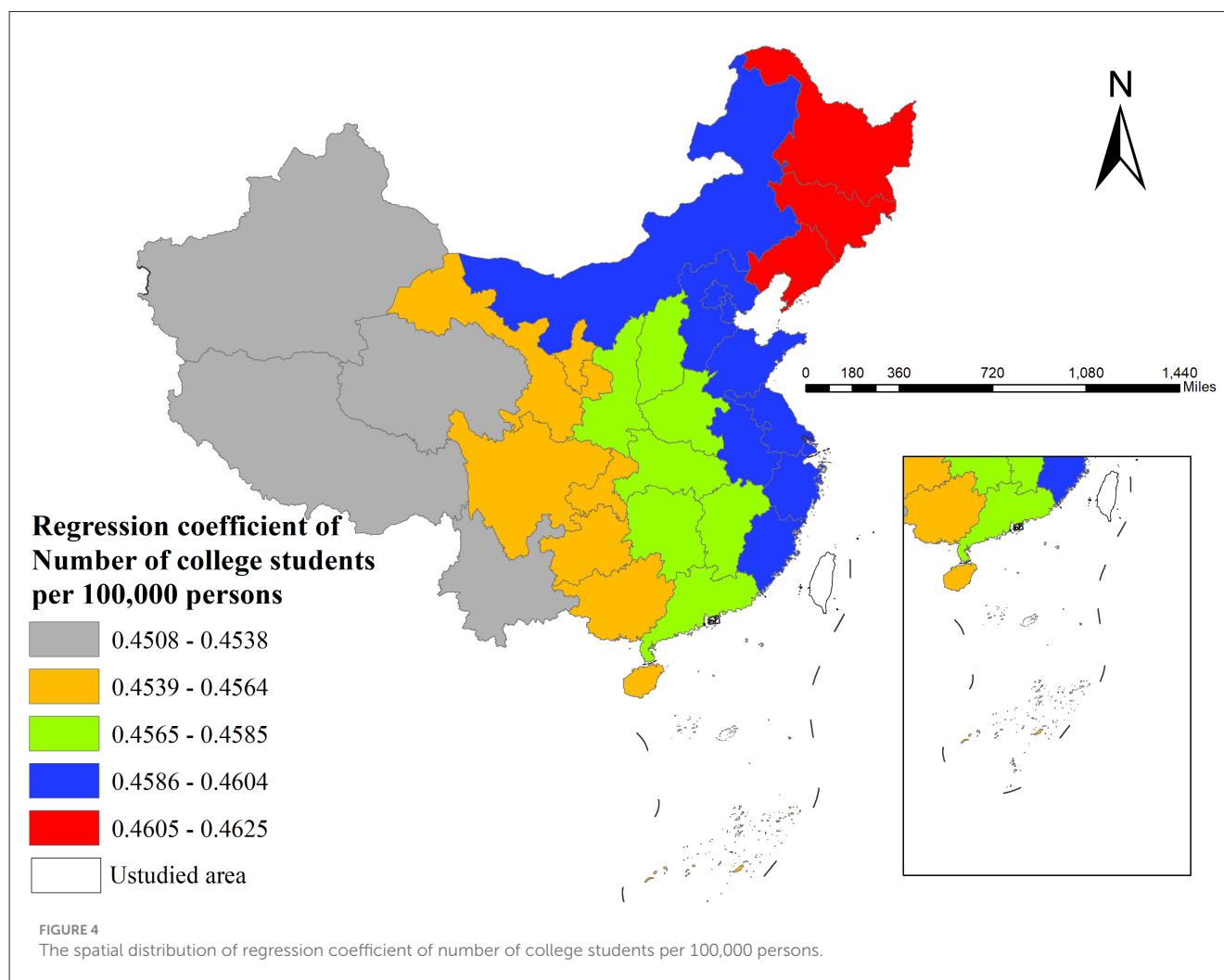
Firstly, per capita GDP showed significant negative correlation with incidence only in western China, while its effect was nonsignificant in central and eastern regions (Figure 3).

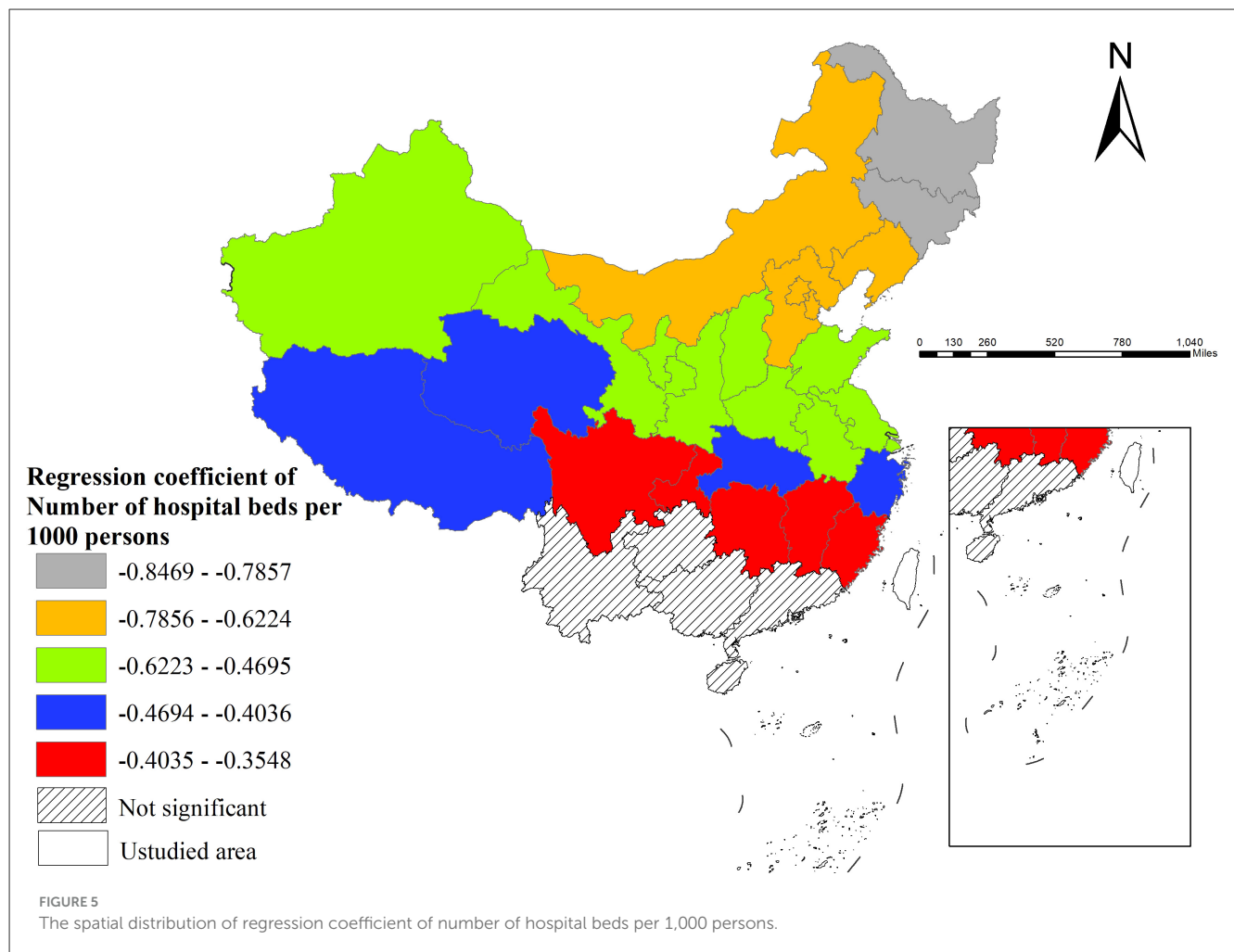
Secondly, the number of college students per 100,000 persons exhibited positive effects, most pronounced in eastern areas (particularly in northeastern provinces) and relatively weaker in western regions (especially Xinjiang and Tibet; Figure 4).

Thirdly, regarding healthcare resources, hospital beds per 1,000 persons demonstrated significant negative effects, particularly prominent in northeastern China, while southern regions (e.g., Guangdong, Guangxi, Yunnan, Guizhou) showed no statistically significant impact (Figure 5).

However, neither population density nor average household size showed statistically significant effects on dysentery incidence across all regions.

Further spatial analysis indicated superior model fit in eastern regions (e.g., northeastern provinces, Beijing, Tianjin), where local R^2 values were generally higher, suggesting strong explanatory power of selected variables for incidence variation. However, the model showed relatively poorer performance in western provinces (e.g., Yunnan, Tibet), implying potential needs for incorporating additional influencing factors to enhance explanatory capability in these areas (Figure 6).





3.5 Sensitivity analysis

The sensitivity analysis revealed that bandwidth selection critically influenced MGWR model outcomes. Comparative evaluation of MGWR models constructed under AIC_c and CV criteria (Table 4) showed that while the AIC_c -based model achieved slightly better goodness-of-fit, its residuals still exhibited significant spatial autocorrelation. Therefore, the CV-based MGWR model demonstrated better robustness and proved more suitable for this study's data analysis.

4 Discussion

This study established a comprehensive analytical framework for dysentery incidence using four methodological approaches: MLR, SEM, GWR, and MGWR. The MLR model first identified key influencing factors; the SEM model confirmed the significant role of spatial dependence; the GWR model revealed the spatial heterogeneity characteristics of influencing factors; and finally, the MGWR model further enhanced the model's explanatory power and predictive accuracy through multiscale analysis. This sequential analysis

not only elucidated the mechanisms influencing dysentery incidence but also identified regional variations in these factors, providing a scientific basis for targeted prevention and control measures.

Empirical results from the MGWR analysis GDP per capita demonstrated a significant negative effect on dysentery incidence in western China. This finding aligns with the conclusions of Zhan et al. (33), confirming the positive role of economic development in disease prevention and control. Potential mechanisms include: economically developed regions possess more robust healthcare resource allocation (5); and higher GDP levels facilitate improvements in public health infrastructure (e.g., water supply and sewage systems) (34). Notably, this negative effect is particularly pronounced in western regions with relatively weaker infrastructure, whereas in eastern regions, where baseline sanitary conditions are already better, the marginal effect of economic factors appears more limited.

The study also revealed the number of college students per 100,000 persons exerted a significant positive on disease incidence. Existing research (5, 35) suggests that the communal living patterns of college students (e.g., sharing tableware and sanitation facilities) may increase disease transmission risks. This effect is particularly pronounced in northeast regions with high concentrations of

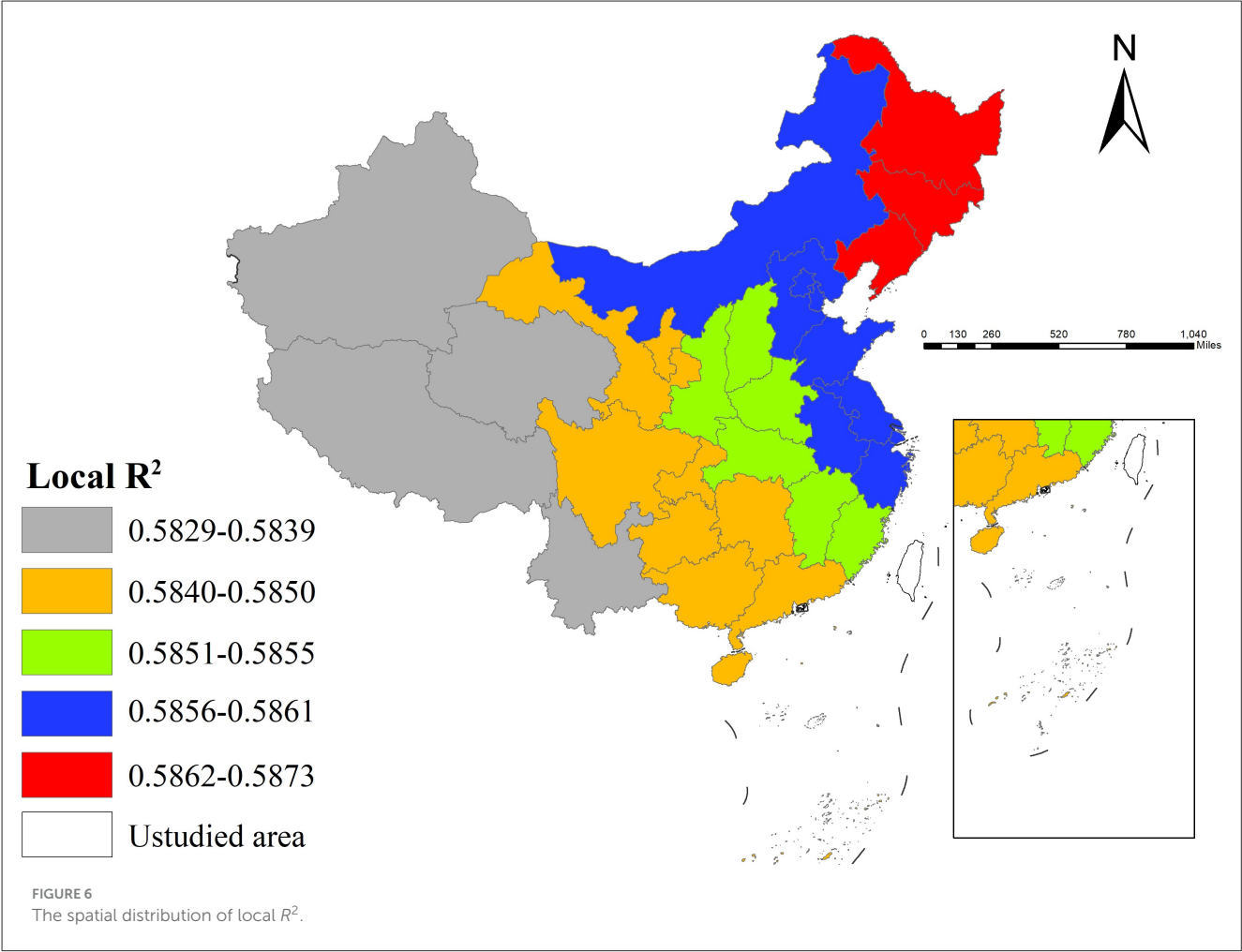


TABLE 4 Comparison of model fit between AIC_c and CV criteria in MGWR.

Model fit	MGWR (CV)	MGWR (AIC _c)
Moran's I	−0.072 ($p = 0.705$)	−0.370 ($p = 0.003$)
R^2	0.584	0.836
AIC	79.874	57.669
Log likelihood	−30.391	−15.957

universities, likely due to higher population density and increased contact frequency.

Hospital beds availability showed a negative effect on dysentery incidence. Increased bed numbers may enhance healthcare accessibility, enabling timely hospitalization of patients and thereby reducing transmission risks (36). The inhibitory effect of hospital beds on dysentery incidence was stronger in northeastern China, which can be attributed to several region-specific factors: relative bed shortages, weaker primary healthcare systems, population outflow alleviating actual medical pressure, and potentially more effective inpatient isolation during cold weather conditions.

Notably, population density showed no significant influence on dysentery incidence. Previous studies (9) suggest this may reflect how robust public health interventions (e.g., vaccination

programs and sanitation infrastructure) in high-density areas effectively counteract the potential risks of population aggregation. Additionally, research (37) has demonstrated that hot and humid climates significantly increase dysentery incidence—such strong climatic effects may overshadow the influence of population density.

Additionally, household size showed no statistically significant effect on dysentery incidence. This finding aligns with previous research (38), potentially because other stronger determinants like sanitation conditions may have overshadowed the role of household size.

Methodologically, this study contributes to the field by innovatively integrating multiple advanced spatial econometric approaches—including MLR, SEM, GWR, and MGWR—to establish a comprehensive analytical framework for identifying determinants of dysentery incidence. This multi-model synergistic approach helps overcome the limitations associated with traditional single-model analyses. By effectively capturing multiscale spatial heterogeneity, the framework provides a useful methodological reference for health geography research and supports the development of regionally tailored public health strategies.

Nevertheless, several limitations should be acknowledged: First, regarding data, the use of provincial-level analysis units due to official statistics availability constraints may insufficiently capture finer-scale (e.g., county/city-level) spatial variations. Meanwhile,

the exclusion of certain important explanatory variables (e.g., environmental factors like temperature and humidity) due to data unavailability might partially compromise the comprehensiveness of findings. Second, methodologically, while the cross-sectional design effectively reveals spatial association patterns, it carries inherent limitations in causal inference. Finally, the models have not accounted for temporal dynamics of key factors like population mobility and prevention policies, which might affect in-depth understanding of dysentery transmission mechanisms.

5 Conclusion

This study systematically investigated the spatial distribution patterns and influencing factors of dysentery incidence in China using multiscale spatial analysis methods. The findings reveal significant spatial heterogeneity in dysentery incidence across China, with a distinct high-incidence cluster identified in the Beijing-Tianjin region. Methodologically, MGWR demonstrated superior analytical performance compared to traditional spatial regression approaches, enabling more precise identification of the spatial variation characteristics of various influencing factors.

Substantial regional differences in dominant factors affecting dysentery incidence. Economic development level showed significant influence in western regions, while educational factors played a particularly prominent role in northeastern China. The impact of healthcare resource allocation also exhibited marked regional variations, with the most pronounced effects observed in northeastern areas. These findings provide novel insights into understanding the spatial patterns of dysentery transmission in China and offer scientific evidence for developing targeted regional prevention and control strategies.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

KH: Conceptualization, Writing – original draft. XY: Conceptualization, Writing – original draft. SO: Conceptualization, Writing – original draft. CL: Methodology, Writing – original draft. XZ: Methodology, Writing – original draft. DX: Validation, Writing – original draft. MY: Supervision, Writing – review & editing.

References

1. Xu C, Li Y, Wang J, Xiao G. Spatial-temporal detection of risk factors for bacillary dysentery in Beijing, Tianjin and Hebei, China. *BMC Public Health*. (2017) 17:743. doi: 10.1186/s12889-017-4762-1
2. Wang L, Xu C, Xiao G, Qiao J, Zhang C. Spatial heterogeneity of bacillary dysentery and the impact of temperature in the Beijing-Tianjin-Hebei region of China. *Int J Biometeorol*. (2021) 65:1919–27. doi: 10.1007/s00484-021-02148-3

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

We thank National Health Commission of the People's Republic of China and National Bureau of Statistics for providing us with data.

Conflict of interest

SO was employed by Honwing Pharma(Guizhou) Company Limited.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2025.1663473/full#supplementary-material>

3. Xu C, Xiao G, Wang J, Zhang X, Liang J. Spatiotemporal risk of bacillary dysentery and sensitivity to meteorological factors in Hunan Province, China. *Int J Environ Res Public Health*. (2017) 15:47. doi: 10.3390/ijerph15010047
4. Yan C, Chen Y, Miao Z, Qin S, Gu H, Cai J. Spatiotemporal characteristics of bacillary dysentery from 2005 to 2017 in Zhejiang Province, China. *Int J Environ Res Public Health*. (2018) 15:1826. doi: 10.3390/ijerph15091826
5. Zhang X, Gu X, Wang L, Zhou Y, Huang Z, Xu C, et al. Spatiotemporal variations in the incidence of bacillary dysentery and long-term effects associated with meteorological and socioeconomic factors in China from 2013 to 2017. *Sci Total Environ*. (2021) 755:142626. doi: 10.1016/j.scitotenv.2020.142626
6. Cociancic P, Zonta ML, Navone GT. A cross-sectional study of intestinal parasitoses in dogs and children of the periurban area of La Plata (Buenos Aires, Argentina): zoonotic importance and implications in public health. *Zoonoses Public Health*. (2018) 65:e44–53. doi: 10.1111/zph.12408
7. Fentahun A, Hailu T, Alemu G. Prevalence of intestinal parasites and *Schistosoma mansoni* and associated factors among fishermen at lake Tana, Northwest Ethiopia. *Biomed Res Int*. (2021) 2021:4534689. doi: 10.1155/2021/4534689
8. Deku JG, Botchway KA, Kinanyok S, Gedzeah CK, Duneeh RV, Duedu KO. Intestinal parasitic infection and associated risk factors among HIV-infected patients seeking healthcare in a rural Hospital in Ghana. *J Pathog*. (2022) 2022:5652637. doi: 10.1155/2022/5652637
9. Gu L, Cai J, Feng Y, Zhan Y, Zhu Z, Liu N, et al. Spatio-temporal pattern and associated factors study on intestinal infectious diseases based on panel model in Zhejiang Province. *BMC Public Health*. (2024) 24:3041. doi: 10.1186/s12889-024-20411-1
10. Bah D, Gebru G, Hakizimana JL, Ogbonna U, Sesay B, Bah B, et al. Prevalence and risk factors of diarrheal diseases in Sierra Leone, 2019: a cross-sectional study. *Pan Afr Med J*. (2022) 41:3. doi: 10.11604/pamj.2022.41.3.32403
11. Shaaban FL, Kabatereine NB, Chami GF. Diarrhoeal outcomes in young children depend on diarrhoeal cases of other household members: a cross-sectional study of 16,025 people in rural Uganda. *BMC Infect Dis*. (2022) 22:484. doi: 10.1186/s12879-022-07468-2
12. Rahman MH, Zafri NM, Ashik FR, Waliullah M, Khan A. Identification of risk factors contributing to COVID-19 incidence rates in Bangladesh: a GIS-based spatial modeling approach. *Heliyon*. (2021) 7:e06260. doi: 10.1016/j.heliyon.2021.e06260
13. Praharaj S, Kaur H, Wentz E. The spatial association of demographic and population health characteristics with COVID-19 prevalence across districts in India. *Geogr Anal*. (2022) 55:427–49. doi: 10.1111/gean.12336
14. Shi B, Fu Y, Bai X, Zhang X, Zheng J, Wang Y, et al. Spatial pattern and spatial heterogeneity of Chinese elite hospitals: a country-level analysis. *Front Public Health*. (2021) 9:710810. doi: 10.3389/fpubh.2021.710810
15. Kianfar N, Mesgari MS. GIS-based spatio-temporal analysis and modeling of COVID-19 incidence rates in Europe. *Spat Spatiotemporal Epidemiol*. (2022) 41:100498. doi: 10.1016/j.sste.2022.100498
16. Nazari Ashani M, Alesheikh AA, Neisani Samani Z, Lotfata A, Bayat S, Alipour S, et al. Socioeconomic and environmental determinants of foot and mouth disease incidence: an ecological, cross-sectional study across Iran using spatial modeling. *Sci Rep*. (2023) 13:13526. doi: 10.1038/s41598-023-40865-4
17. Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci Total Environ*. (2020) 728:138884. doi: 10.1016/j.scitotenv.2020.138884
18. Wu J, Sha W, Zhang P, Wang Z. The spatial non-stationary effect of urban landscape pattern on urban waterlogging: a case study of Shenzhen City. *Sci Rep*. (2020) 10:7369. doi: 10.1038/s41598-020-64113-1
19. Draidí Areed W, Price A, Thompson H, Hassan C, Malseed R, Mengersen K. Bayesian cluster geographically weighted regression for spatial heterogeneous data. *R Soc Open Sci*. (2024) 11:231780. doi: 10.1098/rsos.231780
20. Cao Y, Tian Y, Tian J, Liu K, Wang Y. Impact of built environment on residential online car-hailing trips: based on MGWR model. *PLoS ONE*. (2022) 17:e0277776. doi: 10.1371/journal.pone.0277776
21. Wang C, Solis P, Villa L, Khare N, Wentz EA, Gettel A. Spatial modeling and analysis of heat-related morbidity in Maricopa County, Arizona. *J Urban Health*. (2021) 98:344–61. doi: 10.1007/s11524-021-00520-7
22. Tong Y, Tang L, Xia M, Li G, Hu B, Huang J, et al. Identifying determinants for the seropositive rate of schistosomiasis in Hunan province, China: a multi-scale geographically weighted regression model. *PLoS Negl Trop Dis*. (2023) 17:e0011466. doi: 10.1371/journal.pntd.0011466
23. Zhang M, Tan S, Zhang X. How do varying socio-economic factors affect the scale of land transfer? Evidence from 287 cities in China. *Environ Sci Pollut Res Int*. (2022) 29:40865–77. doi: 10.1007/s11356-021-18126-6
24. Duque AM, Peixoto MV, Lima SVMA, Goes MAO, Santos AD, Araújo KCGM, et al. Analysis of the relationship between life expectancy and social determinants in a north-eastern region of Brazil, 2010–2017. *Geospat Health*. (2018) 13:345–52. doi: 10.4081/gh.2018.702
25. Wang S, Luo K. Life expectancy impacts due to heating energy utilization in China: distribution, relations, and policy implications. *Sci Total Environ*. (2018) 610–611:1047–56. doi: 10.1016/j.scitotenv.2017.08.195
26. Hu K, Zhang X, Yang X, Yu M. A study on the spatial distribution of life expectancy and its air pollution factors in China based on geographically weighted regression. *Front Public Health*. (2025) 13:1565744. doi: 10.3389/fpubh.2025.1565744
27. Sridharan S, Tunstall H, Lawder R, Mitchell R. An exploratory spatial data analysis approach to understanding the relationship between deprivation and mortality in Scotland. *Soc Sci Med*. (2007) 65:1942–52. doi: 10.1016/j.socscimed.2007.05.052
28. Ellsworth SG, van Rossum PSN, Mohan R, Lin SH, Grassberger C, Hobbs B. Declarations of independence: how embedded multicollinearity errors affect dosimetric and other complex analyses in radiation oncology. *Int J Radiat Oncol Biol Phys*. (2023) 117:1054–62. doi: 10.1016/j.ijrobp.2023.06.015
29. Taylor-Haas JA, Garcia MC, Rauh MJ, Peel S, Paterno MV, Bazett-Jones DM, et al. Cadence in youth long-distance runners is predicted by leg length and running speed. *Gait Posture*. (2022) 98:266–70. doi: 10.1016/j.gaitpost.2022.09.085
30. Suerungruang S, Sornlorm K, Laohasiriwong W, Mahato RK. Spatial association and modelling of under-5 mortality in Thailand, 2020. *Geospat Health*. (2023) 18. doi: 10.4081/gh.2023.1220
31. Khedhiri S. A spatiotemporal analysis of the COVID-19 pandemic in North Africa. *Geohealth*. (2022) 6:e2022GH000630. doi: 10.1029/2022GH000630
32. Anderson T, Herrera D, Mireku F, Barner K, Kokkinakis A, Dao H, et al. Geographical variation in social determinants of female breast cancer mortality across US counties. *JAMA Netw Open*. (2023) 6:e2333618. doi: 10.1001/jamanetworkopen.2023.33618
33. Zhan Y, Gu H, Li X. Study on association factors of intestinal infectious diseases based-Bayesian spatio-temporal model. *BMC Infect Dis*. (2023) 23:720. doi: 10.1186/s12879-023-08665-3
34. Zhou W, Gu Y, Wang X. Access to sanitary toilets and health outcomes: a panel data analysis using two-way fixed effects model. *Math Biosci Eng*. (2021) 18:8815–30. doi: 10.3934/mbe.2021435
35. Cha YE, Fu YZ, Yao W. Knowledge, practice of personal hygiene, school sanitation, and risk factors of contracting diarrhea among rural students from five western Provinces in China. *Int J Environ Res Public Health*. (2021) 18:9505. doi: 10.3390/ijerph18189505
36. Wirawan GBS, Januraga PP. Correlation of demographics, healthcare availability, and COVID-19 outcome: Indonesian ecological study. *Front Public Health*. (2021) 9:605290. doi: 10.3389/fpubh.2021.605290
37. Liu J, Wu X, Li C, Xu B, Hu L, Chen J, et al. Identification of weather variables sensitive to dysentery in disease-affected county of China. *Sci Total Environ*. (2017) 575:956–62. doi: 10.1016/j.scitotenv.2016.09.153
38. Bekele D, Merdassa E, Desalegn M, Mosisa G, Turi E. Determinants of diarrhea in under-five children among health extension model and non-model families in Wama Hagelo District, West Ethiopia: community-based comparative cross-sectional study. *J Multidiscip Healthc*. (2021) 14:2803–15. doi: 10.2147/JMDHS.324846