



OPEN

Spatial homogeneity pursuit of regression coefficients for hand, foot and mouth disease in Xinjiang Uygur Autonomous Region in 2018

Xiaoshuang Zhong^{1,2}, Peilin Wang^{1,2} & Huiguo Zhang^{1,2}✉

To explore the complex spatial pattern between the incidence of hand, foot, and mouth disease (HFMD) and meteorological factors [average temperature (AT), average relative humidity (ARH), average air pressure (AP), average wind speed (AW)], this paper constructed a Spatial Clustering coefficient (SCC) regression model to detect spatial clustering patterns of each regression coefficients in different seasons. The results revealed that compared with geographically weighted regression (GWR), the coefficients estimated by SCC method were more smooth with clearly identified spatial and improved edge effects. Therefore, interesting spatial patterns were easy to identify in the SCC estimated coefficients. And then, the SCC method had better estimation accuracy in estimating the relationship between potential meteorological factors and HFMD cases. Meteorological factors had different significance in their effect on HFMD incidence depending on the season. Specifically, the influence of AT on HFMD was negatively correlated in summer and winter, especially in the Altay region, Bayingoleng Mongolian Autonomous Prefecture, Turpan region and Hami region. Second, AW had positive effects with HFMD in summer, but the AW played a negative role in the whole Xinjiang in winter. In Tianshan district, Shayibake district, Shuimogou district, etc. in summer, ARH showed a strong negative correlation, but in Alar city it had a high positive correlation, however, in winter ARH showed a high negative correlation in Altay regions, Aksu region and other places had negative effects, and it showed a strong positive correlation in Shayibak district. Finally, AP had a strong positive correlation with HFMD in summer in Shayibak district, but in winter, AP showed a strong negative correlation in Altay district and Buxel Mongolia Autonomous county. In summary, Xinjiang should adapt measures to local conditions, and formulate appropriate HFMD prevention strategies according to the characteristics of different regions, time, and meteorological factors.

Hand, Foot and Mouth Disease (HFMD) is a viral (enterovirus), infectious and global disease that occurs frequently in children under the age of 5. The Ministry of Health of China designated HFMD as a "C" infectious disease in May 2008. In recent years, it has occurred all over China throughout the year. The threat of HFMD to public health has promoted scholars to study the characteristics and causes of the disease. Many researchers focused on the impact of meteorological factors on HFMD. Studies have shown that meteorological factors such as average temperature and average relative humidity will affect the incidence of HFMD¹⁻⁵. Ma et al. found that the HFMD visit rate in Hong Kong was positively correlated with the average temperature, daily temperature difference, relative humidity, and wind speed⁶. And in Guangzhou and Shenzhen, by time series analysis, they revealed that temperature was positively correlated with the incidence of HFMD^{7,8}. But a Japanese study found that the number of days per week with an average temperature exceeding 25 °C was negatively correlated with the incidence of HFMD⁹. However, the above studies did not consider the spatial differences in the incidence of HFMD. The GWR model, as a spatial statistics method can effectively to model and analyze the heterogeneity of space. Hu et al. used GWR model to investigate the potential factors affecting the incidence of HFMD in children¹⁰. Similarly, Hong et al. used GWR model to discuss the relationship between HFMD and meteorological factors in Inner Mongolia, and they discovered that there was a correlation between these factors and HFMD, and each factor had its unique spatial heterogeneity^{11,12}, as demonstrated by Wang et al. and Bo et al.¹³.

¹College of Mathematics and System Science, Xinjiang University, Urumqi 830017, China. ²These authors contributed equally: Xiaoshuang Zhong, Peilin Wang and Huiguo Zhang. ✉email: zhanghg@xju.edu.cn

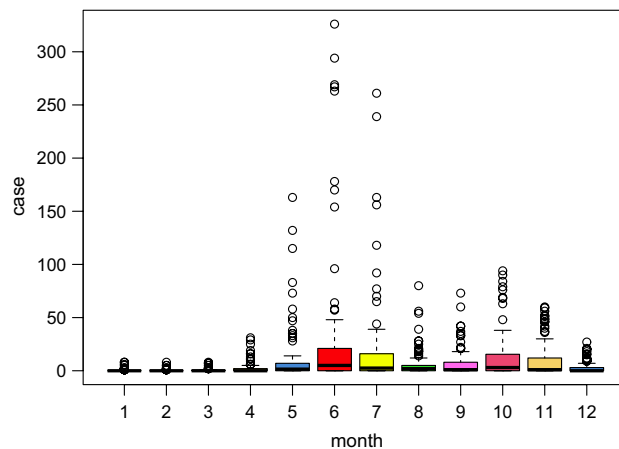


Figure 1. The monthly distribution of HFMD cases in Xinjiang, China, 2018.

It is common knowledge that GWR models tend to estimate regression coefficients as continuous functions, however for data collected from a large area, the relationship between response variables and predictors may exhibit complex spatial dynamic patterns, and the variation of regression coefficients is not necessarily continuous. In particular, relationships between spatial variables may change abruptly at the boundaries of adjacent clusters but remain relatively homogeneous within clusters^{14–16}. Fortunately, the study on the homogeneity pursuit of regression coefficients in high-dimensional data analysis will help to solve this problem. In these studies^{17,18}, pairwise coefficient differences are penalized to encourage homogeneity among coefficients. According to these ideas, Li et al. proposed Spatial Clustering Coefficient (SCC) regression to detect the spatial clustering pattern of regression coefficients, which integrates spatial domain information by constructing appropriate regularization to automatically detect mutation points in regression coefficients, and due to its relatively strong local adaptability, it can also estimate continuously changing regression coefficients¹⁹. Therefore, this paper was to detect and analyze the spatial clustering patterns of AT, MT, ARH, AP, and AW on HFMD occurrence in Xinjiang in 2018 using the SCC model at county-level and monthly-scale.

Results

Statistical analysis. From January 1 to December 31, 2018, the total number of HFMD cases in 100 counties in Xinjiang was 10,260. The monthly distribution of HFMD cases in Xinjiang is shown in Fig. 1. The peak of HFMD occurred in May–July, and the second peak occurred in September–November. Therefore, May–December is chosen as the main research period.

The distribution of HFMD cases in Xinjiang from May 1 to December 31, 2018 is shown in Fig. 2. The results revealed that the incidence of HFMD in northern Xinjiang is higher than that in southern Xinjiang, and the high incidence areas include Changji city, Xinshi district, Shayibak district, Shuimogou district, Midong district, Tianshan district, Shihezi city, Yining city, Jinghe county, Karamay district and Yizhou district.

It can be clearly seen from Fig. 3 that the temperature distribution in Altay region and Tacheng region is similar, and the temperature distribution in Ili Kazakh Autonomous Prefecture, Aksu, Kizilsu Kirgiz Autonomous Prefecture, Kashi region and Hotan region are the same. For ARH, Altay region is a category, Ili Kazakh Autonomous Prefecture is a category, Aksu region and Kizilsu Kirgiz Autonomous Prefecture are clustered into a category, and Kashi region, Hotan region and part of Bayingoleng Mongolian Autonomous Prefecture are a category. The spatial clustering of AP is similar to that of ARH, and AW distribution is similar to that of AT. The specific spatial clustering results are shown in Fig. 3.

Correlation analysis of meteorological factors. Spearman correlation coefficient analysis is used to explore the relationship between the incidence of HFMD (NUM) and meteorological factors in Xinjiang. The detailed results are shown in Table 1. The results reveal that the correlation between AT and MT is 0.968, which is highly correlated. Therefore, meteorological factors other than MT are selected as explanatory variables when establishing the SCC model.

Spatial cluster analysis of regression coefficient. In order to compare the spatial clustering patterns of the regression coefficients in different time periods, this paper mainly selects the data of June and November for analysis, because the statistical analysis shows that the number of HFMD cases in June and November is higher. Next, the partial coefficient estimates obtained by the SCC method are shown in Figs. 4 and 5. At the same time, this paper also shows the coefficient estimates obtained by the GWR method, which is convenient for comparison and verification.

By comparing Figs. 4 and 5, it can be seen that the estimated coefficients $\beta(\mathbf{u}, \mathbf{v})$ vary spatially throughout different seasons. First of all, the incidence of HFMD in southeastern Xinjiang is higher than that in other regions in June 2018. (the first map of SCC in Figs. 4 and 5). Secondly, the AT in June has a negative correlation with HFMD, and the impact categories are mainly divided into three categories, among which it plays an important

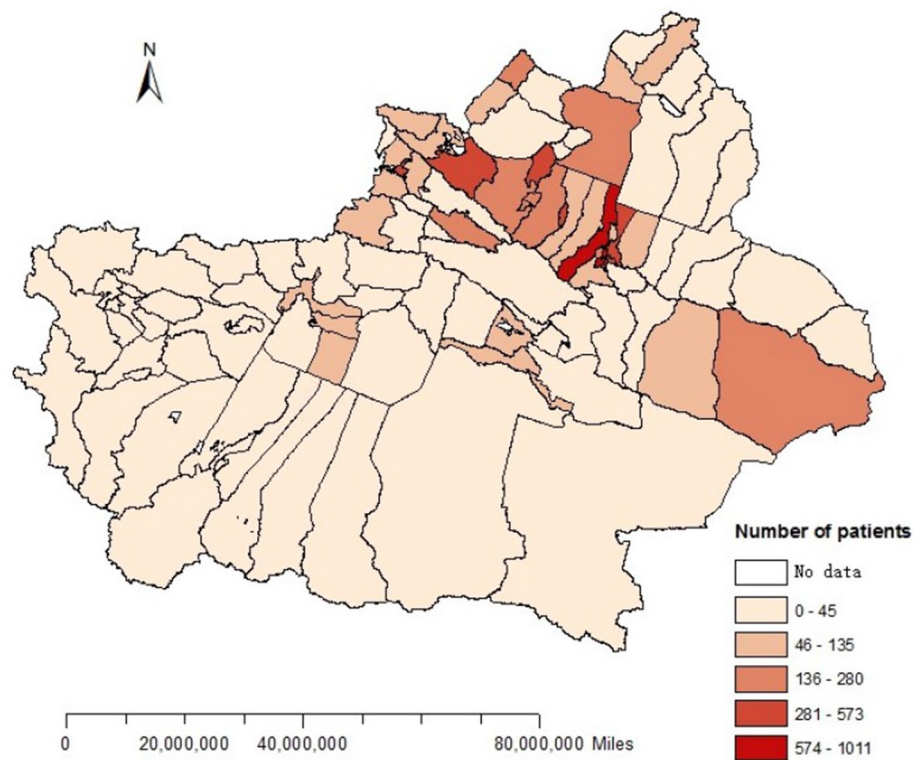


Figure 2. Distribution characteristics of HFMD cases from May to December in Xinjiang.

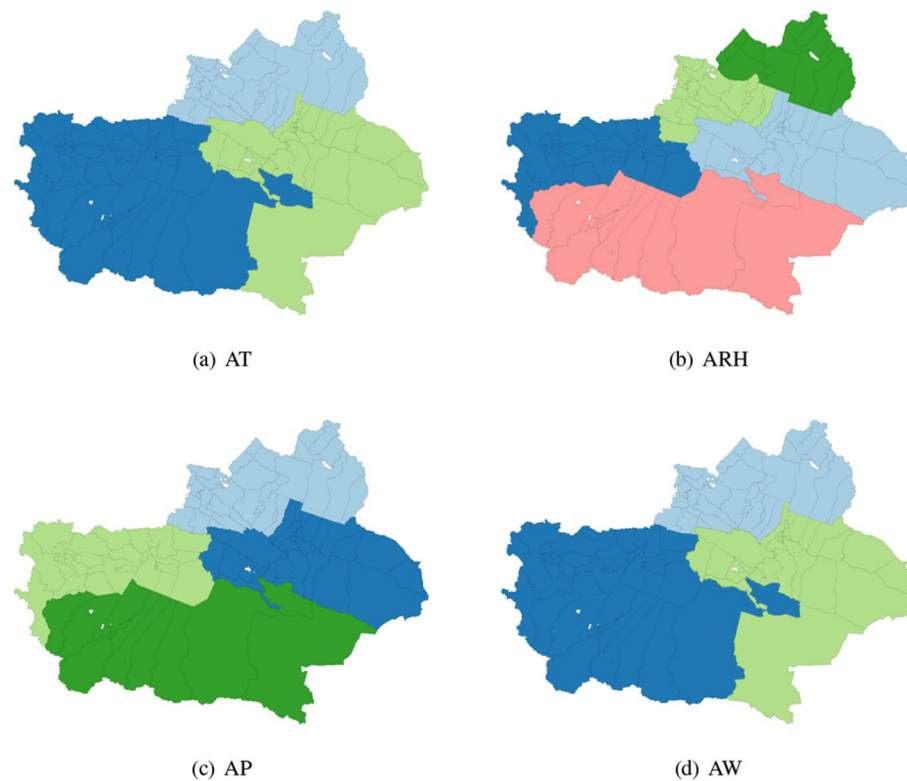


Figure 3. Spatial cluster distribution of some meteorological factors (AT, ARH, AP, AW) in Xinjiang in June 2018.

	NUM	AT	MT	ARH	AP	AW
NUM	1.000					
AT	- 0.439***	1.000				
MT	- 0.486***	0.950	1.000			
ARH	0.511***	- 0.665	- 0.632	1.000		
AP	0.597***	- 0.271	- 0.283	0.480	1.000	
AW	0.064	- 0.354	- 0.376	- 0.127	0.099	1.000

Table 1. Correlation analysis of NUM, AT, MT, ARH, AP, and AW, where the values represent correlation, and ***means that p-value is less than 0.01, that is, there is significant correlation between variables.

	NSCC	Spatial Clustering Area
AT	3	I: Bayingoleng Mongolian Autonomous Prefecture, Turpan region, Hami region
		II: Altay region, and Busier Mongolia Autonomous County, Wuerhe district, Baijiantan district, Karamay district
		III: The northern part of Tacheng area, Bortala Mongolian Autonomous Prefecture, Ili Kazakh Autonomous Prefecture, Aksu Prefecture, Kizilsu Kirgiz Autonomous Prefecture, Kashgar Prefecture, Hotan Prefecture
ARH	5	I: Tianshan district, Shayibak district, Shuimogou district, Midong district, Wujiaqu city, Fukang city
		II: Hami area, Mulei Kazakh Autonomous county, Qitai county, Jimsar county
		III: Remaining areas except I, II, IV, V
		IV: Altay region, Tacheng region, Manas county, Shihezi city, Hutubi county, Changji Hui Autonomous Prefecture
		V: Aral city
AP	5	I: Aksu city
		II: Kashgar city, Artush city, Aheqi county, Keping county, Wushi county, Wensu county, Awati county, Alar city
		III: Altay region (except Jimunai county), Fukang city, Jimsar county, Qitai county
		IV: Remaining areas except I, II, III, V
		V: Xinshi district, Sayibak district, Tianshan district, Shuimogou district
AW	3	I: Altay region, Tacheng region, Changji Hui Autonomous Prefecture, Hami region, Turpan region, Bayingoleng Mongolian Autonomous Prefecture
		II: Remaining areas except I, III
		III: Aksu region, Arche county, Wushi county, Keping county, Bachu county, Tumushuke city, Awati county, Wensu county

Table 2. The SCC results of the regression coefficients in June, where NSCC means number of spatial clustering categories.

role in Bayingoleng Mongolian Autonomous Prefecture, Turpan region, Hami region, Changji, Shihezi and Urumqi, in the Altay region and Hebusaier Mongolian Autonomous Prefecture, Wuerhe county, Baijiantan district and Karamay district, the impact is moderate and in the northern part of Tacheng region, Bortala Mongolian Autonomous Prefecture, Ili Kazakh Autonomous Prefecture, Aksu Region, Kizil Sukyrgiz Autonomous Prefecture, Kashgar and Hotan have weaker effects. The ARH in Tianshan district, Shayibak district, Shuimogou district, Midong district, Wujiaqu city and Fukang county is highly negatively correlated with HFMD cases, but the ARH is strongly positively correlated with the incidence of HFMD in Aral city. The AT has a positive correlation in most areas of Xinjiang, with strong explanatory power in Xinshi district, Shayibak district, Tianshan district and Shuimogou district, but it has a negative correlation in Aksu, which is mainly divided into two levels, among which Aksu shows a high negative correlation, and the influence in other areas is small. The AW shows a positive correlation in the whole Xinjiang, which is roughly divided into three levels of positive correlation, especially in Aksu and some parts of Kashgar, it is strongly correlated. To show the clustering results more clearly, the spatial clustering patterns of these regression coefficients are summarized in Table 2.

In November 2018, the HFMD cases in Xinjiang is mainly distributed according to two characteristics. Specifically, the incidence of HFMD in Bayingoleng Mongolia Autonomous Prefecture, Turpan, Hami and other regions was 12.191 and HFMD cases in other areas of Xinjiang is 10.66. AT shows a negative impact, among which the central and southeast Xinjiang shows a strong effect, and the influence intensity in the Altay region is constant - 0.89. The ARH shows a strong positive correlation in parts of Kyzylsu Kirgiz Autonomous Prefecture, but in the Altay region, the Tacheng region, Borta Lameng Autonomous Prefecture, Ili Kazakh Autonomous Prefecture, Kashgar region of Aksu region, and Hotan region, it shows a strong negative correlation. The AP shows a strong negative correlation in the Altay region, but a high positive correlation in the Saybak region. The AW shows a negative correlation in the whole Xinjiang, which is roughly divided into two levels of positive correlation.

Compared with GWR, the coefficient estimated by SCC has a clearer clustering effect and clear spatial pattern recognition. The reason maybe is the coefficients estimated by GWR are rough and noisy in the study area. Therefore, in the coefficients estimated by GWR, the spatial patterns of clusters are difficult to identify. And by

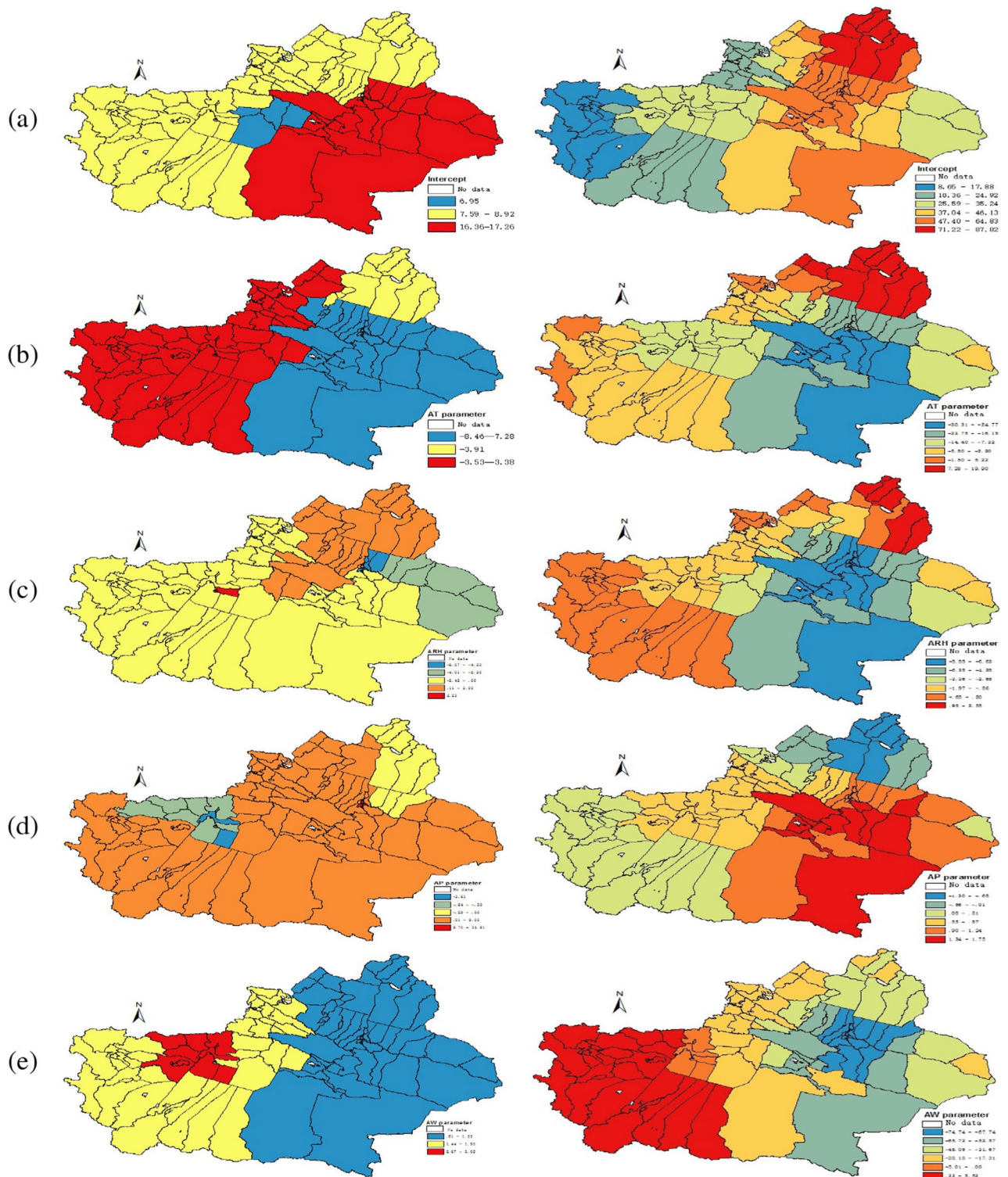


Figure 4. Spatial variation surface of SCC (left panel) and GWR (right panel) model estimation parameters in June 2018, and the estimated regression coefficients of Intercept, AT, ARH, AP, AW in rows (a–e), respectively.

observing Figs. 3 and 4, it is found that the spatial clustering pattern of the regression coefficients is highly similar to the spatial distribution characteristics of meteorological factors, which indicates that the SCC method can accurately identify the real spatial characteristics of the regression coefficients studied in this paper.

The SCC Model evaluation and analysis. In order to intuitively judge the superiority of SCC, this paper calculates the coefficient of determination of the regression model based on the data of June and November, the estimated number of HFMD cases and the mean square error (MSE) of the number of cases per month.

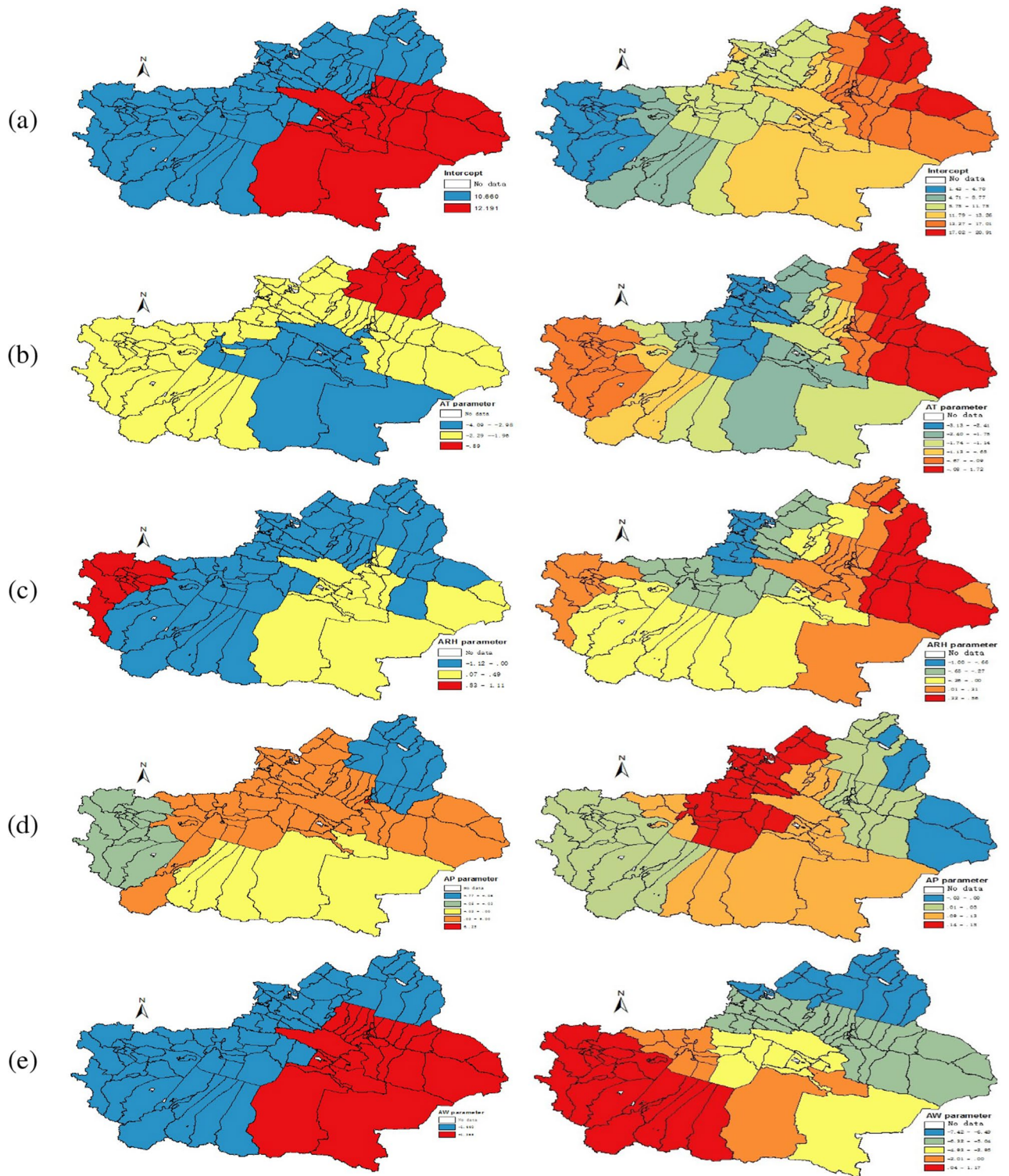


Figure 5. Spatial variation surface of SCC (left panel) and GWR (right panel) model estimation parameters in Nov. 2018, and the estimated regression coefficients of Intercept, AT, ARH, AP, AW in rows (a–e), respectively.

From the observation of Fig. 6 and Table 3, it can be seen that the curve drawn by the actual patient value and the estimated value based on SCC model presents a 45 straight line, and the coefficient of determination estimated based on the SCC method is close to 1, indicating that the estimated value and the true value are almost the same, thus illustrating the accuracy of the SCC model estimation. To further illustrate the accuracy of SCC model estimation, this paper compares SCC and GWR model estimation results, and the results are shown in Fig. 7. The MSE values estimated by SCC are all around 0 and better than the estimated by GWR.

	June	November
SCC	0.999	0.999
GWR	0.383	0.271

Table 3. The coefficient of determination under GWR estimation and SCC estimation.

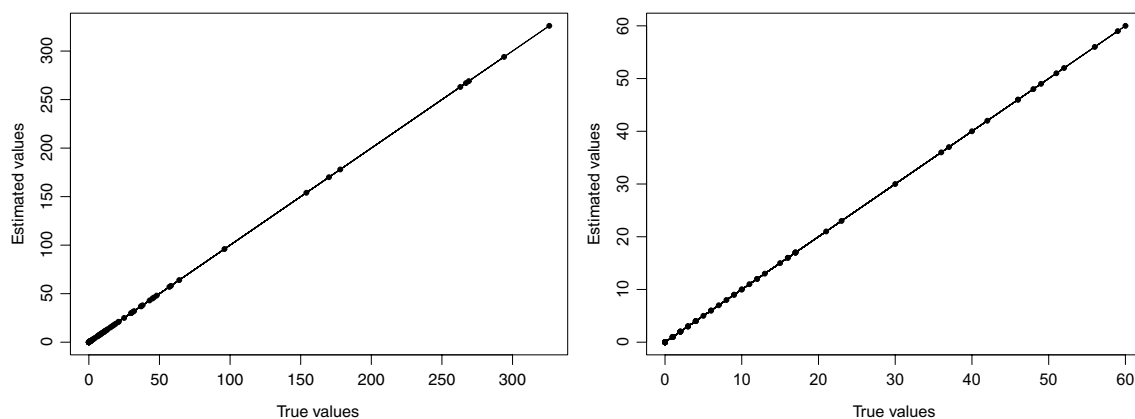


Figure 6. The curve of HFMD incidence estimated based on SCC model and actual incidence, and left panel is in June, 2018 and the right panel is in November, 2018.

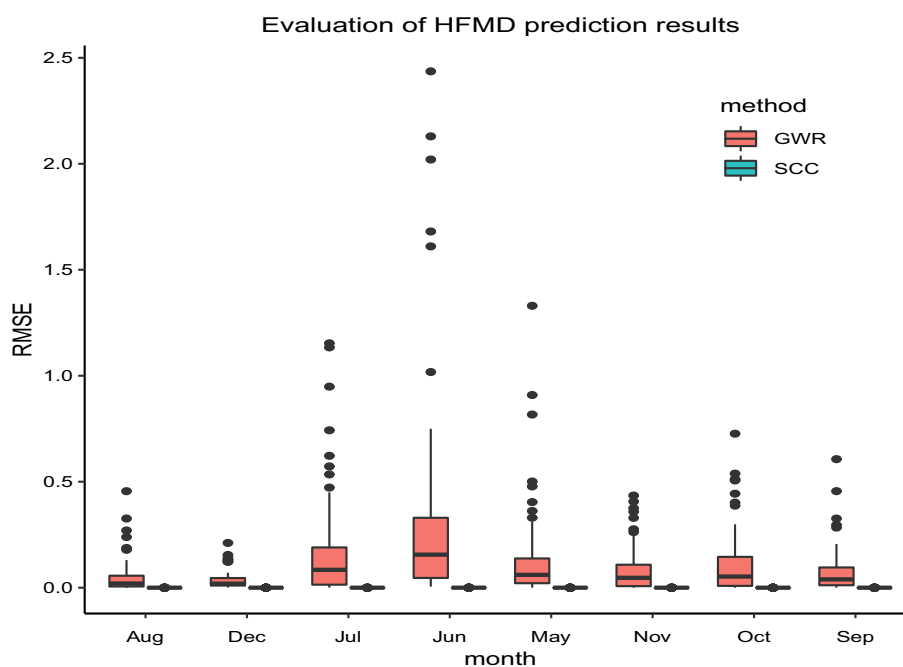


Figure 7. RMSE comparison results based on SCC and GWR.

Discussion

Spatial epidemiology can clearly understand the spatial clustering of diseases, explore the clustering location and exact scope through spatial analysis models²⁰. At present, many studies have carried out spatial correlation analysis on the characteristics of the number of people infected with various infectious diseases, which have not been considered. The influence of other factors on the number of infectious diseases^{21–23}. Moreover, there are few spatial analyses of the incidence of HFMD in Xinjiang, and relevant studies have focused on the analysis of epidemiological signs²⁴.

In this study, based on the dataset of 100 counties and cities in Xinjiang, Spatial Clustering Coefficient (SCC) regression is used to detect the spatial clustering patterns of regression coefficient and estimates the regression coefficient values, so as to speculate the influence pattern of meteorological factors on HFMD cases. The results

show that the regression coefficients have different spatial clustering patterns for different seasons. Compared with previous studies^{25–27}, the AT, ARH, AP and AW factors all have an impact on the number of HFMD cases in Xinjiang. The existing studies have shown that the incidence of HFMD in northern Xinjiang is higher because of its high population density, which is consistent with the results of this study. Secondly, the study finds HFMD cases decline as AT rise in June. Some studies have been shown that the influence of temperature on HFMD is inverted “U”. When the temperature is lower than a certain value, the incidence of HFMD increases with the increase of temperature, whereas the risk of HFMD decreases with the increase of temperature^{4,7}. And Xinjiang is in the midsummer season in June, because the daytime average temperature is as high as 28 °C, which leads to a decrease in the number of people going out, thus reducing the incidence of HFMD¹. On the contrary, AW has a positive impact on the number of HFMD cases. It may be that the increased wind speed promotes the spread of the virus and the higher wind speed will cause cool weather, which will further promote people’s outdoor activities, so the incidence of HFMD increases. The regression coefficients of ARH have five spatial clustering patterns. First, with the increase of ARH, the incidence of HFMD decreases fastest in Urumqi and its surroundings compared with other areas, and at the same time in Aral City, with the increase of ARH, the incidence of HFMD increased rapidly. AP plays an important role in Xinshi District, Sayibak District, and Tianshan District, while it shows a strong negative correlation in Aksu District, and a study in Guangdong revealed that for every 1 hPa increase in air pressure, the number of cases decreased by 6.8²⁸. This may be because lower air pressure may weaken the human’s/organism’s immune system²⁹. In early winter (November), similar to June, AT has a negative correlation with the incidence of HFMD and its regression coefficients have three clusters, but its influence intensity is relatively weak, and it has a strong explanatory power in Bayingoleng Autonomous Prefecture and other areas. The reason may be that as the temperature dropped, the virus began to multiply in large numbers, and susceptible people gathered, which accelerated the spread of the virus²⁸. The above results fully reflect the seasonal characteristics of enterovirus^{6,9}. AW and HFMD cases shows a negative correlation in Xinjiang. According to a systematic review on the association between ventilation and infection suggested that higher ventilation rates could decrease infection rates or outbreaks of some airborne diseases³⁰. Hence, it is likely that strong natural ventilation could serve as a barrier for the spread of respiratory droplets. For the regression coefficients of ARH, it has relatively complex spatial patterns, and it may seem that the ARH in each region has a different impact on the number of HFMD cases due to the complex terrain of Xinjiang. According to previous research, in days with high relative humidity, the infected people may excrete more enteroviruses into the environment³¹, and these enteroviruses can easily attach to the surface of toys or small particles in the air, and the process may cause the virus to accumulate in the environment, thereby accelerating the spread of HFMD³². This is consistent with the results of this study in November (except for the northern, northwestern, and southwestern regions of Xinjiang).

In summary, the coefficient spatial clustering patterns of AT, ARH, AP and AW in June and November in Xinjiang are detected by using the SCC method, and the influence characteristics of meteorological factors on HFMD are analyzed. It is hoped that this will provide some theoretical basis for prevention and control of HFMD in Xinjiang.

Methods

Data source. The research data of this paper are the number of confirmed HFMD cases in 100 cities and counties (municipal districts) in Xinjiang from January 2018 to December 2018, and the daily incidence of HFMD in Urumqi, Xinjiang in 2018, as well as five meteorological factors: daily average temperature (°C), minimum temperature (°C), average relative humidity (%), average air pressure (Pa) and average wind speed (m/s). The data are from the Public Health Science Data Center of Chinese Center for Disease Control and Prevention (<http://www.phsciencedata.cn/Share/index.jsp>).

Model. Suppose the spatial data $(x(u_i, v_i), y(u_i, v_i)), i = 1, \dots, n$ is the observation of the observation position $(u_1, v_1), \dots, (u_n, v_n) \in R^2$, where the response variable $y(u_i, v_i)$ is spatially correlated and $\mathbf{x}(u_i, v_i) = (x_1(u_i, v_i), \dots, x_p(u_i, v_i))$ is p dimensional explanatory variables at (u_i, v_i) . Consider a spatial variable coefficient model

$$y(u_i, v_i) = \sum_{k=1}^p x_k(u_i, v_i) \beta_k(u_i, v_i) + \epsilon(u_i, v_i). \quad (1)$$

Clearly, this problem needs to be regularized because there are more variables and parameter than observations. For spatial problems, the association between the response variable and the explanatory variables at nearby locations is expected to be highly homogenous, which prompted us to assign β a regularization function that reflects this spatial homogeneity.

Specifically, the paper minimizes the SCC model to estimate β

$$\frac{1}{n} \sum_{i=1}^n [y(u_i, v_i) - \sum_{k=1}^p x_k(u_i, v_i) \beta_k(u_i, v_i)]^2 + \lambda \sum_{k=1}^p \sum_{(i,j) \in E} |\beta_k(u_i, v_i) - \beta_k(u_j, v_j)|. \quad (2)$$

where E is an edge set consisting of n vertices, each of which corresponds to an observation position, and $|\cdot|$ is a generalized lasso penalty function to judge the corresponding positions of the two regression coefficients (u_i, v_i) and (u_j, v_j) are connected by an edge in E , thus encouraging homogeneity between the two regression coefficients. λ is a penalty parameter, which is selected using the BIC criterion in this example. Once an edge set E is given, (2) is written in matrix form

$$\frac{1}{n} \sum_{i=1}^n [y(u_i, v_i) - \sum_{k=1}^p x_k(u_i, v_i) \beta_k(u_i, v_i)]^2 + \lambda \sum_{k=1}^p |\mathbf{H} \boldsymbol{\beta}_k|_1. \quad (3)$$

where \mathbf{H} is an $m \times n$ matrix consisting of an edge set E with m edges. For the edge connecting two positions (u_i, v_i) and (u_j, v_j) this paper denote the penalty term $|\beta_k(u_i, v_i) - \beta_k(u_j, v_j)|$ as $|\mathbf{H}_m \boldsymbol{\beta}_k|$, \mathbf{H}_m represents the row vector of \mathbf{H} , which contains only two non-zero elements, the i -th element is 1, and the j -th element is -1 .

The marginal set E is a key component of the SCC model because it reflects prior assumptions about the structure of the regression coefficients. This paper constructs the edge set E by generating a minimum spanning tree (MST).

Once the MST is constructed, the resulting penalty no longer contains redundant terms, so the original problem can be easily transformed into a lasso or lasso-type problem after appropriate reparameterization. Define new parameters $\boldsymbol{\theta}_k, k = 1, \dots, p$ as $\boldsymbol{\theta}_k = \begin{pmatrix} \mathbf{H} \\ \frac{1}{n} \mathbf{1}^T \end{pmatrix} \boldsymbol{\beta}_k = \tilde{\mathbf{H}} \boldsymbol{\beta}_k$.

This new design matrix is written as $\tilde{\mathbf{X}} = [\text{diag}(\mathbf{X}_1) \tilde{\mathbf{H}}^{-1}, \dots, \text{diag}(\mathbf{X}_n) \tilde{\mathbf{H}}^{-1}]$, where $\tilde{\mathbf{H}}$ is an $n \times n$ invertible matrix. Since \mathbf{H} rows are full rank, there is a one-to-one transformation between $\boldsymbol{\beta}_k$ and $\boldsymbol{\theta}_k$. Then, the SCC model in equation 4 can be rewritten as

$$\frac{1}{n} \|\mathbf{y} - \tilde{\mathbf{X}} \boldsymbol{\theta}\|_2^2 + \lambda \sum_{l \in B} |\theta_l|. \quad (4)$$

where $\boldsymbol{\theta} = (\theta'_1, \dots, \theta'_p)'$ is a vector of np , and B represents the index set $B = \{l : \text{mod}(l, n) \neq 0, l = 1, \dots, np\}$, does not include the n -th element. For simplicity, the paper will denote $\sum_{l \in B} |\theta_l|$ by $|\boldsymbol{\theta}_B|_1$. Therefore, by solving the lasso problem in (4) with respect to the parameter $\boldsymbol{\theta}$, the solution of the SCC model (2) with lasso penalty can be obtained. The estimator of $\boldsymbol{\beta}$ is given by $\hat{\boldsymbol{\beta}}_k = \tilde{\mathbf{H}}^{-1} \hat{\boldsymbol{\theta}}_k, k = 1, \dots, p$.

Ethics declarations. This study does not involve human participants, and uses public data from the Public Health Science Data Center of Chinese Center for Disease Control and Prevention, so it was not approved by the Ethical Committee.

Data availability

Hand, Foot and Mouth Disease datasets generated during the current study are available in the Public Health Science Data Center of Chinese Center for Disease Control and Prevention (<http://www.phsciencedata.cn/Share/index.jsp>). The meteorological datasets are available in China Meteorological Data Network (<http://data.cma.cn/>).

Received: 29 August 2022; Accepted: 7 December 2022

Published online: 12 December 2022

References

- Dong, W. *et al.* The effects of weather factors on hand, foot and mouth disease in Beijing. *Sci. Rep.* **6**(1), 1–9 (2016).
- Xu, Y. *et al.* Pathogenic characteristics of hand, foot and mouth disease in Shaanxi Province, China, 2010–2016. *Sci. Rep.* **10**(1), 1–11 (2020).
- Du, Z., Zhang, W., Zhang, D., Yu, S. & Hao, Y. The threshold effects of meteorological factors on Hand, foot, and mouth disease (HFMD) in China, 2011. *Sci. Rep.* **6**, 36351 (2016).
- Wang, P., Goggins, W. B. & Chan, E. Y. Y. Hand, foot and mouth disease in Hong Kong: A time-series analysis on its relationship with weather. *PLoS ONE*. **11**(8), e0161006 (2016).
- Wu, H., Wang, H., Wang, Q., Xin, Q. & Lin, H. The effect of meteorological factors on adolescent hand, foot, and mouth disease and associated effect modifiers. *Glob. Health Action*. **7**(1), 24664 (2014).
- Ma, E., Lam, T., Wong, C. & Chuang, S. K. Is hand, foot and mouth disease associated with meteorological parameters?. *Epidemiol. Infect.* **138**(12), 1779–1788 (2010).
- Huang, Y. *et al.* Effect of meteorological variables on the incidence of hand, foot, and mouth disease in children: A time-series analysis in Guangzhou, China. *BMC Infect. Dis.* **13**(1), 1–8 (2013).
- Lin, H. *et al.* Short-term effect of El Nino-Southern Oscillation on pediatric hand, foot and mouth disease in Shenzhen, China. *PLoS ONE*. **8**(7), e65585 (2013).
- Urashima, M., Shindo, N. & Okabe, N. Seasonal models of herpangina and hand-foot-mouth disease to simulate annual fluctuations in urban warming in Tokyo. *Jpn. J. Infect. Dis.* **56**(2), 48–53 (2003).
- Hu, M. *et al.* Determinants of the incidence of hand, foot and mouth disease in China using geographically weighted regression models. *PLoS ONE*. **7**(6), e38978 (2012).
- Hong, Z. *et al.* Exploration of potential risks of hand, foot, and mouth disease in Inner Mongolia Autonomous Region, China using geographically weighted regression model. *Sci. Rep.* **8**(1), 1–10 (2018).
- Wang, H. *et al.* Detecting the association between meteorological factors and hand, foot, and mouth disease using spatial panel data models. *Int. J. Infect. Dis.* **34**, 66–70 (2015).
- Bo, Y. C., Song, C., Wang, J. F. & Li, X. W. Using an autologistic regression model to identify spatial risk factors and spatial risk patterns of hand, foot and mouth disease (HFMD) in Mainland China. *BMC Public Health*. **14**(1), 1–13 (2014).
- Gómez-Rubio, V., Ferrández-Ferragud, J. & López-Quilez, A. Detecting clusters of disease with R. *J. Geogr. Syst.* **7**(2), 189–206 (2005).
- Conley, J., Gahegan, M. & Macgill, J. A genetic approach to detecting clusters in point data sets. *Geogr. Anal.* **37**(3), 286–314 (2005).
- Zhou, T., Enyeart, P. J. & Wilke, C. O. Detecting clusters of mutations. *PLoS ONE* **3**(11), e3765 (2008).
- Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. & Knight, K. Sparsity and smoothness via the fused lasso. *J. R. Stat. Soc. B (Stat. Methodol.)* **67**(1), 91–108 (2005).
- Ke, Z. T., Fan, J. & Wu, Y. Homogeneity pursuit. *J. Am. Stat. Assoc.* **110**(509), 175–194 (2015).
- Li, F. & Sang, H. Spatial homogeneity pursuit of regression coefficients for large datasets. *J. Am. Stat. Assoc.* **114**(527), 1050–1062 (2019).

20. Wang, T., Xue, F., Chen, Y., Ma, Y. & Liu, Y. The spatial epidemiology of tuberculosis in Linyi city, China, 2005–2010. *BMC Public Health*. **12**, 885 (2012).
21. Ning, G., Yin, Z., Wu, D., Li, J. & Li, Y. Spatial clustering of Japanese encephalitis in China 2013. *Chin. J. Vacc. Immuniz.* **21**(4), 365–368 (2015) (In Chinese).
22. Yang, G. *et al.* Analysis of spatial autocorrelation of HIV/AIDS in Chongqing from 2004 to 2012. *J. Shanghai Jiaotong Univ. Med. Sci.* **35**(1), 112–116 (2015) ((In Chinese)).
23. Bi, S., Bie, S., Hu, X. & Zhang, H. Analysis on the characteristics of spatio-temporal evolution and aggregation trend of early COVID-19 in mainland China. *Sci. Rep.* **12**(1), 1–7 (2022).
24. Liu, W., Huang, Y., Ma, L., Ayiguli, Y. & Deng, H. Epidemiological characteristics of hand foot and mouth disease in Xinjiang over 2011–2015 years. *J. Xinjiang Med. Univ.* **41**(3), 357–361 (2018) (In Chinese).
25. Huang, R. *et al.* Correlation analysis on meteorological factors the incidence of hand, foot and mouth disease in Xinjiang Uygur Autonomous Region, 2011–2018. *Chin. J. Epidemiol.* **40**(12), 1563–1568 (2011) (In Chinese).
26. Xie, L., Huang, R., Wang, H. & Liu, S. Spatial-temporal heterogeneity and meteorological factors of hand-foot-and-mouth disease in Xinjiang, China from 2008 to 2016. *PLoS ONE* **16**(8), e0255222 (2021).
27. Gao, Y. *et al.* Spatial and temporal characteristics of hand-foot-and-mouth disease and their influencing factors in Urumqi, China. *Int. J. Environ. Res. Public Health*. **18**(9), 4919 (2021).
28. Li, T., Yang, Z., Liu, X., Kang, Y. & Wang, M. Hand-foot-and-mouth disease epidemiological status and relationship with meteorological variables in Guangzhou, southern China, 2008–2012. *Rev. Inst. Med. Trop. São Paulo*. **56**, 533–539 (2014).
29. Maes, M. & Meyer, F. D. Relationships of climatic data to immune and hematologic variables in normal human. *Neuroendocrinol. Lett.* **21**(2), 127–136 (2000).
30. Chartier, Y. & Pessoa-Silva, C. *Natural Ventilation for Infection Control in Health-Care Settings* (World Health Organization, 2009).
31. Fong, T. & Lipp, E. Enteric viruses of humans and animals in aquatic environments: Health risks, detection, and potential water quality assessment tools. *Microbiol. Mol. Biol. Rev.* **69**(2), 357–371 (2005).
32. Fletcher, J. The importance of correctly choosing a bandage and bandaging technique. *Nurs. Times*. **100**(32), 52–53 (2004).

Acknowledgements

This study is supported by The Ministry of Education of Humanities and Social Science project (Grant Number 19YJA910007) and the Natural Science Foundation of Xinjiang (Grant Number 2019D01C045).

Author contributions

X.S.Z. analyzed the data and wrote the paper. P.L.W. and H.G.Z. provided amendments. H.G.Z. reviewed the final paper. All authors approved the final paper and accepted the accountability for all aspects of the work.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.Z.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022