Assessment for Spatial Driving Forces of HFMD Prevalence in Beijing, China

Jiaojiao Wang  
The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences  
No. 95, Zhongguancun East Road, Haidian District, Beijing, China  
+86-10-82544716  
jiaojiao.wang@ia.ac.cn

Zhidong Cao  
The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences  
No. 95, Zhongguancun East Road, Haidian District, Beijing, China  
+86-10-82544716  
zhidong.cao@ia.ac.cn

Quanyi Wang  
Beijing Center for Disease Prevention and Control  
No.16 He Pingli Middle Street, Dongcheng District, Beijing, China  
+86-10-64407122  
bjcdcxm@126.com

Xiaoli Wang  
Beijing Center for Disease Prevention and Control  
No.16 He Pingli Middle Street, Dongcheng District, Beijing, China  
+86-10-64407122  
wangxiaoli198215@163.com

Daniel Dajun Zeng  
The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences  
University of Chinese Academy of Sciences  
No. 95, Zhongguancun East Road, Haidian District, Beijing, China  
+86-10-82544716  
dajun.zeng@ia.ac.cn

ABSTRACT

Hand-foot-mouth disease (HFMD) outbreak greatly threatened Beijing city, the capital city of China, in 2008. The control prevention of HFMD has become an urgent mission for Beijing Center for Disease Control and Prevention and a focus problem for the citizens. Medical, social and environmental situations account for much of HFMD morbidity. The spatial driving forces of HFMD occurrence vary across geographical regions, whereas the factors that play a significant role in HFMD prevalence may be concealed by global statistics analysis. This study aims at the identification of the association between the spatial driving forces and HFMD morbidity across the study area and the epidemiological explanation of the results. HFMD spatial driving forces are represented by 6 factors which was obtained by Pearson Correlation analysis and Stepwise Regression method. Compared to Classical Linear Regression Model (CLRM), Geographically weighted regression (GWR) techniques were implemented to predict HFMD morbidity and examine the nonstationary of HFMD spatial driving forces. Informative maps of estimated HFMD morbidity and statistically significant spatial driving forces were generated and rigorously evaluated in quantitative terms. Prediction accuracy by GWR was higher than that by CLRM. The residual led to by CLRM suggested a significant degree of spatial dependence, while by GWR indicated no significant spatial dependence. In the three regions plotted by Beijing city Ring Roads, HFMD morbidity was found to have significantly positive or negative association with the 6 kinds of spatial driving forces. GWR model can effectively represent the spatial heterogeneity of HFMD driving forces, significantly improve the prediction accuracy and greatly decrease the spatial dependence. The results improve current explanation of HFMD spread in the study area and provide valuable information for adequate disease intervention measures.

CCS Concepts

• Modeling and simulation → Model verification and validation

Keywords

Driving forces; Geographically weighted regression (GWR); Spatial heterogeneity; Hand-foot-mouth disease (HFMD)

1. INTRODUCTION

The first Hand-foot-mouth disease (HFMD) case was reported in Shanghai, China in 1981. Since then, HFMD epidemic has occurred in many provinces and cities in China. Over the past few years, HFMD prevalence in China kept showing an increasing trend as the scale gradually became larger, posing a great threat to public health and security. In May 2nd, 2008, HFMD was brought into infectious diseases management of Grade C in China [1]. Beijing city, the capital of China, is vastly threatened by HFMD. During the last 3 years, the number of HFMD cases has been always holding the primacy among Grade C infectious diseases in Beijing city. In 2008, Beijing Center for Disease Control and Prevention (CDC) established “Beijing Hand-foot-mouth disease prevention and control scheme”.

In recent years, Chinese scientists have conducted Beijing HFMD-related research from the perspectives of molecular biology [2, 3], clinical medicine [4,5,6], etiology [7,8] and epidemiology [9,10], and some progresses have been made in these areas. Although the epidemiological features of Beijing HFMD epidemic have been verified, there is still much work in the research on the driving forces (such as natural, social environment and human factors) of HFMD prevalence. In traditional linear regression framework, the elastic coefficients of independent variables are assumed to be constant over space, however, infectious diseases prevalence actually shows spatial...
heterogeneity. The estimated results by traditional regression framework represent the average level throughout the study area, covering up the very significant geographic and socio-economic phenomena. Geographically weighted regression (GWR) technique allows local parameters to be estimated and assumes there is spatial nonstationary. GWR model can effectively explain the relationship between HFMD morbidity and its spatial driving forces, which plays a key role in the regional disease prevention and control.

GWR was developed as a new technique which extends the traditional regression framework [11, 12]. GWR has been adopted in various fields since it was put forward. Fotheringham and Charlton [13] described the spatial distribution of limiting long-term illness (LLTI) based on the data from 605 census wards in the northeast of England. Brunsdon and Fotheringham [14] analyzed an example data set based on house prices in Kent in the U.K. Fotheringham [15] assessed the performance of all state primary schools in Britain. In recent years, many researchers have examined GWR technique as follow:

1. **Empirical application.** (1) **Geography.** Socioeconomic factors of urban heat islands and land-cover [16], regional analysis of wealth and the land cover in Massachusetts [17], spatially varying relationships between land use and water quality [18], white-tailed deer distribution [19], geographical factors of theft crime [20]. (2) **Ecology and environment.** The relationship between plant diversity and climate-environment [21], spatial patterns of species richness in New World coral snakes [22], estimation of LAI [23], spatiotemporal dynamics of forest net primary production in China [24]. (3) **Economics.** Persistent pockets of extreme American poverty and job growth [25], spatial variation in housing attribute prices [26]. (4) **Public health.** Geographical variations in poverty-obesity relationships [27], local determinants of neural tube defects [28], spatial patterns of mortality in Atlanta metropolitan area [29], spatial variations in heart disease mortality [30].

2. **Parameter estimation.** Statistical tests for spatial nonstationary and mixed GWR. GWR-SEM model measuring the value of air quality [31], GWR Logistic model for urban growth [32], assessment of coefficient accuracy [33], testing the importance of the explanatory variables [34], GWPR model for disease association mapping [35].

The aim of this study was to assess the spatial driving forces of Beijing 2008 HFMD prevalence in China. The exploration was necessary for the adoption of the treatment to the infected individuals and provided the important clues for the measures to safeguard the residents in Beijing and other regions where there is HFMD outbreak.

2. **METHODS**

**Study area.** Beijing is situated at the northern tip of North China Plain (Figure 1), which opens to the south and east of the city, with mountains to the north, northwest and west shielding the city and northern China’s agricultural heartland from the encroaching desert steppes. The northwestern part of the municipality, especially Yangqing County and Huaiou District, are dominated by the Jundu Mountains, while the western part of the municipality is framed by Xishan Mountains. Beijing is also the northern terminus of the Grand Canal of China and Miyun Reservoir is crucial to its water supply. The urban area of Beijing is situated in the south-central part of the municipality and occupies a small but expanding part of the municipality’s area. It spreads out in bands of concentric road rings. The city’s climate is a monsoon-influenced humid continental climate, characterized by hot, humid summers due to the East Asian monsoon, and generally cold, windy, dry winters that reflect the influence of the vast Siberian anticyclone. Majority of precipitation falls in the summer months. It is divided into 16 districts consisting of 309 towns and villages with a total area of 16,801.25 km². It is a major transportation hub with dozens of railways, roads and motorways passing through the city. It is the destination of many international flights, and the political, educational, cultural center of China.

![Figure 1. Location of Beijing in China.](image)

**HFMD morbidity.** Beijing CDC received 18,446 observed HFMD cases in 2008 covering 309 towns and villages in 16 administrative districts through epidemiology survey.

**Potential Risk factors.** There are many risk factors influencing HFMD epidemic, such as urbanization (agricultural land, constructed land, unused land), environmental (greenspace), demographic (kindergarten, population density), socioeconomic (GDP per capita, Unit GDP energy consumption, disposable income of urban residents), health service (healthcare organizations, medical practitioners, registered nurses, beds in hospital), and climate (annual average temperature, annual average relative humidity). The preceding study has indicated that health service play an important role in HFMD prevalence [9]. Therefore, we selected the following indices which could directly reflect regional health service such as the number of healthcare organizations per 10,000 people, the number of doctors/nurses/hospital beds per 1,000 people as risk factors.

**Risk factors selection.** The relativity between HFMD morbidity and potential risk factors referred to above was denoted by Pearson correlation. Factor ‘agricultural land’ and ‘registered nurses’ was eliminated because of being not significantly associated with HFMD morbidity. Factor ‘healthcare organizations’ exhibited the strongest association with HFMD morbidity (-0.436**). In order to solve the multicollinearity problem, Stepwise Regression method was adopted to remove the redundant variables. Finally, 6 risk factors significantly associated with HFMD morbidity were found, being disposable income of urban residents (noted as DIS INC), healthcare organizations (noted as HEA ORG), beds in hospital (noted as HOS BEDS), population density (noted as POP DEN), annual average temperature (noted as TEMP) and annual average relative humidity (noted as REL HUM). Figure 2 shows the spatial distribution of the above 6 risk factors which was selected as spatial driving forces in this study.
where $y$ is the HFMD morbidity, $x_0$ is the intercept, respectively represents the parameter estimate for independent variables, $e$ is the error term, $x_1$ is POP DEN, $x_2$ is DIS INC, $x_3$ is HEA ORG, $x_4$ is HOS BEDS, $x_5$ is REL HUM, $x_6$ is TEMP. The unknown parameters are estimated via Ordinary Least Squares (OLS) technique.

GWR is an extension of the traditional standard regression framework by allowing local rather than global parameters to be estimated. It can produce a set of local parameter estimates showing how a relationship varies over space and then to examine the spatial pattern of the local estimates to get some understanding of hidden possible driving forces of this pattern [15]. GWR model can be written as [11]:

$$y = x_0 + x_1x_1 + ... + x_6x_6 + e$$

where $n$ is the sample size, $i$ is the geographical coordinate for location, $b$ is the intercept for location, $a_i$ is the local parameter estimate for independent variable at location, $e_i$ is the estimation for dependent variable at location, $w_{ij}$ is the random error term for location (satisfy the Assumption of zero mean, homoscedastic and independent). If $b$ keeps constant, GWR will be degenerated to CLRM. The estimation for $b$ is calculated by the equation as follow:

$$b = \frac{\sum_i a_i w_{ij} y_i}{\sum_i a_i^2}$$

where $y_i$ represents the weight of observation at point $i$ for observation at another point in the study area, which can be represented as distance decay function and suggest the importance of the observation location for parameters estimate. The essence of GWR is Locally Weighted Least Squares technique in which the weight is a continuous function of the distance between any two observations in the study area. In this study, GWR technique was implemented to explore the spatial driving forces of HFMD epidemic.

### 3. RESULTS

#### 3.1 CLRM Estimation

The global performance of CLRM is shown in Table 1. It reports that intercept, DIS INC and Temp were positively associated with HFMD morbidity, while POP DEN, HEA ORG, HOS BEDS and REL HUM were negatively associated with HFMD morbidity. The $R^2$ value by CLRM was 0.50. Moran’s I index is often used to test for global autocorrelation [38], of which the value of 0.4787 ($p<0.05$) suggested that the residual led to by CLRM still remained significant autocorrelation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>T value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP DEN</td>
<td>-0.0052</td>
<td>-7.5209</td>
<td>0.0000</td>
</tr>
<tr>
<td>DIS INC</td>
<td>0.0108</td>
<td>3.2315</td>
<td>0.0014</td>
</tr>
<tr>
<td>HEA ORG</td>
<td>-28.7730</td>
<td>-4.0036</td>
<td>0.0001</td>
</tr>
<tr>
<td>HOS BEDS</td>
<td>-17.3783</td>
<td>-5.0093</td>
<td>0.0000</td>
</tr>
<tr>
<td>REL HUM</td>
<td>-14.8308</td>
<td>-5.2638</td>
<td>0.0000</td>
</tr>
<tr>
<td>TEMP</td>
<td>20.5377</td>
<td>3.6371</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

#### 3.2 GWR Performance

Figure 3 presents the maps of GWR local regression, local $R^2$ ranging from 0.42 to 0.76 and the average $R^2$ reaching 0.68. The Moran’s I value by GWR was 0.0412 ($p>0.05$), representing that residual spatial autocorrelation by GWR was greatly eliminated. Figure 4 illustrates the spatial distribution of HFMD morbidity in three forms. Figure 5 plots the local coefficients for every variable together with significance levels, in which the shadowed regions represented the parameters estimate of these areas was not significant. Table 2 summarizes the associations between spatial driving forces and HFMD morbidity in different regions throughout Beijing city.
4. DISCUSSION

As Figure 4 shows, Beijing 2008 HFMD cases were mostly aggregated in urban-rural transition zone and relatively less distributed in central area and suburban area. It would be possibly
caused by the following reasons: 1) There was higher population density, better housing conditions and medical infrastructures in central area. 2) There was lower population density and much more excellent natural- original-ecological environment although its relatively poor public health conditions in suburban area. 3) While in urban-rural transition area, although population density ranged between the other two regions, there was more population migration and poorer public health conditions. As a result, a higher prevalence of HFMD would probably occur in the transition area.

According to the above CLRM estimation and GWR performance, it suggests that GWR model effectively demonstrates the spatial heterogeneity and local driving forces of HFMD prevalence: 1) It is clear that in central area, with high population density, relative humidity, temperature and disposable income, better medical and health conditions and stronger consciousness of diseases prevention and health promotion, HFMD morbidity is low. 2) In Suburban Area, although relative humidity is high and public health conditions are poor, with lower population density and excellent natural ventilation, low HFMD morbidity displays in the region. 3) There are probably many driving forces of HFMD epidemic representing in urban-rural transition area, such as frequent contact among crowds of people, high-density residential zones, inadequate medical cares, low disposable income, high relative humidity and temperature and poor natural ventilation, therefore HFMD morbidity exhibits higher.

5. CONCLUSIONS
This study has been carried out based on 18,446 HFMD cases covering 309 township units in Beijing during the year 2008. Results shows that the relationship between HFMD morbidity and driving forces presents spatial heterogeneity, which is overlooked by CLRM while solved by GWR. As a result of neglecting the spatial dependence among the disease cases, CLRM clearly shows two drawbacks: 1) Low prediction accuracy. 2) Estimation bias in prediction, which would mislead management, decision, understanding, simulation and prevention of infectious diseases. Aiming at these defects, GWR technique performs better than CLRM in solving non-stationary spatial problems, removing residual spatial dependence and exploring the local variation in parameters estimate and HFMD spread. Although the conclusions could be partially limited, the significant spatial driving forces of HFMD prevalence provides important clues for local infectious diseases prevention and control.

To sum up, the local statistics approach GWR can identify the townships where the HFMD morbidity is significantly associated with spatial driving forces. In the three regions plotted by the city ring roads, HFMD morbidity was found to have significantly positive or negative association with the 6 kinds of driving forces. In Beijing, it is urgent to adopt an appropriate assessment for spatial driving forces of infectious diseases, to permit the adoption of precautions and measures and to emphasize the fundamentals of basic hygiene in order to safeguard the public health of the city. Detecting and improving the knowledge of spatial patterns of epidemic transmission will help the prevention and management of infectious hazard.

6. ACKNOWLEDGMENTS
This research was supported by the National Natural Science Foundation of China (Grant No. 91546112, 91224008, 71402177, 71621002, 71602184, 71603253), the National Major Research Program of China (Grant No. 2016YFC1200702) and the Early Career Development Award of SKLMCCS (Y3S9021F37, Y6S9011F4F).

7. REFERENCES


