Using Geographically Weighted Regression to Study the Seasonal Influence of Potential Risk Factors on the Incidence of HFMD on the Chinese Mainland

Jingtao Sun, Sensen Wu, Zhen Yan, Yadong Li, Cheng Yan, Feng Zhang, Renyi Liu, and Zhenhong Du

Abstract: Hand, foot, and mouth disease (HFMD) is an epidemic infectious disease in China. Its incidence is affected by a variety of natural environmental and socioeconomic factors, and its transmission has strong seasonal and spatial heterogeneity. To quantify the spatial relationship between the incidence of HFMD (I-HFMD) and eight potential risk factors (temperature, humidity, precipitation, wind speed, air pressure, altitude, child population density, and per capita GDP) on the Chinese mainland, we established a geographically weighted regression (GWR) model to analyze their impacts in different seasons and provinces. The GWR model successfully describes the spatial changes of the influence of potential risks, and shows greatly improved estimation performance compared with the ordinary linear regression (OLR) method. Our findings help to understand the seasonally and spatially relevant effects of natural environmental and socioeconomic factors on the I-HFMD, and can provide information to be used to develop effective prevention strategies against HFMD at different locations and in different seasons.

Keywords: GWR; HFMD; t-test; spatial non-stationary; seasonal non-stationary

1. Introduction

Hand, foot, and mouth disease (HFMD) is a common infectious disease that is usually found in children under 5 years old. This disease is caused by viruses, such as Human enterovirus 71 (EV71) and Coxsackie virus A16 strain (CoxA16), and can result in symptoms in the hand, mouth, or foot, including fever, blisters, and ulcers [1]. In most cases, the disease is mild and self-limiting, but severe neurological symptoms, such as meningitis, encephalitis, poliomyelitis (such as paralysis), and pulmonary edema may occur, especially in patients aged 5 or under [2]. Surveillance data from 2008 to 2017 show that the incidence of HFMD (I-HFMD) was considerably high and that the high-risk areas were mainly in central, southern, and eastern China [3]. A large number of epidemiological studies on HFMD have been carried out in provinces with a high I-HFMD, such as Guangdong [4–6], Sichuan [7], Henan [8,9], and Shandong [10].

Many studies have indicated that the I-HFMD is strongly related to socioeconomic factors and natural environmental factors, such as temperature [1,11–14], precipitation [1,5,15,16], humidity [13–17], altitude [18], wind speed [1,19], air pressure [20], child population density [21,22] and per capita GDP [23–26]. Quantifying the influence of potential risk factors on the I-HFMD is conducive to the prevention and treatment of the disease. Most studies of HFMD have used exploratory data analysis methods to analyze
the spatial distribution of HFMD in the form of graphs and tables, or classical data models to predict HFMD and estimate its potential risk factors, for which the model types mainly include dynamic models [27–29], linear regression models [5,14], seasonal moving average models [24,30], and Bayes networks [24].

Current studies have generally used datasets derived from disease surveillance systems with spatiotemporal information. As the classical data model cannot effectively mine the geographical information in the HFMD monitoring data, some spatial statistical methods, such as the geographically weighted regression (GWR) model [21,22,31], have also been applied to study HFMD. However, these studies have mainly focused on a single dimension of time and space, rather than attempting a combined analysis on a large scale and in the long term. Furthermore, studies in China have mainly focused on a particular city or province and have usually adopted a time span of a single season or year.

Few studies have investigated the spatial and temporal effects of the I-HFMD across the entire Chinese mainland over the long term. Because the spread of HFMD in China can be quite large in scale and significantly seasonal over the long term, studying HFMD in a local area (e.g., a city or a province) over a short period (e.g., a season or a year) can reflect some local-level laws but not the spatiotemporal changes and their mechanisms across the entire Chinese mainland. In addition, due to the complex landform and vast area of the Chinese mainland, managing the spatiotemporal heterogeneity in this region is quite difficult and needs further investigation. Therefore, it is of great importance to study HFMD at the scale of the whole Chinese mainland, as this can capture more accurate and useful information on which to base suggestions for restraining the spread of HFMD in China. In this study, the spatial association between the I-HFMD and socioeconomic and natural environmental factors was therefore investigated to explore the potential spatiotemporal non-stationary characteristics of the hypothetical relationship between these factors and the I-HFMD in the Chinese mainland.

This study had three main objectives: (1) study the spatial and seasonal relationship between the I-HFMD and potential influencing factors based on monthly average data for 11 years in the Chinese mainland with 31 provinces; (2) construct the GWR model to estimate the I-HFMD in the Chinese mainland and compare its performance with the ordinary linear regression (OLR) model; and (3) analyze local heterogeneity and put forward appropriate measures to limit the I-HFMD in each season. Our article is organized as follows. In the second section, we describe the study area and dataset and introduce the methods of the spatial autocorrelation test and the GWR model. Sections 3 and 4 describe the results of the GR model for I-HFMD estimation and analyze the seasonal characteristics of the influence of each potential risk factor. Finally, we draw conclusions and summarize the paper in Section 5.

2. Materials and Methods

2.1. Research Region and Data

The study area used in this study was the Chinese mainland, including 31 provinces (except Hong Kong, Macau, and Taiwan). The spatial measured unit was the provincial administrative unit, and the time span was from 2007 to 2017 with a time resolution of 1 month. The I-HFMD data were the monthly incidence data for each province, collected by the Chinese Center for Disease Control, Prevention and Health. The collected I-HFMD data were divided into four seasons with each season including 31 provinces and 36 months with more than 1000 records. The distribution of the cumulative incidence rate from 2007 to 2017 in each season for the 31 provinces of the Chinese mainland is shown in Figure 1. Northwestern, northeastern, northern and central China had a significantly low I-HFMD in the four seasons, whereas the I-HFMD in southwestern and southern China was relatively high. Furthermore, the I-HFMD of Guangxi, Hainan and Guangdong provinces were always among the highest in the country.
Figure 1. The seasonal distribution of cumulative incidence of hand, foot, and mouth disease (I-HFMD) on the Chinese mainland.

Many studies have shown that socioeconomic and natural environmental factors significantly affect the I-HFMD. This study analyzed six natural environmental factors and two socioeconomic factors as potential influences. The natural environmental factors were temperature (TEMP), precipitation (PREC), humidity (HU), altitude (ALT), wind speed (WS), and air pressure (AP). The socioeconomic factors were child population density (CHD) and per capita GDP (mGDP). The sources and details of the experimental data are shown in Table 1.

The values of TEMP, PREC, HU, AP, and WS were derived from the Chinese regional surface meteorological element-driven dataset of the China Qinghai-Tibet Plateau Scientific Data Center [32,33]. This dataset is based on the international Princeton reanalysis data, Global Land Data Assimilation System (GLDAS) data, Global Energy and Water Exchanges-Surface Radiation Budget (GEWEX-SRB) radiation data and Tropical Rainfall Measuring Mission (TRMM) precipitation data. It merges the conventional meteorological observation data of the China Meteorological Administration. The original data come from meteorological bureau observation data, reanalysis data, and remote sensing data. The values of non-physical range have been removed using ANU-Spline statistical interpolation. The accuracy of this dataset is between the meteorological bureau observation data and satellite remote sensing data, which is better than the international reanalysis data [32,33]. The obtained data were the monthly averages for China. We used the vector data at the
province level in China to calculate the average values in the raster range of each province for the six environmental factors. ETOPO1 is a 1 arc-minute global relief model of Earth’s surface that integrates land topography and ocean bathymetry. ALT was calculated from the ETOPO1 data for each province. Data for the two socioeconomic variables (i.e., CHD and mGDP) were extracted from national statistical yearbooks and are accurate to the provincial administrative unit level.

Table 1. Sources and details of potential risk factors and I-HFMD data.

| Variables               | Data Type        | Variables Name | Time Resolution | Space Resolution | Data Source                                      | Units        |
|-------------------------|------------------|----------------|-----------------|------------------|-------------------------------------------------|--------------|
| Incidence of HFMD       | Statistical yearbook | I-HFMD         | Monthly mean    | -                | Public Health Science Data Center               | 1/100,000    |
| Temperature             | Remote sensing   | TEMP           | Monthly mean    | 1 km             | National Tibetan Plateau Data Center            | Celsius      |
| Amount of precipitation | Remote sensing   | PREC           | Monthly mean    | 1 km             | National Tibetan Plateau Data Center            | milliliter   |
| Specific humidity       | Remote sensing   | HU             | Monthly mean    | 1 km             | National Tibetan Plateau Data Center            | g/kg         |
| Air pressure            | Remote sensing   | AP             | Monthly mean    | 1 km             | National Tibetan Plateau Data Center            | kPa          |
| Wind speed              | Remote sensing   | WS             | Monthly mean    | 1 km             | National Tibetan Plateau Data Center            | m/s          |
| Altitude                | Remote sensing   | ALT            | All years       | 0.1°             | NOAA ETOPO1 Global Relief Model                 | m            |
| Children population     | Statistical yearbook | CHD           | Seasonal average | -                | National Bureau of Statistics of China          | Person/km²   |
| per capita              | Statistical yearbook | mGDP          | Seasonal average | -                | National Bureau of Statistics of China          | Yuan/Person  |

2.2. Research Framework

The research framework of this study was as follows. First, we analyzed the correlation and multicollinearity of the potential risk factors. Second, we adopted Moran’s index to evaluate the spatial autocorrelation in the I-HFMD dataset. Third, we used the GWR model to study the spatial and seasonal heterogeneity of the relationship between the I-HFMD and potential risk factors. During the analysis, it was found that the independent and dependent variables had extreme values, which may have negative effects on the regression models, leaving the results susceptible to deviations. The five variables of PREC, ALT, CHD, mGDP and I-HFMD were therefore logarithmically transformed to reduce the influence of extreme values. Fourth, we optimized the bandwidth parameter of the GWR model and evaluated its accuracy. Fifth, we compared the performance of the GWR model with that of the OLR model in each season to evaluate its applicability and superiority. Sixth, we conducted statistical significance test of the GWR model to examine whether the spatially varying coefficients of the GWR model were statistically significant, and then analyze the influence characteristics of various factors on the I-HFMD over space in each season, providing an explanation and understanding of the spatial pattern and relationship between the I-HFMD and climate change, economic development, and changes in the child population on the Chinese mainland.

2.3. Methods

2.3.1. Correlation Analysis and Multicollinearity Test

The Pearson correlation test was used to explore the correlation of potential risks with the I-HFMD. This analysis was carried out in SPSS v.20 using the bilateral significance test.
The Pearson correlation coefficient between two variables is defined as the quotient of the covariance and standard deviation, as follows:

\[
\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y},
\]

where \(E[(X - \mu_X)(Y - \mu_Y)]\) is the covariance of \(X\) and \(Y\), \(\sigma_X\), \(\sigma_Y\) are the standard deviation of \(X\), \(Y\).

As the multicollinearity of independent variables will seriously affect the experimental results of regression models [34], the variance inflation factor (VIF) was used to measure the severity of multicollinearity in the regression models. The VIF value represents the quotient of the variance in a model with multiple terms by the variance of a model with one term alone, and is expressed as follows:

\[
\text{VIF} = \frac{1}{1 - R_i^2},
\]

where \(R_i\) is the multi-correlation coefficient of \(X_i\) with others variable \(X_j(i \neq j)\).

2.3.2. Spatial Autocorrelation Test

To explore the spatial aggregation effect of HFMD, Moran’s index was adopted to evaluate the spatial autocorrelation in the I-HFMD dataset [35]. The Moran’s index value, along with the z-score and \(p\)-value, helps to evaluate the significance of the index, and is expressed as follows:

\[
I = \frac{n}{S_0} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x}) \frac{(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2},
\]

\[
S_0 = \frac{n}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}},
\]

where \(x_i\) and \(x_j\) are the attribute values of features \(i\) and \(j\), \(\bar{x}\) is the average of \(n\) cells’ attribute values; \(w_{ij}\) is the spatial weight matrix.

2.3.3. Geographically Weighted Regression

The GWR model, extending from the OLR model, was used to study the relationship between potential risk factors and the I-HFMD. OLR is a traditional method of estimating global regression coefficients and does not change the coefficients by geographic location. The regression formula of the classic OLR model is expressed as follows:

\[
I - \text{HFMD}_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \epsilon_i \quad i = 1, 2, \cdots, n,
\]

where \(y_i\) and \(x_{i1}, x_{i2}, x_{i3}, \cdots, x_{ip}\) are the dependent variable and independent variables, respectively; \(\beta_0, \beta_1, \beta_2, \ldots, \beta_p\) are the regressive coefficients; and \(\epsilon_i\) is the error term. GWR extends OLR by allowing local estimates across space, and the form used in this study is the same as in [36–38]:

\[
I - \text{HFMD}_i = \beta_{i0} + \beta_{i1} \times HU_i + \beta_{i2} \times TEMP_i + \beta_{i3} \times PREC_i + \beta_{i4} \times WS_i + \beta_{i5} \times AP_i + \beta_{i6} \times ALT_i + \beta_{i7} \times CHD_i + \beta_{i8} \times mGDP_i + \epsilon_i \quad i = 1, 2, \cdots, n,
\]

where \(\beta_{ik}\) denotes the series coefficients of point \(i\). GWR captures the local effects through the spatially varying coefficients \(\hat{\beta}_i\), calculated as follows:

\[
\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i (I - \text{HFMD})_i,
\]
where \( W_i \) is an \( n \times n \) geographical weighting matrix with the diagonal elements representing non-stationary weights. A certain weight kernel of GWR should be specified to calculate the weight matrix. The widely used fixed Gaussian and adaptive bi-square weighting functions were adopted in our experiments. In addition, GWR was used with the corrected Akaike information criterion (AICc) value as its performance criterion to achieve the best model and was implemented in MATLAB 2013a. Two types of GWR model were tested in this study, as defined in Table 2, and their performance were compared with the OLR model for each season to evaluate the applicability.

### Table 2. The definition of GWR-AAB and GWR-AFG.

| Model Name | Bandwidth Optimization Criteria | Kernel Function |
|------------|---------------------------------|-----------------|
| GWR-AFG    | AICc                            | Fixed Gaussian  |
| GWR-AAB    | AICc                            | Adaptive Bi-square |

It is usually necessary to examine whether the GWR results have significant spatial non-stationarity through diagnostic analysis. A F1-test based on the extremely approximate distribution of the residual square sums proposed by Leung [39] and Wu [40] was adopted to test the significance of the spatial non-stationarity in the GWR results.

Moreover, the local \( R^2 \), which indicates how well the local regression model fits the I-HFMD, was used to represent the local fitting accuracy and was computed as follows [36]:

\[
R^2 = \frac{TSS^w - RSS^w}{TSS^w},
\]

where \( TSS^w \) is the geographically weighted total sum of squares, and \( RSS^w \) is the geographically weighted residual sum of squares.

### Model Evaluation

The following indices were used to evaluate model performance: determination coefficient \( (R^2) \), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and AICc. The formula for each indicator is as follows [36,37,41]:

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2},
\]

RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}},

MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n},

MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100%,

AICc = n \log(e) (\hat{\sigma}^2) + n \log(e) (2\pi) + n \left( \frac{n + \text{tr}(S)}{n - 2 - \text{tr}(S)} \right),

where \( y \) represents the average of the observed values and \( \hat{\sigma}^2 \) is the mean square error of the model.

### 3. Results

#### 3.1. Correlation Analysis and Multicollinearity Test

The results of the correlation analysis and significance test for each potential risk factor in the four seasons are shown in Table 3. All potential risk factors were significantly correlated with the I-HFMD in each season. Specifically, HU, TEMP, PREC, AP, CHD, and
mGDP were positively related to the I-HFMD, whereas ALT was negatively related to the I-HFMD in each season. Notably, WS was positively related to the I-HFMD in spring and winter but negatively related to the I-HFMD in summer and autumn, suggesting significant temporal non-stationarity in the association between WS and the I-HFMD. In addition, CHD had the strongest correlation with the I-HFMD in spring (0.421), summer (0.420), and autumn (0.416), whereas mGDP had the strongest correlation with the I-HFMD in winter (0.402). Multicollinearity among all potential risk variables was tested using VIF values, as shown in Table 4. All VIF values were less than 10, suggesting that the eight potential risks did not show multicollinearity. Therefore, the subsequent regression analysis selected all potential risks as explanatory variables.

Table 3. Pearson correlations between the I-HFMD and eight potential risk factors for each season.

| Independent          | HU   | TEMP | PREC | WS   | AP   | ALT  | CHD  | mGDP |
|----------------------|------|------|------|------|------|------|------|------|
| I-HFMD in Spring     | 0.124** | 0.187** | 0.254** | −0.233** | 0.244** | −0.202** | 0.421** | 0.238** |
| I-HFMD in Summer     | 0.239** | 0.228** | 0.166** | 0.152* | 0.133** | −0.328** | 0.420** | 0.297** |
| I-HFMD in Autumn     | 0.289** | 0.183** | 0.228** | 0.233* | 0.178** | −0.292** | 0.416** | 0.299** |
| I-HFMD in Winter     | 0.285** | 0.378** | 0.148** | −0.168** | 0.234** | −0.206** | 0.359** | 0.402** |

** Stands for significance at the 1% level. * Stands for significance at the 5% level.

Table 4. VIF values of multicollinearity testing of potential risks for each season.

| Independent          | HU   | TEMP | PREC | WS   | AP   | ALT  | CHD  | mGDP |
|----------------------|------|------|------|------|------|------|------|------|
| I-HFMD in Spring     | 6.045 | 5.681 | 2.547 | 2.623 | 3.927 | 3.891 | 5.036 | 1.229 |
| I-HFMD in Summer     | 4.682 | 8.055 | 4.224 | 2.108 | 1.580 | 4.719 | 5.336 | 1.235 |
| I-HFMD in Autumn     | 9.319 | 9.600 | 2.377 | 1.702 | 4.113 | 3.431 | 4.914 | 1.277 |
| I-HFMD in Winter     | 7.936 | 9.805 | 1.965 | 1.580 | 4.150 | 3.523 | 5.959 | 1.218 |

3.2. Diagnosis of Spatial Autocorrelation of the I-HFMD

To investigate the spatial autocorrelation of the I-HFMD, the global Moran’s index of the I-HFMD for each season was calculated by ArcGIS. Moran’s index was 0.89 (p < 0.01) in spring, −0.50 (p < 0.01) in summer, 0.32 (p < 0.01) in autumn, and 0.79 (p < 0.01) in winter, which indicates that the I-HFMD was spatially positively clustered in spring, autumn and winter but spatially negatively clustered in summer. In addition, Moran’s index of the I-HFMD for each season was statistically significant, indicating that there was strong spatial autocorrelation in the incidence rate. Thus, spatial regression models were likely to be needed to determine the relationship between the I-HFMD and potential risk factors.

3.3. Model Performance of OLR and GWR

Table 5 shows the model performance of OLR and GWR for I-HFMD estimation in the four seasons. GWR-AFG and GWR-AAB showed great improvements over the traditional global regression for all statistical indicators, benefiting from its consideration of spatial heterogeneity. The p-values of the F1 test also showed significant spatial non-stationarity in the estimated relationship. In addition, the estimation accuracy of the GWR model with adaptive bi-square kernel (GWR-AAB) was better than that of the GWR model with fixed Gaussian kernel (GWR-AFG) in spring, summer, and winter, but worse than that of GWR-AFG in autumn. This suggests that the adaptive kernel is more suitable than the fixed kernel to characterize the spatially non-stationary relationship between the I-HFMD and the eight potential risk factors, which is most likely a reflection of the complex landform and vast area of the Chinese mainland.
Table 5. Model performance of OLR and GWR for I-HFMD estimation in the four seasons.

| Model               | $R^2$  | RMSE  | MAE   | MAPE  | AICc  | F1    | $p$-Value of F1 | Bandwidth |
|---------------------|--------|-------|-------|-------|-------|-------|-----------------|-----------|
| OLR in Spring       | 60.10% | 1.096 | 0.792 | 27.81 | 2837.6| -     | -               | -         |
| GWR-AFG in Spring   | 73.61% | 0.892 | 0.614 | 20.05 | 2598.3| 0.72  | 0.01            | 37,083.6  |
| GWR-AAB in Spring   | 75.47% | 0.856 | 0.596 | 19.54 | 2584.1| 0.69  | 0.01            | 295.3     |
| OLR in Summer       | 29.09% | 0.851 | 0.656 | 5.07  | 2285.6| -     | -               | -         |
| GWR-AFG in Summer   | 62.94% | 0.615 | 0.476 | 3.21  | 1928.4| 0.60  | 0.01            | 27,641.9  |
| GWR-AAB in Summer   | 68.95% | 0.563 | 0.429 | 2.90  | 1902.5| 0.54  | 0.01            | 146.5     |
| OLR in Autumn       | 39.35% | 0.841 | 0.660 | 10.06 | 2267.5| -     | -               | -         |
| GWR-AFG in Autumn   | 76.29% | 0.523 | 0.393 | 5.19  | 1690.7| 0.46  | 0.01            | 26,182.0  |
| GWR-AAB in Autumn   | 74.31% | 0.547 | 0.416 | 5.77  | 1702.8| 0.48  | 0.01            | 262.2     |
| OLR in Winter       | 56.46% | 1.273 | 0.952 | 47.15 | 2979.6| -     | -               | -         |
| GWR-AFG in Winter   | 76.77% | 0.930 | 0.688 | 33.93 | 2664.3| 0.62  | 0.01            | 30,001.8  |
| GWR-AAB in Winter   | 78.04% | 0.904 | 0.674 | 33.21 | 2640.2| 0.60  | 0.01            | 279.2     |

The better of the two GWR models was chosen in each season for analysis. The estimation accuracy ($R^2$) for the four seasons increased from 60.01% with OLR to 75.47% with GWR-AAB in spring, from 29.09% with OLR to 68.95% with GWR-AAB in summer, from 39.35% with OLR to 76.29% with GWR-AFG in autumn, and from 56.46% with OLR to 78.04% with GWR-AAB in winter. Regarding other evaluation indices, the RMSE, MAE, MAPE, and AICc values of the GWR model were also considerably lower than those of the OLR model in each season.

For the statistically significant seasonal GWR model, the local $R^2$ values at the provincial level are shown in Figure 2. In Figure 2, the local $R^2$ values are distributed in the range 0.23–0.70, and winter and spring show higher local accuracy than summer and autumn. The distributions of local $R^2$ vary over time and space, which demonstrates the comprehensive statistical effect and fit accuracy of the influential risks on the I-HFMD. Most of the provinces of mainland China show robust estimation performance, indicating the well-fitted properties of the GWR model.

Figure 2. Local $R^2$ distributions of the GWR model in each season.
### 3.4. Statistical Significance Test of the GWR Model

The t-test, which indicates whether the spatially varying coefficients of the GWR model were statistically significant, was used to evaluate whether the influence of the potential risk factors on the I-HFMD was significant in each province [36,39]. According to the t-test table, an absolute value of the t-test of over 1.96 is statistically significant. The pass rate proposed by Hong [21], which was defined as the ratio between the number of statistically significant provinces and the total number of provinces, was adopted in our analysis. We calculated the pass rate of all potential risks in the four seasons and the results are shown in Table 6. If the pass rate is zero, it obviously suggests no statistical significance. According to previous studies, the pass rate was considered to be statistically significant if it was more than 35% [21,40]. Of the potential risks, mGDP had the highest pass rate, ranging from 48.4% to 100% in the four seasons. The pass rates of AP and ALT were over 40% in all seasons. The pass rates of HU and TEMP were not statistically significant in summer and autumn, and that of PREC was not statistically significant in spring and autumn. Furthermore, the pass rate of WS was not statistically significant in spring and that of CHD was not statistically significant in summer. Of the four seasons, summer had the lowest statistical significance with an average of 35.2%, whereas winter had the highest statistical significance with an average of 56.9%.

| Season  | Intercept | HU  | TEMP | PREC | WS  | AP  | ALT | CHD | mGDP |
|---------|-----------|-----|------|------|-----|-----|-----|-----|------|
| Spring  | 70.9%     | 41.9%| 38.3%| 22.6%| 35.3%| 61.2%| 58.1%| 41.9%| 64.5%|
| Summer  | 41.9%     | 32.3%| 3.2% | 38.7%| 45.2%| 41.9%| 31.9%| 22.6%| 48.4%|
| Autumn  | 87.10%    | 22.6%| 25.8%| 22.6%| 48.4%| 54.8%| 31.9%| 67.7%| 64.5%|
| Winter  | 54.8%     | 54.8%| 45.1%| 58.1%| 51.6%| 58.1%| 34.1%| 45.1%| 100% |

### 3.5. Seasonal and Spatial Variations of the Relationship between Potential Risk Factors and the I-HFMD

The spatially varying coefficients of the potential risks generated by the GWR model in each season of each province from 2007 to 2017 were averaged to study the overall seasonal variations of their influence on the I-HFMD, as shown in Figure 3. In spring, HU, TEMP, WS, AP, ALT, and mGDP had positive effects on the I-HFMD, which suggests that these variables can promote the incidence of HFMD. CHD had a negative effect on the I-HFMD in spring, suggesting that it can inhibit the incidence of HFMD. PREC had a positive effect on the I-HFMD, but it was not statistically significant (22.6% in Table 6). In summer, PREC, WS, AP, and mGDP had positive influences on the I-HFMD and HU, TEMP, CHD, and ALT had no statistically significant relationship with the I-HFMD. In autumn, WS, AP, and mGDP had positive influences on the I-HFMD, whereas CHD had a negative influence on the I-HFMD. ALT, HU, TEMP, and PREC had no statistically significant relationship with the I-HFMD in autumn (see Table 6). In winter, HU, TEMP, AP, and mGDP had positive effects on the I-HFMD, whereas PREC, WS, and CHD had negative effects on the I-HFMD. ALT had no statistically significant relationship with the I-HFMD in winter (34.1% in Table 6).

Based on the results of the GWR model, we further investigated the seasonal relationship between natural environmental factors, socioeconomic factors, and the I-HFMD in different provinces and analyzed the spatiotemporal heterogeneity of each potential risk factor. The spatial distributions of the coefficients of the potential risk factors are shown in Figures 4–7.
Figure 3. Average coefficients of the potential risk factors in the four seasons.

Figure 4. The spatial variations of the coefficients of the GWR model in spring.

Figure 5. The spatial variations of the coefficients of the GWR model in summer.
Figure 6. The spatial variations of the coefficients of the GWR model in autumn.

Figure 7. The spatial variations of the coefficients of the GWR model in winter.

In spring, mGDP had the greatest influence on the I-HFMD in the southwestern and southeastern regions with a high I-HFMD, as shown in Figure 4. AP showed obvious spatial heterogeneity with inhibitory effects in northeastern, southwestern, and northwestern China but positive effects in the southeastern and northern regions. Furthermore, AP was positively associated with the I-HFMD in higher AP provinces but negatively associated with the I-HFMD in lower AP provinces. The influence of ALT also showed spatial non-stationarity, with an inhibitory effect in northeastern, southwestern, and northwestern China but a positive effect in northern, central, southern, and eastern China. Furthermore, ALT played an inhibitory role in high-altitude provinces and a positive role in low-altitude provinces. The more mountainous high-altitude provinces might isolate people, hampering the spread of HFMD. HU had a positive effect in all provinces and a higher coefficient
in provinces with higher incidence. TEMP had a positive impact in about 70% of the provinces, mostly those with low incidence.

In summer, WS had a statistically significant influence on the I-HFMD in most provinces. WS had a positive effect mainly in northeastern and eastern China, accelerating the spread of HFMD in Inner Mongolia [21]. In addition, mGDP had a positive effect on the I-HFMD in most provinces in summer, which may to some extent be due to high mGDP reflecting more economic activity and social contact and therefore a greater spread of HFMD.

In autumn, CHD had a negative effect in the high-risk provinces in northern, central, and southern China and mGDP had a positive effect in most provinces across the Chinese mainland. AP had a positive effect on the I-HFMD in northern and eastern China, with relatively high AP values compared to other parts of China. This indicates that high AP in autumn can promote the spread of HFMD. WS had a positive effect in 80.6% of the provinces, including Hainan, Guangxi, Guangdong, Fujian, and Shanghai.

In winter, mGDP had a positive effect on the I-HFMD in all provinces. AP positively influenced the I-HFMD in 75% of the provinces and had a positive effect in areas at high risk of HFMD, which may be due to AP being higher in winter than in other seasons; this is consistent with previous studies that have shown that high AP contributes to the spread of HFMD viruses [21]. PREC had a negative effect in all provinces, which may be because low precipitation inhibits the reproduction of HFMD viruses. HU had a positive effect in 90% of the provinces due to relatively moderate HU in winter promoting the reproduction of HFMD viruses. Except for Guangdong province, TEMP had a positive effect on the I-HFMD in other provinces because outdoor activities are inhibited in winter and children can easily transmit the virus to their younger siblings or neighbors indoors [21]. CHD inhibited the I-HFMD in most provinces, and more significantly in high-risk provinces such as Hainan, Guangxi, Yunnan, and Shanghai. WS promoted the I-HFMD in 77% of the provinces.

4. Discussion

In this study, we established a GWR model to quantify the spatial relationship between the I-HFMD and eight socioeconomic and natural environmental factors on the Chinese mainland and then analyzed their impacts in different seasons and provinces. The GWR model successfully described the spatial changes in the influence of potential risks and showed a great improvement in model performance compared with the global regression method. Socioeconomic factors and natural environmental factors were found to be potential contributors to the I-HFMD in most parts of the Chinese mainland. In addition, the intensity and direction of influence between these factors and the I-HFMD were significantly spatially non-stationary at the local geographical level.

The finding that meteorological factors had significant influences on the I-HFMD is consistent with previous studies [1,2,4,6,8,16,17,21,22,24,25,29,38,42–45]. However, the exact mechanism of the association between meteorological parameters and the I-HFMD is still not clear. It is generally assumed that meteorological parameters influence its incidence rate by affecting HFMD transmission [1]. As for each potential risk factor, we found that HU had a positive correlation with the I-HFMD in spring and winter, which is similar to Wang’s finding that HU was positively related to the I-HFMD in Hong Kong and was the most influential factor [44]. In contrast, HU had no statistically significant effect in summer and autumn, when HU is generally higher than in other seasons. Other studies have also been unable to establish a general pattern in relation to higher HU [5,15,16].

TEMP showed influences similar to HU, being generally positively correlated with the I-HFMD in spring and winter but having no statistically significant effect in summer and autumn. There are two potential reasons for the complex association between TEMP and the I-HFMD: (1) there is virological evidence for the temperature-sensitive nature of enteroviruses and other human enteric viruses [12,13], and (2) more outdoor activities in moderately warmer weather increase close contact between individuals and thus increase
HFMD transmission [1]. The studies by Xu et al. [14] in Beijing and Huang et al. [45] in Guangzhou found similar patterns. Furthermore, PREC was positively related to the I-HFMD in summer and negatively related to the I-HFMD in winter, and had no statistically significant effect in spring and autumn. However, studies have found that higher PREC was positively associated with the I-HFMD, including studies in Singapore [46], Hong Kong [44], and mainland China [2,3,17,18,22,47,48]. A possible reason is that PREC in China is highest in summer and lowest in winter.

WS was generally positively correlated with the I-HFMD [1]. However, although WS was positively associated with the I-HFMD in spring, summer, and autumn, it was negatively associated with the I-HFMD in winter, mainly because WS is higher in winter than in other seasons. Suya Yi [19] came to similar conclusions with WS generally positively correlated with the I-HFMD in Beijing but showing a negative correlation with the I-HFMD as it increased. AP was found to typically present spatial non-stationarity in all seasons, with AP being positively related to the I-HFMD in higher AP provinces and negatively related to the I-HFMD in lower AP provinces. This conclusion is the opposite to that of a study at the local scale, which revealed a 6.8% drop in cases of HFMD for every 1 hPa increase in AP in Guangdong [4]. ALT positively affected the I-HFMD in spring and had no statistically significant effect in other seasons. This may be due to ALT usually playing a special and indirect role in the influence of I-HFMD: for example, Xiong Xiao found that ALT may have modification effects on the TEMP–HFMD relationship in a multi-city study on the Chinese mainland [18].

In our study, we also took two socioeconomic factors into account to analyze their impacts in different seasons and provinces. CHD had a negative correlation with the I-HFMD in spring, autumn, and winter and no statistically significant effect in summer. There is no consensus in the literature over the influence of CHD on the I-HFMD. Studies have found that the strength [22] and direction of the association between CHD and the I-HFMD have obvious spatial non-stationarity at the local geographical level. In some areas, CHD was not significantly related to the I-HFMD, whereas in other areas, it was positively related to the I-HFMD or had an inverse association with the I-HFMD of a different strength. In addition, mGDP was found to have a positive correlation with the I-HFMD in all seasons, which is similar to a previous finding that mGDP was positively associated with disease incidence [23]. Gou also found that mGDP and CHD had a significant effect on HFMD transmission [24]. Moreover, urban and urban–rural border areas with a much higher mGDP value have been found to be at high risk of occurrence of HFMD due to the frequent migration of people to these areas [25,26].

5. Conclusions

A correct understanding of the spatial relationship between natural environmental and socioeconomic factors and the I-HFMD is important for health care workers to formulate reasonable epidemic prevention and control measures. This study provides a new perspective to explore the main spatiotemporal patterns and potential risk factors of HFMD, using the GWR model to study the seasonal influence characteristics of potential risk factors on the I-HFMD at a large geospatial range and over a long period. Based on the results of the GWR model, we provide an explanation and understanding of the spatial pattern and relationships between the I-HFMD and climate change, economic development and changes in the child population in the Chinese mainland. The findings of this study may help predict the risk of HFMD in children in different seasons and regions based on climatic, economic and demographic information in China.

However, there are some limitations to this study. First, medical resources, such as the number of doctors and hospital beds per capita, may have a significant impact on the risk of HFMD. We did not analyze this effect as we do not currently have suitable medical resource data. Second, the combined effects of the two environmental factors were not analyzed in this study. Studies have shown that different combinations of climatic factors may produce different types of climate, thus affecting the infection and spread of HFMD.
These considerations should be addressed in future studies. Despite these limitations, we believe that our findings can help to understand the seasonal and spatially relevant effects of socioeconomic and natural environmental factors on the I-HFMD, which is conducive to developing effective prevention strategies for different locations and different seasons.

**Author Contributions:** Jingtao Sun was involved in the design of the study, interpretation of data, drafted the major revisions, and performed the experiments; Sensen Wu contributed to the study design and algorithm improvement; Zhen Yan drafted part of the manuscript; Yadong Li conceived the experiments and improved the manuscript; Cheng Yan improved the manuscript; Zhenhong Du was involved in the data acquisition and analyses of the data and experiments; and Feng Zhang was involved in the revision of the manuscript. Renyi Liu improved the manuscript; All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China [No. 41922043, No. 41871287, No. 42001323], National Key Research and Development Program of China [No. 2018YFB0505000].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data, models, and code that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors would like to thank the support from our university and laboratory and thank all anonymous reviewers for their constructive comments that improved the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Li, L.; Qiu, W.; Xu, C.; Wang, J. A spatiotemporal mixed model to assess the influence of environmental and socioeconomic factors on the incidence of hand, foot and mouth disease. *BMC Public Health* **2018**, *18*, 274. [CrossRef] [PubMed]

2. Ji, T.; Han, T.; Tan, X.; Zhu, S.; Yan, D.; Yang, Q.; Song, Y.; Cui, A.; Zhang, Y.; Mao, N. Surveillance, epidemiology, and pathogen spectrum of hand, foot, and mouth disease in mainland of China from 2008 to 2017. *Biosaf. Health* **2019**, *1*, 32–40. [CrossRef]

3. Zhuang, D.; Hu, W.; Ren, H.; Ai, W.; Xu, X. The influences of temperature on spatiotemporal trends of hand-foot-and-mouth disease in mainland China. *Int. J. Environ. Health Res.* **2014**, *24*, 1–10. [CrossRef] [PubMed]

4. Du, Z.; Lawrence, W.R.; Zhang, W.; Zhang, D.; Yu, S.; Hao, Y. Interactions between climate factors and air pollution on daily HFMD cases: A time series study in Guangdong, China. *Sci. Total Environ.* **2019**, *650*, 1358–1364. [CrossRef] [PubMed]

5. Zhang, Z.; Xie, X.; Chen, X.; Li, Y.; Lu, Y.; Mei, S.; Liao, Y.; Lin, H. Short-term effects of meteorological factors on hand, foot and mouth disease among children in Shenzhen, China: Non-linearity, threshold and interaction. *Sci. Total Environ.* **2016**, *539*, 576–582. [CrossRef]

6. Guo, C.; Yang, J.; Guo, Y.; Ou, Q.; Shen, S.; Ou, C.; Liu, Q. Short-term effects of meteorological factors on pediatric hand, foot, and mouth disease in Guangdong, China: A multi-city time-series analysis. *BMC Infect. Dis.* **2016**, *16*, 524. [CrossRef]

7. Liao, J.; Qin, Z.; Zuo, Z.; Yu, S.; Zhang, J. Spatial-temporal mapping of hand foot and mouth disease and the long-term effects associated with climate and socio-economic variables in Sichuan Province, China from 2009 to 2013. *Sci. Total Environ.* **2016**, *563*, 152–159. [CrossRef]

8. Xu, C.; Zhang, X.; Xiao, G. Spatiotemporal decomposition and risk determinants of hand, foot and mouth disease in Henan, China. *Sci. Total Environ.* **2019**, *657*, 509–516. [CrossRef]

9. Huang, X.; Wei, H.; Wu, S.; Du, Y.; Liu, L.; Su, J.; Xu, Y.; Wang, H.; Li, X.; Wang, Y. Epidemiological and etiological characteristics of hand, foot and mouth disease in Henan, China, 2008–2013. *Sci. Rep.* **2015**, *5*, 1–9. [CrossRef]

10. Kou, Z.; Jia, J.; Liu, X.; Luo, T.; Xin, X.; Gong, J.; Zhang, J.; Sun, D.; Jiang, F.; Gao, R. Epidemiological characteristics and spatial-temporal clusters of hand, foot, and mouth disease in Qingdao City, China, 2013-2018. *PLoS ONE* **2020**, *15*, e233914. [CrossRef]

11. Kulldorff, M.; Heffernan, R.; Hartman, J.; Assunção, R.; Mostashari, F. A space–time permutation scan statistic for disease outbreak detection. *PLoS Med.* **2005**, *2*, e59. [CrossRef] [PubMed]

12. Kung, Y.; Huang, S.; Kuo, P.; Kiang, D.; Ho, M.; Liu, C.; Yu, C.; Su, I.; Wang, J. Introduction of a strong temperature-sensitive phenotype into enterovirus 71 by altering an amino acid of virus 3D polymerase. *Virology* **2010**, *396*, 1–9. [CrossRef]

13. Rzeżutka, A.; Cook, N. Survival of human enteric viruses in the environment and food. *FEMS Microbiol. Rev.* **2004**, *28*, 441–453. [CrossRef]

14. Xu, M.; Yu, W.; Tong, S.; Jia, L.; Liang, F.; Pan, X. Non-linear association between exposure to ambient temperature and children’s hand-foot-and-mouth disease in Beijing. *PLoS ONE* **2015**, *10*, e126171. [CrossRef]
42. Zheng, S.; Wang, M.; Wang, S.; Shang, K.; Hu, L.; Dong, J. Effect of meteorological conditions on occurrence of hand, foot and mouth disease in Wuwei City, Northwestern China. In Proceedings of the 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI), Shanghai, China, 15–17 October 2011; pp. 1893–1895.

43. Dong, W.; Yang, P.; Liao, H.; Wang, X.; Wang, Q. The effects of weather factors on hand, foot and mouth disease in Beijing. Sci. Rep. 2016, 6, 1–9. [CrossRef]

44. Wang, P.; Goggins, W.B.; Chan, E.Y. Hand, foot and mouth disease in Hong Kong: A time-series analysis on its relationship with weather. PLoS ONE 2016, 11, e161006. [CrossRef] [PubMed]

45. Huang, Y.; Deng, T.; Yu, S.; Gu, J.; Huang, C.; Xiao, G.; Hao, Y. 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. 2013, 13, 134. [CrossRef] [PubMed]

46. World Health Organization. Using Climate to Predict Infectious Disease Epidemics; WHO: Geneva, Switzerland, 2005.

47. Wang, C.; Li, X.; Zhang, Y.; Xu, Q.; Huang, F.; Cao, K.; Tao, L.; Guo, J.; Gao, Q.; Wang, W.; et al. Spatiotemporal Cluster Patterns of Hand, Foot, and Mouth Disease at the County Level in Mainland China, 2008–2012. PLoS ONE 2016, 11, e0147532. [CrossRef] [PubMed]

48. Huang, J.; Wang, J.; Bo, Y.; Xu, C.; Hu, M.; Huang, D. Identification of health risks of hand, foot and mouth disease in China using the geographical detector technique. Int. J. Environ. Res. Public Health 2014, 11, 3407–3423.