


Article

Analysis of HIV/AIDS Epidemic and Socioeconomic Factors in Sub-Saharan Africa

Shuman Sun, Zhiming Li , Huiguo Zhang, Haijun Jiang and Xijian Hu *

College of Mathematics and System Science, Xinjiang University, Urumqi 830046, China; Sunshuman2222@163.com (S.S.); zml@xju.edu.cn (Z.L.); zhang_huiguo@163.com (H.Z.); jianghaixju@163.com (H.J.)

* Correspondence: xijianhu@xju.edu.cn

Received: 21 September 2020; Accepted: 26 October 2020; Published: 29 October 2020



Abstract: Sub-Saharan Africa has been the epicenter of the outbreak since the spread of acquired immunodeficiency syndrome (AIDS) began to be prevalent. This article proposes several regression models to investigate the relationships between the HIV/AIDS epidemic and socioeconomic factors (the gross domestic product per capita, and population density) in ten countries of Sub-Saharan Africa, for 2011–2016. The maximum likelihood method was used to estimate the unknown parameters of these models along with the Newton–Raphson procedure and Fisher scoring algorithm. Comparing these regression models, there exist significant spatiotemporal non-stationarity and auto-correlations between the HIV/AIDS epidemic and two socioeconomic factors. Based on the empirical results, we suggest that the geographically and temporally weighted Poisson autoregressive (GTWPAR) model is more suitable than other models, and has the better fitting results.

Keywords: HIV/AIDS epidemic; regression model; Newton–Raphson procedure; Fisher scoring algorithm; time series

1. Introduction

Acquired immunodeficiency syndrome (AIDS) is a malignant infectious disease with a high fatality rate caused by human immunodeficiency virus (HIV). The HIV/AIDS epidemic has been one of the greatest global public health and social development problems since 1981, particularly in Sub-Saharan Africa. As of 31 December 2016, over 30 million people had died from the disease [1]. More than 70% of the 35 million people are infected with the HIV/AIDS disease in Sub-Saharan Africa. Thus, the HIV/AIDS epidemic of Sub-Saharan Africa has attracted extensive attention from researchers around the world [2–4].

In earlier studies, Janet et al. [5] and Hallman et al. [6] demonstrated the relationship between the disease and socioeconomic status. Chris et al. [7] indicated socioeconomic factors to explain this disease outperformed cultural ones in South Africa. Mathematical models always play an important role in evaluating the trends of the HIV/AIDS epidemic [8]. For example, regression models have been widely used in the study of the relationship between this disease and influencing factors. Shiboski et al. [9] considered a generalized linear model to obtain the statistical analysis of the HIV/AIDS disease. A mixed-effects linear regression model was used to analyze the correlation between national population and antenatal care [10]. Laurence et al. [11] applied a spatial regression model to show that the epidemic had substantial geographic variance across Sub-Saharan Africa.

This paper proposes several regressive models to investigate the relationships between the HIV/AIDS epidemic, the gross domestic product (GDP) per capita and the population density in ten countries of Sub-Saharan Africa. The Poisson regression model is introduced in Section 2.1. Sections 2.2 and 2.3 describe two spatial models, respectively. A spatiotemporal autoregressive model

is proposed in Section 2.4. The maximum likelihood method is used to obtain the iterative formulas of coefficient estimations in Section 3. The main results are shown in Section 4, followed by discussion in Section 5.

2. Methodologies

2.1. Poisson Regression Model

Regression models are a set of statistical processes for estimating the relationships between response and explanatory variables. The classical model is a linear regression. Nelder and Wedderburn [12] extended the linear model to a generalized linear regression for solving the discrete data problem. This kind of models are very important in ecology, medicine and economics [13–15]. Suppose that $Y = (Y_1, Y_2, \dots, Y_n)$ is the response variable, where $Y_i (i = 1, \dots, n)$ are independent. The density function is

$$f(y_i; \theta_i, \phi_i) = \exp\left(\frac{y_i \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i)\right),$$

where $a(\cdot), b(\cdot), c(\cdot, \cdot)$ are known functions, and θ_i, ϕ_i are unknown parameters for $i = 1, 2, \dots, n$. Denote $\mu_i = E(Y_i)$, and $g(\mu_i) = \ln(\mu_i)$ is a link function. Let X_{ij} be explanatory variables for the i th observation in the j th variable. Then, the Poisson regression (PR) model is given by

$$g(\mu_i) \triangleq \eta_i = \sum_{j=1}^p \beta_j X_{ij}, \quad (1)$$

where $i = 1, 2, \dots, n$, and $\beta_j (j = 1, 2, \dots, p)$ are unknown parameters.

2.2. Geographically Weighted Poisson Regression Model

With in-depth study, regression models have been frequently applied in epidemiology and health geography for trying to investigate the persistent geographical variations in disease [16]. Based on the generalized linear regression, Brunson et al. [17] proposed the geographically weighted regression model to analyze the spatial non-stationary processes of discrete data. The disease maps arising from this process are considered through the establishment of the geographically weighted Poisson regression (GWPR) model [18–20] below

$$g(\mu_i) \triangleq \eta_i = \sum_{j=1}^p \beta_j(u_i, v_i) X_{ij}, \quad (2)$$

where $(u_i, v_i) (i = 1, 2, \dots, n)$ are the geographical locations, and $\beta_j(u_i, v_i) (j = 1, 2, \dots, p)$ are unknown parameters at the position (u_i, v_i) .

2.3. Geographically Weighted Poisson Autoregressive Model

Another issue deserving of special attention is whether there exists an interaction between different regions in terms of spatial data. Previous studies [21–24] showed that spatial data has not only spatial non-stationarity but also correlation. Zhang [25] proposed the geographically weighted Poisson autoregressive (GWPAR) model as follows:

$$g(\mu_i) \triangleq \eta_i = \rho \sum_{k=1}^n c_{ik} \eta_k + \sum_{j=1}^p \beta_j(u_i, v_i) X_{ij}, \quad (3)$$

where ρ is a scalar autoregressive parameter, and $c_{ik}(i, k = 1, 2, \dots, n)$ is the adjacency relation between the i th and k th locations. Let c^i be the number of regions adjacent to the i th position. If the k th position is next to the i th's, then $c_{ik} = 1/c^i$. Otherwise, $c_{ik} = 0$.

2.4. Geographically and Temporally Weighted Poisson Autoregressive Model

Recently, many spatiotemporal models have been proposed to describe the spatiotemporal variations in the relationships of response and explanatory variables [26,27]. Concerning the modeling of spatiotemporal data, there are two important properties: non-stationarity and auto-correlation. The non-stationarity indicates that there exists more than one linear relation between response and explanatory variables. It can be used to identify where interesting relationships are likely to occur or where detailed investigation is necessary in the study areas [28]. Spatiotemporal auto-correlation is an important factor to determine the temporal correlations of observations [29]. These two problems always appeared together [30]. A geographically and temporally weighted autoregressive (GTWPAR) model can be applied to account for non-stationary and auto-correlated effects simultaneously.

Let Y be the response variable, and $Y_{ik}(i = 1, 2, \dots, n_k, k = 1, 2, \dots, T)$ be the independent variables of Y in the i th position and the k th time. The density function can be defined as follows:

$$f(y_{ik}; \theta_{ik}, \phi_{ik}) = \exp\left(\frac{y_{ik}\theta_{ik} - b(\theta_{ik})}{a(\phi_{ik})} + c(y_{ik}, \phi_{ik})\right),$$

where the parameters are similar to Section 2.1. Denote $\mu_{ik} = E(Y_{ik})$, and $g(\mu_{ik}) = \ln(\mu_{ik})$. Let $X_{ijk}(j = 1, 2, \dots, p)$ be the j th explanatory variable. The GTWPAR model is expressed by

$$g(\mu_{ik}) \triangleq \eta_{ik} = \rho \sum_{m=1}^T \sum_{l=1}^{n_k} c_{lm}^{(ik)} \eta_{lm} + \sum_{j=1}^p \sum_{k=1}^T \beta_{jk}(u_{ik}, v_{ik}, t_k) X_{ijk}, \tag{4}$$

where $\{\beta_{jk}(u_{ik}, v_{ik}, t_k)\}$ is a set of unknown parameters at the i th position in the k th time, and $c_{lm}^{(ik)}$ is the adjacent relation between the location (u_{ik}, v_{ik}, t_k) and (u_{lm}, v_{lm}, t_m) . Following the work of [31], the spatiotemporal distance between the locations (u_{ik}, v_{ik}, t_k) and (u_{lm}, v_{lm}, t_m) can be defined as

$$d_{lm}^{(ik)} = \sqrt{\lambda[(u_{ik} - u_{lm})^2 + (v_{ik} - v_{lm})^2] + \mu(t_k - t_m)^2},$$

where μ and λ are used to balance spatiotemporal distances. Suppose that

$$c_{lm}^{(ik)} = \begin{cases} 1/c^{ik}, & 0 < d_{lm}^{(ik)} < d, \\ 0, & \text{otherwise,} \end{cases}$$

where d is a constant and satisfies $\min\{d_{lm}^{(ik)}\} < d < \max\{d_{lm}^{(ik)}\}$.

Next, we rewrite the model (4) in a matrix form

$$\eta = \rho C \eta + \mathbf{B}' X',$$

where $\eta = (\eta_{11}, \dots, \eta_{n_1 1}, \eta_{12}, \dots, \eta_{n_2 2}, \dots, \eta_{1T}, \dots, \eta_{n_T T})'$, $C = (c_{lm}^{(ik)})$, $X = (X_{ijk})$ and $\mathbf{B} = (\beta_{jk}(u_{ik}, v_{ik}, t_k))$. For convenience, define η_K as the K th element of η ; C_{IK} and X_{IK} are the I th row and the K th column of the matrices C and X , respectively. The detailed expressions of C , X and \mathbf{B} are given in Appendix A.1.

Remark 1. For the GTWPAR model (4), if $\rho = 0$ and $\beta_{jk}(u_{ik}, v_{ik}, t_k)$ is independent of the spatiotemporal effect, the model is a PR model. If $\rho = 0$ and $\beta_{jk}(u_{ik}, v_{ik}, t_k)$ is dependent on spatial effect but independent of temporal effect, the model becomes GWPR model. If $\rho \neq 0$ and $\beta_{jk}(u_{ik}, v_{ik}, t_k)$ is independent of temporal effect, it is the GWPAR model. Thus, PR, GWPR and GWPAR models are the special cases of the GTWPAR model.

3. Coefficient Estimation

In this section, we only provide the estimation method of the GTWPAR model since the PR, GWPR and GWPARG models are its special cases (Remark 1). Let $(u_{ik}, v_{ik}, t_k) (i = 1, 2, \dots, n_k, k = 1, 2, \dots, T)$ be any point in the studied spatiotemporal region. We fix a point (u_{00}, v_{00}, t_0) and assume that $\beta_{jk}(u_{ik}, v_{ik}, t_k) \approx \beta_{j0}(u_{00}, v_{00}, t_0) (j = 1, 2, \dots, p)$. Then, the model (4) can be rewritten by

$$\eta_{ik} = g(\mu_{ik}) = \rho \sum_{m=1}^T \sum_{l=1}^{n_k} c_{lm}^{(ik)} \eta_{lm} + \sum_{j=1}^p \sum_{k=1}^T \beta_{j0}(u_{00}, v_{00}, t_0) X_{ijk}. \tag{5}$$

Denote $\beta(u_{00}, v_{00}, t_0) = (\beta_{10}, \dots, \beta_{p0})'$, $\mathbf{X} = \text{diag}(X_i)$ and $X_i = (X_{i1}, \dots, X_{ip})$. The corresponding matrix form can be represented as $\eta = \rho C \eta + \beta'(u_{00}, v_{00}, t_0) \mathbf{X}'$.

3.1. Estimation of Parameter Vector β

For the fixed point (u_{00}, v_{00}, t_0) , we define a spatiotemporal distance $d_{ik}^{(0)}$ from this point to (u_{ik}, v_{ik}, t_k) as $d_{ik}^{(0)} = \sqrt{\lambda[(u_{00} - u_{ik})^2 + (v_{00} - v_{ik})^2] + \mu(t_0 - t_k)^2}$. The Gauss kernel function of these two points can be written by

$$\begin{aligned} w_{ik}(u_{00}, v_{00}, t_0) &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{d_{ik}^{(0)}}{h_{ST}} \right)^2 \right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \frac{\lambda[(u_{00} - u_{ik})^2 + (v_{00} - v_{ik})^2] + \mu(t_0 - t_k)^2}{h_{ST}^2} \right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(u_{00} - u_{ik})^2 + (v_{00} - v_{ik})^2}{h_S^2} + \frac{(t_0 - t_k)^2}{\tau h_S^2} \right) \right\}, \end{aligned}$$

where h_{ST} and h_S are the space-time bandwidth and space bandwidth, respectively. Meanwhile, we have $h_{ST}^2 = \lambda h_S^2$, and $\tau = \lambda/\mu$ is a spatiotemporal factor. Without loss of generality, let $\lambda = 1$. Then, the weighted maximum likelihood of $Y_{ik} (i = 1, 2, \dots, n_k, k = 1, 2, \dots, T)$ at the point (u_{00}, v_{00}, t_0) is

$$L(\beta_{10}, \beta_{20}, \dots, \beta_{p0}) = \prod_{k=1}^T \prod_{i=1}^{n_k} f(y_{ik}; \theta_{ik}, \phi_{ik}) w_{ik}(u_{00}, v_{00}, t_0),$$

where $f(y_{ik}; \theta_{ik}, \phi_{ik})$ is the density function. The log-likelihood can be obtained as follows:

$$\mathbf{L}_1(\beta(u_{00}, v_{00}, t_0)) = \sum_{k=1}^T \sum_{i=1}^{n_k} \left(\frac{y_{ik} \theta_{ik} - b(\theta_{ik})}{a(\phi_{ik})} + c(y_{ik}, \phi_{ik}) \right) w_{ik}(u_{00}, v_{00}, t_0).$$

Note that $c(y_{ik}, \phi_{ik}) = -\ln(y_{ik}!)$, $b(\theta_{ik}) = \mu_{ik} = \exp(\theta_{ik})$, and $a(\phi_{ik}) = \phi_{ik} = 1$. Thus, $E(Y_{ik}) = b'(\theta_{ik}) = \exp(\theta_{ik}) = \mu_{ik}$, $Var(Y_{ik}) = b''(\theta_{ik}) a(\phi_{ik}) = \exp(\theta_{ik}) = \mu_{ik}$. Differentiating \mathbf{L}_1 with respect to $\beta(u_{00}, v_{00}, t_0)$ yields

$$\frac{\partial \mathbf{L}_1}{\partial \beta_{r0}} = \sum_{k=1}^T \sum_{i=1}^{n_k} \left(\frac{y_{ik} - \mu_{ik}}{a_{ik} \phi} \frac{\partial \theta_{ik}}{\partial \beta_{r0}} \right) w_{ik}(u_{00}, v_{00}, t_0) = 0, \tag{6}$$

where $\beta_{r0} = \beta_r(u_{00}, v_{00}, t_0) (r = 1, 2, \dots, p)$, and

$$\frac{\partial \theta_{ik}}{\partial \beta_{r0}} = \left(\frac{\partial \mu_{ik}}{\partial \theta_{ik}} \right)^{-1} \frac{\partial \mu_{ik}}{\partial g(\mu_{ik})} \frac{\partial g(\mu_{ik})}{\partial \beta_{r0}} = \frac{1}{b''(\theta_{ik})} \frac{1}{g'(\mu_{ik})} \frac{\partial \eta_{ik}}{\partial \beta_{r0}}.$$

For convenience, let $N = \sum_{k=1}^N n_k$ and $W = (w_{ik}(u_{00}, v_{00}, t_0))_{N \times N}$. Denote $A = (I_N - \rho C)^{-1}$, $Y = (Y_{11}, \dots, Y_{n_1 1}, \dots, Y_{1T}, \dots, Y_{n_T T})'$, $\mu = (\mu_{11}, \dots, \mu_{n_1 1}, \dots, \mu_{1T}, \dots, \mu_{n_T T})'$, $\theta = (\theta_{11}, \dots, \theta_{n_1 1}, \dots, \theta_{1T}, \dots, \theta_{n_T T})'$, $\phi = (\phi_{11}, \dots, \phi_{n_1 1}, \dots, \phi_{1T}, \dots, \phi_{n_T T})'$. Suppose that Y_K, μ_K, θ_K and ϕ_K are the K th elements of Y, μ, θ and ϕ , respectively. Then, we take the derivative of the model (5) with respect to β_{r0} , and obtain

$$\frac{\partial \eta_l}{\partial \beta_{r0}} = \sum_{h=1}^N A_{lh} X_{hr} = A_l \cdot X_r, \quad l = 1, 2, \dots, N.$$

The calculation process is given in Appendix A.2. Thus, the Equation (6) can be rewritten as

$$\frac{\partial \mathbf{L}_1}{\partial \beta_{r0}} = \frac{1}{\phi} \sum_{l=1}^N T_l A_l \cdot X_r (Y_l - \mu_l) g'(\mu_l) W_l(u_{00}, v_{00}, t_0) = 0.$$

However, there is not a close-form solution for $\beta(u_{00}, v_{00}, t_0)$. The Newton–Raphson procedure and Fisher scoring algorithm are used to get the estimation of β . The iterative formula is expressed as

$$\begin{aligned} \hat{\beta}^{(m+1)}(u_{00}, v_{00}, t_0) &= \hat{\beta}^{(m)}(u_{00}, v_{00}, t_0) + \mathbf{I}^{-1}(\hat{\beta}^{(m)}(u_{00}, v_{00}, t_0)) S(\hat{\beta}^{(m)}(u_{00}, v_{00}, t_0)) \\ &= ((A^{(m)} X)' T^{(m)} W(u_{00}, v_{00}, t_0) (A^{(m)} X))^{-1} \\ &\quad \times (A^{(m)} X)' T^{(m)} W(u_{00}, v_{00}, t_0) Z^{(m)}, \end{aligned} \tag{7}$$

where the Fisher information matrix $\mathbf{I}(\beta) = E(I(\beta))$, and

$$S(\hat{\beta}^{(m)}(u_{00}, v_{00}, t_0)) = \left(\frac{\partial \mathbf{L}_1}{\partial \beta_{10}}, \frac{\partial \mathbf{L}_1}{\partial \beta_{20}}, \dots, \frac{\partial \mathbf{L}_1}{\partial \beta_{p0}} \right)'$$

is the scalar vector. The detail process is provided in Appendix A.2. For the fixed point $(u_{ik}, v_{ik}, t_k) (i = 1, 2, \dots, n_k; k = 1, 2, \dots, T)$, $\hat{\beta}_{jk}(u_{ik}, v_{ik}, t_k)$ can be obtained by (7).

Remark 2. The estimations $\hat{\beta}(u_{ik}, v_{ik}, t_k) (i, l = 1, 2, \dots, n_k, k, m = 1, 2, \dots, T)$ are related to the temporal and spatial effects in the GTWPAR model. If $m \neq k, w_{ik}(u_{lm}, v_{lm}, t_m) = 0$ and $c_{lm}^{(ik)} = 0$, then $\hat{\beta}(u_{ik}, v_{ik}, t_k) = \hat{\beta}(u_i, v_i)$ correspond to the parameter estimations of the GWPAR model. If $w_{ik}(u_{lm}, v_{lm}, t_m) = 0 (m \neq k)$ and $C = \mathbf{0}$, they are the estimations of the GWPR model. If $W = \mathbf{0}$ and $C = \mathbf{0}$, then $\hat{\beta}(u_{ik}, v_{ik}, t_k) = \hat{\beta}$ are the global estimation values of the PR model.

3.2. Estimation of Parameter ρ

Based on the density function, the log-likelihood function of ρ is

$$\mathbf{L}_2(\rho) = \sum_{k=1}^T \sum_{i=1}^{n_k} \left(\frac{y_{ik} \theta_{ik} - b(\theta_{ik})}{a(\phi_{ik})} + c(y_{ik}, \phi_{ik}) \right).$$

Differentiating $\mathbf{L}_2(\rho)$ with respect to ρ , we have

$$\frac{\partial \mathbf{L}_2}{\partial \rho} = \sum_{k=1}^T \sum_{i=1}^{n_k} \left(\frac{y_{ik} - \mu_{ik}}{a_{ik} \phi} \frac{\partial \theta_{ik}}{\partial \rho} \right) = 0, \tag{8}$$

where $\frac{d\theta_{ik}}{d\rho} = \frac{1}{b''(\theta_{ik})g'(\mu_{ik})} \frac{d\eta_{ik}}{d\rho}$. Then, we take the derivative of the model (5) with respect to ρ as follows:

$$\frac{d\eta_l}{d\rho} = \frac{dg(\mu_l)}{d\rho} = \sum_{h=1}^N A_l \cdot C_{.h} \eta_h.$$

The detail calculation is given in Appendix A.3. Then, the Equation (8) can be rewritten in the following nonlinear form

$$\frac{dL_2}{d\rho} = \sum_{l=1}^N \frac{(Y_l - \mu_l) \sum_{h=1}^N A_l C_h \eta_h}{a_l \phi V(\mu_l) g'(\mu_l)} = 0.$$

According to the Newton–Raphson procedure and Fisher scoring algorithm, the iterative formula of $\hat{\rho}^{(m+1)}$ is

$$\begin{aligned} \hat{\rho}^{(m+1)} &= \hat{\rho}^{(m)} + \mathcal{I}^{-1}(\hat{\rho}^{(m)}) S(\hat{\rho}^{(m)}) \\ &= \hat{\rho}^{(m)} + ((A^{(m)} C \eta^{(m)})' T^{(m)} (A^{(m)} C \eta^{(m)}))^{-1} \\ &\quad \times (A^{(m)} C \eta^{(m)})' T^{(m)} (Z^{(m)} - \eta^{(m)}), \end{aligned} \tag{9}$$

where the scalar vector $S(\hat{\rho}^{(m)}) = \frac{1}{\phi} (AC\eta)' T (Z - \eta)$ and the Fisher information matrix $\mathcal{I}(\rho) = \frac{1}{\phi} (AC\eta)' T (AC\eta)$. The calculation process of the scalar vector $S(\hat{\rho}^{(m)})$ and the information matrix \mathcal{I} is given in Appendix A.3.

4. Main Results

In this section, we apply the PR, GWPR, GWPARG and GTWPARG models to analyze the relationships between the HIV/AIDS epidemic, the GDP per capita and population density in ten countries of Sub-Saharan Africa from 2011 to 2016. The ten countries are Angola, Botswana, Lesotho, Malawi, Mozambique, Namibia, South Africa, Swaziland, Zimbabwe and Zambia. The parameters of these four models are estimated by the Newton–Raphson procedure and Fisher scoring algorithm. The coefficient of determination R^2 , the corrected Akaike information criterion (AICc), the deviation (D) and mean-square error (MSE) are used to compare the performances of the four models [18].

4.1. The HIV/AIDS Epidemic Models

The data of HIV/AIDS incidence, GDP per capita and population density were derived from <http://data.cnki.net/InternationalData/Report>. Readers should note that authorization is required to access the database on this website. Figure 1 describes the HIV/AIDS incidence in ten countries from 2011 to 2016. It shows that the incidence varies significantly in different regions. Angola has a minimum incidence of less than 5%, while Botswana and Swaziland have higher incidences of more than 20% every years. Therefore, it may be necessary to consider the temporal and spatial factors in analyzing the HIV/AIDS epidemic.

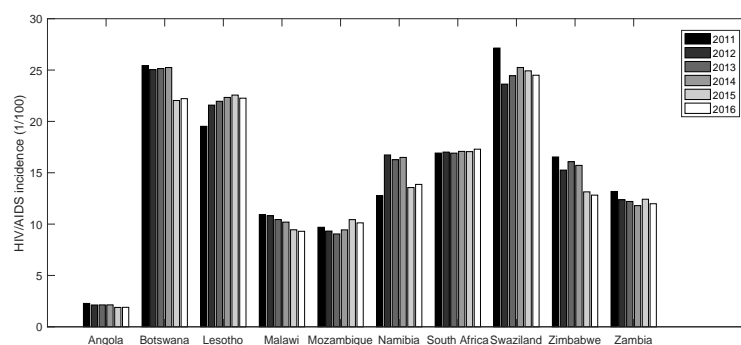


Figure 1. Spatiotemporal HIV/AIDS incidence of ten countries, 2011–2016.

The distributions of HIV/AIDS cases, GDP per capita and population density are displayed in Figure 2. The Pearson correlation coefficients between these cases and GDP per capita and population density are 0.2739 and -0.1179 , respectively. Meanwhile, the two socioeconomic factors have different effects on the HIV/AIDS cases at the spatiotemporal locations. These reflect a spatiotemporal

non-stationarity between the cases and two factors in ten countries from 2011 to 2016. Table 1 lists the p -values of the first-order autocorrelation of HIV/AIDS cases in the different years of the same region or the different regions of the same year. Each region has a significant spatial autocorrelation (p -value < 0.01) each year. Lesotho and South Africa had temporal autocorrelation during 2011 to 2016. Thus, the spatial and temporal autocorrelation should not be ignored.

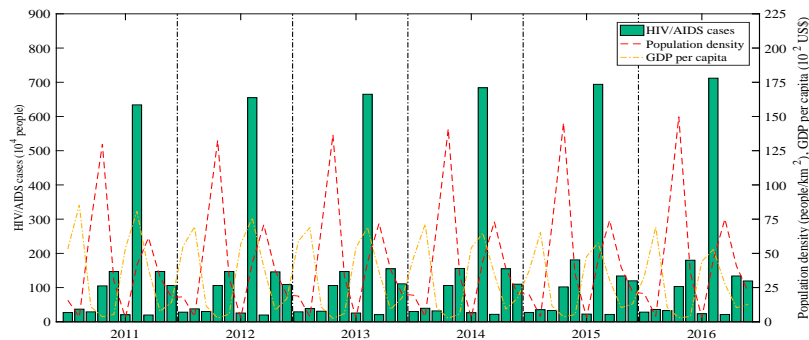


Figure 2. Distributions of HIV/AIDS cases, GDP per capita and population density of ten countries, 2011–2016.

Table 1. p -values of the spatial and temporal autocorrelation analysis.

Time	2011	2012	2013	2014	2015	2016
p -value	0.0024	0.0028	0.0021	0.0019	0.0017	0.0022
Regions	Angola	Botswana	Lesotho	Malawi	Mozambique	Namibia
p -value	0.9094	0.8807	0.0092	0.5300	0.1289	0.8267
	South Africa	Swaziland	Zimbabwe	Zambia		
	0.0045	0.1491	0.5688	0.2231		

Next, we standardized the two socioeconomic factors. The multiplex collinear test [32] was performed by the condition number $k = \sqrt{\lambda_{\max}/\lambda_{\min}} = 1.804 (\leq 15)$ (λ is the eigenvalue of explanatory variable matrix). If $k > 15$, then the data have collinearity. Otherwise, there is no collinearity. Thus, there is no collinearity between the two factors. Let μ_{ik} , r_{ik} and P_{ik} be the annual HIV/AIDS cases (Unit: 1/1000 people), incidence (Unit: 1/100) and total population (unit: 100,000 people) in the k th year of the i th region, respectively. Denote $g(\mu_{ik}) = \eta_{ik} = \ln \mu_{ik} = \ln r_{ik} + \ln P_{ik} (\mu_{ik} = r_{ik}P_{ik}, i = 1, 2, \dots, 10$ and $k = 1, 2, \dots, 6)$. Let X_{i1k} and X_{i2k} be the GDP per capita and population density in the i th region at the k th year, respectively. The PR model is written by

$$g(\mu_{ik}) = \beta_0 + \beta_1 X_{i1k} + \beta_2 X_{i2k}, \quad i = 1, 2, \dots, 10, k = 1, 2, \dots, 6, \tag{10}$$

where $\beta_j (j = 0, 1, 2)$ are unknown constants. The GWPR model is introduced as

$$g(\mu_{ik}) = \beta_0(u_{ik}, v_{ik}) + \beta_1(u_{ik}, v_{ik})X_{i1k} + \beta_2(u_{ik}, v_{ik})X_{i2k}, \quad i = 1, 2, \dots, 10, \tag{11}$$

where k is a fixed constant taken from $\{1, 2, \dots, 6\}$, and $\beta_j(u_{ik}, v_{ik})$ are unknown spatial parameters for the i th country (u_{ik}, v_{ik}) in the k th year. Let ρ be a scalar autoregressive parameter, and c_{ij} be a constant that represents an adjacency relation. The GWPAR model is

$$g(\mu_{ik}) = \rho \sum_{l=1}^n c_{il} \eta_{lk} + \beta_0(u_{ik}, v_{ik}) + \beta_1(u_{ik}, v_{ik})X_{i1k} + \beta_2(u_{ik}, v_{ik})X_{i2k}, \tag{12}$$

where $n = 10$, k is a fixed constant, and $\beta_j(u_{ik}, v_{ik})$ are defined as above. Let $c_{lm}^{(ik)}$ be a spatiotemporal adjacency relation, and $\beta_{jk}(u_{ik}, v_{ik}, t_k)$ ($k = 1, 2, \dots, 6$) be unknown spatiotemporal parameters in the i th country (u_{ik}, v_{ik}) in the k th year. The GTWPAR model is established as follows:

$$g(\mu_{ik}) = \rho \sum_{m=1}^T \sum_{l=1}^{n_k} c_{lm}^{(ik)} \eta_{lm} + \beta_{0k}(u_{ik}, v_{ik}, t_k) + \beta_{1k}(u_{ik}, v_{ik}, t_k) X_{i1k} + \beta_{2k}(u_{ik}, v_{ik}, t_k) X_{i2k}, \tag{13}$$

where $T = 6$; $n_k = 10$ for every k years; and ρ is defined as above.

Algorithms I, II, III and IV of PR, GWPR, GWPAR and GTWPAR models are provided in Appendix A.4, respectively.

4.2. Statistical Analysis

For the PR model, we get the estimated values of unknown parameters by Algorithm I. Then, the best space bandwidth is chosen by the cross-validation method. Following Huang et al. [28], the range [0.09, 2.49] of the space bandwidth is selected according to the minimum and maximum distance of the geographical positions. In the GWPR model, the best space bandwidth is $h = 0.62, 0.59, 0.62, 0.61, 0.60, 0.60$, and the estimations of coefficient functions are given by Algorithm II. The optimal space bandwidth of the GWPAR model is selected as $h = 1.2895, 1.1316, 1.1316, 1.0526, 1.0526, 1.0526$. Based on Algorithm III, we can get the estimations of coefficient functions and the scalar autoregressive parameter $\hat{\rho} = 0.267, 0.269, 0.263, 0.264, 0.264, 0.264$. For the GTWPAR model, we chose $h_s = 1.1316, 0.9737, 0.8947, 0.9211, 0.7842, 0.8789$ and $\tau = 0.1$, where $\tau (> 0)$ is a balanced parameter. The coefficient estimations and scalar autoregressive parameter $\hat{\rho} = 0.126$ can be obtained by Algorithm IV. The quantile and mean values of coefficient estimations and response variables are shown in Table 2. We note that the GWPR, GWPAR and GTWPAR models can reflect the non-stationarity property of the influencing factors; the PR model cannot. Moreover, the GTWPAR model has a better performance than other models by comparing the true and fitted values.

Table 2. The quantile and mean values of coefficient estimations and response variables.

Model	Coefficient	Min	1st Qu	Median	3rd Qu	Max	Mean
True	η	5.293	5.644	6.455	7.288	8.871	6.559
PR	$\hat{\beta}_1$	0.581	0.581	0.581	0.581	0.581	0.581
	$\hat{\beta}_2$	0.385	0.385	0.385	0.385	0.385	0.385
	$\hat{\eta}$	6.344	6.527	7.185	7.447	8.162	7.078
GWPR	$\hat{\beta}_1$	-0.519	-0.249	0.009	0.641	1.774	0.288
	$\hat{\beta}_2$	-0.672	-0.139	0.068	2.044	3.803	0.747
	$\hat{\eta}$	5.618	6.422	7.022	7.202	8.813	6.932
GWPAR	$\hat{\beta}_1$	-6.286	-1.258	-0.910	0.240	3.036	-0.556
	$\hat{\beta}_2$	-2.089	-0.669	0.019	2.281	9.152	0.912
	$\hat{\eta}$	5.176	5.625	6.435	7.219	8.940	6.539
GTWPAR	$\hat{\beta}_1$	-5.642	-0.956	-0.677	0.186	2.992	-0.443
	$\hat{\beta}_2$	-1.865	-0.464	0.215	2.134	9.104	0.988
	$\hat{\eta}$	5.200	5.589	6.387	7.215	8.803	6.489

The average estimated coefficients are visualized in Figure 3. For the PR model, the GDP per capita and population density had the same effect on the HIV/AIDS epidemic for ten countries in six years. However, there exist significant spatial non-stationarity and auto-correlation for different countries under the GWPR, GWPAR and GTWPAR models. Figure 4 shows the spatial distribution

of the average MSE of their response variables. The lighter the color, the smaller the average error is. Thus, the GWPAR and GTWPAR models have the better fitting results.

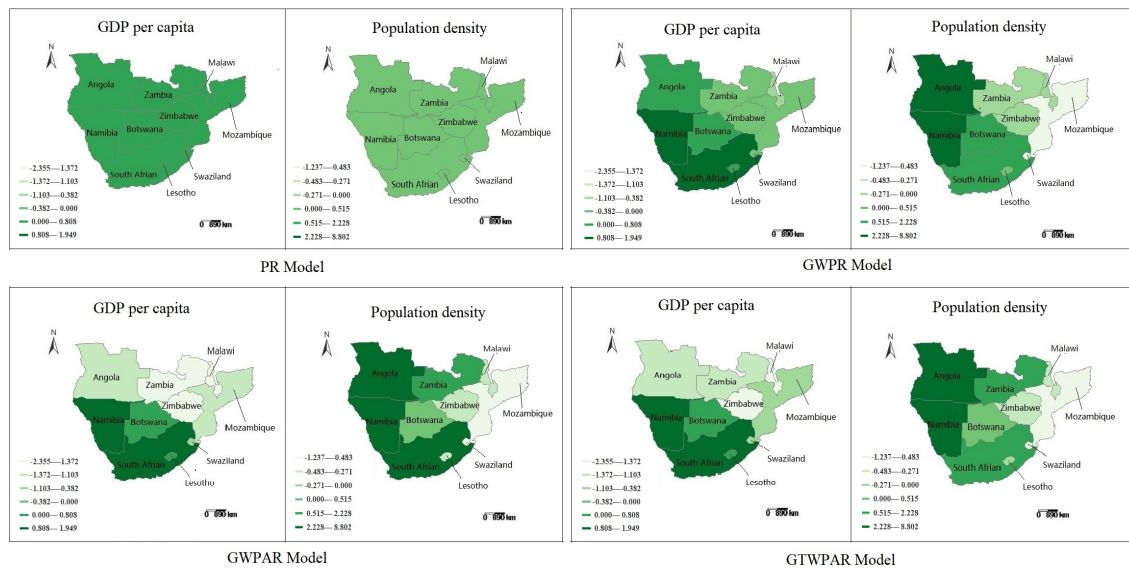


Figure 3. The spatial distribution of the average coefficient estimations in four models.

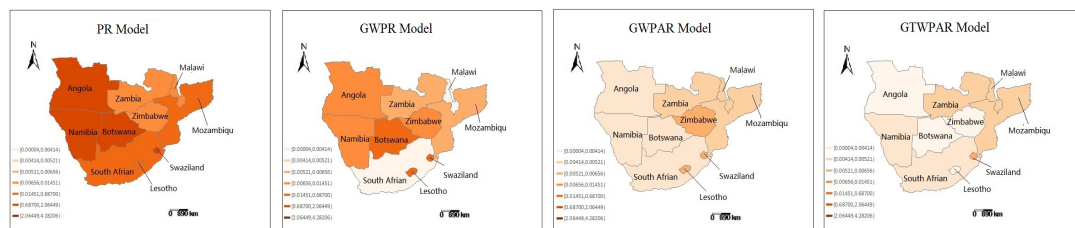


Figure 4. The average MSE of response variables.

These four indicators can effectively compare the performances of the proposed models (Table 3). The calculation formulas of R^2 , AICc, D and MSE are given in Appendix A.5. The coefficient of determination R^2 gradually increases from 12.91% of the PR model to 99.57% of the GTWPAR model. The MSE, AICc and D values of the GTWPAR model are smaller than those of other models. Therefore, the GTWPAR model is more suitable to investigate the spatiotemporal HIV/AIDS epidemic.

Table 3. The comparison of the four models.

Model	R^2	AICc	D	MSE
PR	0.1291	42,504.40	42,624.07	1.6488
GWPR	0.6139	6495.12	6613.02	0.4326
GWPAR	0.9940	155.46	236.02	0.0067
GTWPAR	0.9957	115.25	190.05	0.0048

Based on the GTWPAR model, the mean values and 95% confidence intervals of the coefficient estimations are shown in Figure 5. The mean estimations are represented by the dot, and the 95% confidence intervals are given by the upper and lower lines. Note that the GDP per capita in Botswana, Namibia and South Africa has a positive effect on the HIV/AIDS cases. Six other countries (except Lesotho) had the opposite results. The population density for five countries had a positive effect on the HIV/AIDS cases—Angola, Botswana, Namibia, South Africa and Zambia. The population density of other five countries had the negative effect. Moreover, the impact of the GDP per capita on HIV/AIDS epidemic had a strong spatiotemporal non-stationarity in Lesotho,

Malawi and Zimbabwe, while the population density had a strong spatiotemporal non-stationarity in Angola.

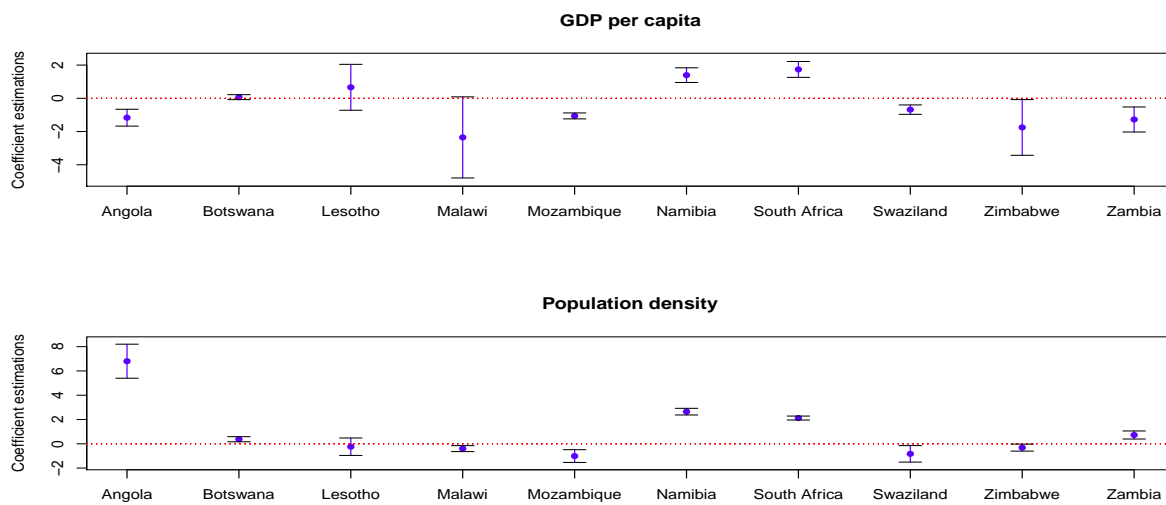


Figure 5. The mean values and 95% confidence intervals of coefficient estimations.

5. Conclusions

In this paper, we propose four regression models, including the PR, GWPR, GWPARG and GTWPAR, to investigate the non-stationary and auto-correlation properties. The relationships between the HIV/AIDS epidemic, GDP per capita and population density were analyzed in ten countries of Sub-Saharan Africa from 2011 to 2016. The unknown parameters of these models can be estimated by the Newton–Raphson procedure and Fisher scoring algorithm.

The PR model is a classical generalized model, which considers the global relationships between the response and explanatory variables. The GWPR and GWPARG models have been introduced to determine the spatial non-stationarity or auto-correlation. The GTWPAR model proposed by this article can be used to investigate not only spatiotemporal non-stationary but also auto-correlation. Thus, the PR, GWPR and GWPARG models are several special cases of the GTWPAR model (see Remark 1 and Remark 2). The performances of these models were evaluated by analyzing the correlations between the HIV/AIDS epidemic and two socioeconomic factors. The parameter estimations of the models can be obtained by Algorithms I, II, III and IV in Appendix A.4.

The results show that the impacts of GDP per capita and population density on HIV/AIDS cases had significant spatiotemporal non-stationarity and auto-correlation. The GWPR, GWPARG and GTWPAR models can reflect the strong spatial or spatiotemporal non-stationarity. The auto-correlation can be reflected in the GWPARG and GTWPAR models. Compared with other models, the GTWPAR model is more effective in terms of four comparison indicators. Thus, we suggest that the GTWPAR model can be used to analyze the spatiotemporal characteristics of the HIV/AIDS epidemic and the influences of the GDP per capita and population density.

Further work also exists in our study. For example, we observed that the effects of the GDP per capita for Lesotho, Malawi and Zimbabwe and the population density for Angola on HIV/AIDS had strong spatiotemporal non-stationarity. These may be the result of local environmental or political factors. Whether the fitting results of these regions will perform better if explanatory variables such as local unique environmental or political factors are added needs to be further investigated.

Author Contributions: Conceptualization, Z.L., H.Z. and X.H.; methodology, S.S., H.Z. and X.H.; software, S.S.; supervision, Z.L. and H.J.; visualization, Z.L.; writing—original draft, S.S.; writing—review & editing, S.S., Z.L. and H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (XJEDU2017M001, 2018Q011), and the National Natural Science Foundation of China (U1703237, 11661076, 12061070).

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

Abbreviations

The following abbreviations are used in this manuscript:

- HIV/AIDS Human immunodeficiency virus/Acquired immunodeficiency syndrome
- PR Poisson regression model
- GWPR Geographically weighted Poisson regression model
- GWPARG Geographically weighted Poisson autoregressive model
- GTWPARG Geographically and temporally weighted Poisson autoregressive model
- GDP Gross domestic product

Appendix A. Detailed Processes

Appendix A.1. The Expressions of C, X and B

In model $\eta = \rho C\eta + B'X'$, the expressions of C, X and B are

$$C = \begin{pmatrix} c_{11}^{(11)} & \cdots & c_{11}^{(n_11)} & \cdots & c_{11}^{(1T)} & \cdots & c_{11}^{(n_T T)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{(n_11)}^{(11)} & \cdots & c_{(n_11)}^{(n_11)} & \cdots & c_{(n_11)}^{(1T)} & \cdots & c_{(n_11)}^{(n_T T)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{1T}^{(11)} & \cdots & c_{1T}^{(n_11)} & \cdots & c_{1T}^{(1T)} & \cdots & c_{1T}^{(n_T T)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{(n_T T)}^{(11)} & \cdots & c_{(n_T T)}^{(n_11)} & \cdots & c_{(n_T T)}^{(1T)} & \cdots & c_{(n_T T)}^{(n_T T)} \end{pmatrix}, X = \begin{pmatrix} X_{111} & X_{121} & \cdots & X_{1p1} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n_111} & X_{n_121} & \cdots & X_{n_1p1} \\ \vdots & \vdots & \vdots & \vdots \\ X_{11T} & X_{12T} & \cdots & X_{1pT} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n_T1T} & X_{n_T2T} & \cdots & X_{n_TpT} \end{pmatrix},$$

where C_{IK}, X_{IK} are respectively the I th row and the K th column of the matrices C and X. Moreover,

$$B = (\beta_{11}(u_{11}, v_{11}, t_1), \cdots, \beta_{p1}(u_{11}, v_{11}, t_1), \cdots, \beta_{11}(u_{n_11}, v_{n_11}, t_1), \cdots, \beta_{p1}(u_{n_11}, v_{n_11}, t_1), \cdots, \beta_{1T}(u_{1T}, v_{1T}, t_T), \cdots, \beta_{pT}(u_{1T}, v_{1T}, t_T), \cdots, \beta_{1T}(u_{n_T1}, v_{n_T1}, t_1), \cdots, \beta_{pT}(u_{n_T1}, v_{n_T1}, t_1))'$$

Appendix A.2. Formula and Information Matrix of $\beta(u, v, t)$

(1) For the matrix form of model (5), we obtain

$$\begin{cases} (1 - \rho C_{11})\eta_1 - \rho C_{12}\eta_2 - \cdots - \rho C_{1N}\eta_N = \sum_{j=1}^p \beta_{j0} X_{1j}, \\ -\rho C_{21}\eta_1 + (1 - \rho C_{22})\eta_2 - \cdots - \rho C_{2N}\eta_N = \sum_{j=1}^p \beta_{j0} X_{2j}, \\ \vdots \\ -\rho C_{N1}\eta_1 - \rho C_{N2}\eta_2 - \cdots + (1 - \rho C_{NN})\eta_N = \sum_{j=1}^p \beta_{j0} X_{Nj}. \end{cases}$$

Differentiating the above equations with β_{r0} yields

$$\begin{cases} (1 - \rho C_{11}) \frac{\partial \eta_1}{\partial \beta_{r0}} - \rho C_{12} \frac{\partial \eta_2}{\partial \beta_{r0}} - \dots - \rho C_{1N} \frac{\partial \eta_N}{\partial \beta_{r0}} = X_{1r}, \\ -\rho C_{21} \frac{\partial \eta_1}{\partial \beta_{r0}} + (1 - \rho C_{22}) \frac{\partial \eta_2}{\partial \beta_{r0}} - \dots - \rho C_{2N} \frac{\partial \eta_N}{\partial \beta_{r0}} = X_{2r}, \\ \vdots \\ -\rho C_{N1} \frac{\partial \eta_1}{\partial \beta_{r0}} - \rho C_{N2} \frac{\partial \eta_2}{\partial \beta_{r0}} - \dots + (1 - \rho C_{NN}) \frac{\partial \eta_N}{\partial \beta_{r0}} = X_{Nr}. \end{cases}$$

Then,

$$\begin{pmatrix} 1 - \rho C_{11} & -\rho C_{12} & \dots & -\rho C_{1N} \\ -\rho C_{21} & 1 - \rho C_{22} & \dots & -\rho C_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ -\rho C_{N1} & -\rho C_{N2} & \dots & 1 - \rho C_{NN} \end{pmatrix} \begin{pmatrix} \frac{\partial \eta_1}{\partial \beta_{r0}} \\ \frac{\partial \eta_2}{\partial \beta_{r0}} \\ \vdots \\ \frac{\partial \eta_N}{\partial \beta_{r0}} \end{pmatrix} = \begin{pmatrix} X_{1r} \\ X_{2r} \\ \vdots \\ X_{Nr} \end{pmatrix}.$$

Denote $\frac{\partial \eta}{\partial \beta_{r0}} = (\frac{\partial \eta_1}{\partial \beta_{r0}}, \frac{\partial \eta_2}{\partial \beta_{r0}}, \dots, \frac{\partial \eta_N}{\partial \beta_{r0}})'$ and

$$A = \begin{pmatrix} 1 - \rho C_{11} & -\rho C_{12} & \dots & -\rho C_{1N} \\ -\rho C_{21} & 1 - \rho C_{22} & \dots & -\rho C_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ -\rho C_{N1} & -\rho C_{N2} & \dots & 1 - \rho C_{NN} \end{pmatrix}^{-1},$$

$$X_{.r} = (X_{1r}, X_{2r}, \dots, X_{Nr})', A_{l.} = (A_{l1}, A_{l2}, \dots, A_{lN})'.$$

Thus, $\frac{\partial \eta}{\partial \beta_{r0}} = AX_{.r}$, that is

$$\frac{\partial \eta_l}{\partial \beta_{r0}} = \sum_{h=1}^N A_{lh} X_{hr} = A_{l.} X_{.r}, \quad l = 1, 2, \dots, N.$$

(2) The element I_{rb} of $I(\beta)$ satisfies

$$\begin{aligned} I_{rb}(\beta) &= -\frac{\partial^2 \mathbf{L}_1}{\partial \beta_{b0} \partial \beta_{r0}} \\ &= -\frac{\partial}{\partial \beta_{b0}} \left(\sum_{l=1}^N \left(\frac{Y_l - \mu_l}{a_l \phi} \right) \left(\frac{A_{l.} X_{.r}}{V(\mu_l) g'(\mu_l)} \right) W_l(u_{00}, v_{00}, t_0) \right) \\ &= -\sum_{l=1}^N \left(\frac{Y_l - \mu_l}{a_l \phi} \right) \frac{\partial}{\partial \beta_{b0}} \left(\frac{A_{l.} X_{.r}}{V(\mu_l) g'(\mu_l)} \right) W_l(u_{00}, v_{00}, t_0) \\ &\quad - \sum_{l=1}^N \left(\frac{A_{l.} X_{.r}}{V(\mu_l) g'(\mu_l)} \right) \frac{\partial}{\partial \beta_{b0}} \left(\frac{Y_l - \mu_l}{a_l \phi} \right) W_l(u_{00}, v_{00}, t_0) \\ &= -\sum_{l=1}^N \left(\frac{Y_l - \mu_l}{a_l \phi} \right) \frac{\partial}{\partial \beta_{b0}} \left(\frac{A_{l.} X_{.r}}{V(\mu_l) g'(\mu_l)} \right) W_l(u_{00}, v_{00}, t_0) \\ &\quad + \sum_{l=1}^N \frac{A_{l.} X_{.r} A_{l.} X_{.b}}{a_l \phi V(\mu_l) (g'(\mu_l))^2} W_l(u_{00}, v_{00}, t_0). \end{aligned}$$

The Fisher information matrix is

$$\begin{aligned} \mathbf{I}(\beta) &= E(I(\beta)) = E((I_{rb}(\beta))_{p \times p}) = \sum_{l=1}^N \frac{A_l \cdot X_r A_l \cdot X_b}{a_l \phi V(\mu_l) (g'(\mu_l))^2} W_l(u_{00}, v_{00}, t_0) \\ &= \frac{1}{\phi} \sum_{l=1}^N T_l A_l \cdot X_r A_l \cdot X_b W_l(u_{00}, v_{00}, t_0). \end{aligned}$$

Appendix A.3. Formula and Information Matrix of ρ

(1) Differentiating (4) with ρ yields

$$\begin{cases} (1 - \rho C_{11}) \frac{d\eta_1}{d\rho} - \rho C_{12} \frac{d\eta_2}{d\rho} - \dots - \rho C_{1N} \frac{d\eta_N}{d\rho} = C_{11}\eta_1 + C_{12}\eta_2 + \dots + C_{1N}\eta_N, \\ -\rho C_{21} \frac{d\eta_1}{d\rho} + (1 - \rho C_{22}) \frac{d\eta_2}{d\rho} - \dots - \rho C_{2N} \frac{d\eta_N}{d\rho} = C_{21}\eta_1 + C_{22}\eta_2 + \dots + C_{2N}\eta_N, \\ \vdots \\ -\rho C_{N1} \frac{d\eta_1}{d\rho} - \rho C_{N2} \frac{d\eta_2}{d\rho} - \dots + (1 - \rho C_{NN}) \frac{d\eta_N}{d\rho} = C_{N1}\eta_1 + C_{N2}\eta_2 + \dots + C_{NN}\eta_N. \end{cases}$$

That is to say

$$\begin{pmatrix} 1 - \rho C_{11} & -\rho C_{12} & \dots & -\rho C_{1N} \\ -\rho C_{21} & 1 - \rho C_{22} & \dots & -\rho C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -\rho C_{N1} & -\rho C_{N2} & \dots & 1 - \rho C_{NN} \end{pmatrix} \begin{pmatrix} \frac{d\eta_1}{d\rho} \\ \frac{d\eta_2}{d\rho} \\ \vdots \\ \frac{d\eta_N}{d\rho} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1N} \\ C_{21} & C_{22} & \dots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N1} & C_{N2} & \dots & C_{NN} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{pmatrix}.$$

Let $\frac{d\eta}{d\rho} = (\frac{d\eta_1}{d\rho}, \dots, \frac{d\eta_N}{d\rho})'$. Then, $\frac{d\eta}{d\rho} = AC\eta$, where

$$\frac{d\eta_l}{d\rho} = \frac{dg(\mu_l)}{d\rho} = \sum_{h=1}^N A_l \cdot C_{.h} \eta_h, \quad l = 1, 2, \dots, N.$$

(2) The scalar vector of ρ is

$$\begin{aligned} S(\rho) &= \frac{dL_2}{d\rho} = \frac{1}{\phi} \sum_{l=1}^N T_l \left(\sum_{h=1}^N A_l \cdot C_{.h} \eta_h \right) (Y_l - \mu_l) g'(\mu_l) \\ &= \frac{1}{\phi} (AC\eta)' T(Z - \eta). \end{aligned}$$

The information matrix is

$$\begin{aligned} I(\rho) &= -\frac{\partial^2 L_2}{\partial \rho^2} = -\frac{d}{d\rho} \sum_{l=1}^N \left(\frac{Y_l - \mu_l}{a_l \phi} \right) \left(\frac{\sum_{h=1}^N A_l \cdot C_{.h} \eta_h}{V(\mu_l) g'(\mu_l)} \right) \\ &= -\sum_{l=1}^N \left(\frac{Y_l - \mu_l}{a_l \phi} \right) \frac{d}{d\rho} \left(\frac{\sum_{h=1}^N A_l \cdot C_{.h} \eta_h}{V(\mu_l) g'(\mu_l)} \right) \\ &\quad - \sum_{l=1}^N \left(\frac{\sum_{h=1}^N A_l \cdot C_{.h} \eta_h}{V(\mu_l) g'(\mu_l)} \right) \frac{d}{d\rho} \left(\frac{Y_l - \mu_l}{a_l \phi} \right), \end{aligned}$$

where

$$\frac{d\mu_l}{d\rho} = \frac{d\mu_l}{dg(\mu_l)} \frac{dg(\mu_l)}{d\rho} = \frac{1}{g'(\mu_l)} \sum_{h=1}^N A_l \cdot C_{.h} \eta_h.$$

Thus,

$$I(\rho) = - \sum_{l=1}^N \left(\frac{Y_l - \mu_l}{a_l \phi} \right) \frac{d}{d\rho} \left(\frac{\sum_{h=1}^N A_l \cdot C_{.h} \eta_h}{V(\mu_l) g'(\mu_l)} \right) + \sum_{l=1}^N \frac{(\sum_{h=1}^N A_l \cdot C_{.h} \eta_h)^2}{a_l \phi V(\mu_l) (g'(\mu_l))^2}.$$

The Fisher information matrix is

$$\begin{aligned} \mathcal{I}(\rho) &= E(I(\rho)) = \sum_{l=1}^N \frac{(\sum_{h=1}^N A_l \cdot C_{.h} \eta_h)^2}{a_l \phi V(\mu_l) (g'(\mu_l))^2} \\ &= \frac{1}{\phi} \sum_{l=1}^N T_l \left(\sum_{h=1}^N A_l \cdot C_{.h} \eta_h \right)^2 = \frac{1}{\phi} (AC\eta)' T(AC\eta). \end{aligned}$$

Appendix A.4. Algorithms of Coefficient Estimation

We provide four algorithms to estimate the unknown coefficients of PR, GWPR, GWPARG and GTWPAR models in Section 4.

Algorithm I: Estimate the unknown parameters in the PR model. Take the initial values $g(\mu_{ik}^{(0)}) = \eta_{ik}^{(0)} = \ln \mu_{ik}^{(0)}, y_{ik} = \mu_{ik}^{(0)}, Z_{ik}^{(0)} = \eta_{ik}^{(0)} + g'(\mu_{ik}^{(0)})(y_{ik} - \mu_{ik}^{(0)})$, and

$$w_{ik}^{(0)} = \frac{1}{a_{ik} V(\mu_{ik}^{(0)}) (g'(\mu_{ik}^{(0)}))^2}, i = 1, 2, \dots, 10, k = 1, 2, \dots, 6.$$

The iterative formula of $\hat{\beta}^{(m+1)}$ is

$$\hat{\beta}^{(m+1)} = \left(X' W^{(m)} X \right)^{-1} X' W^{(m)} Z^{(m)}.$$

Repeat the above step until convergence yields. The estimated value $\hat{\beta} = \hat{\beta}^{(m)}$ can be obtained.

Algorithm II: Estimate the unknown coefficients in the GWPR model. (Note that k is a fixed constant taken from $\{1, 2, \dots, 6\}$ and the following steps should be repeated six times independently). Take the initial values $\eta_{ik}^{(0)} = g(\mu_{ik}^{(0)}), y_{ik} = \mu_{ik}^{(0)}, Z_{ik}^{(0)} = \eta_{ik}^{(0)} + g'(\mu_{ik}^{(0)})(y_{ik} - \mu_{ik}^{(0)})$, and

$$t_{ik}^{(0)} = \frac{1}{a_{ik} V(\mu_{ik}^{(0)}) (g'(\mu_{ik}^{(0)}))^2}, i = 1, 2, \dots, 10.$$

Let $Z^{(0)} = (Z_{1k}^{(0)}, Z_{2k}^{(0)}, \dots, Z_{10k}^{(0)})'$ and $T^{(0)} = \text{diag}(T_{1k}^{(0)}, T_{2k}^{(0)}, \dots, T_{10k}^{(0)})$. The iterative formula of $\hat{\beta}^{(m+1)}$ at the location (u_{0k}, v_{0k}) is

$$\hat{\beta}^{(m+1)}(u_{0k}, v_{0k}) = \left(A'(u_{0k}, v_{0k}) T^{(m)} W(u_{0k}, v_{0k}) A(u_{0k}, v_{0k}) \right)^{-1} A'(u_{0k}, v_{0k}) T^{(m)} W(u_{0k}, v_{0k}) Z^{(m)},$$

where $W(u_{0k}, v_{0k}) = \text{diag}(w_1(u_{0k}, v_{0k}), w_2(u_{0k}, v_{0k}), \dots, w_{10}(u_{0k}, v_{0k}))$ and

$$w_i(u_{0k}, v_{0k}) = \frac{1}{\sqrt{2\pi}} \exp \left(- \frac{1}{2} \left(\frac{d_{ik}^{(0)}}{h} \right)^2 \right).$$

Repeat the above step until convergence. When (u_{0k}, v_{0k}) takes all the locations (u_{ik}, v_{ik}) , we will get the estimated value $\hat{\beta} = \hat{\beta}^{(m)}$ in a fixed the k th year.

Algorithm III: Estimate unknown coefficients in the GWPARG model. (Note that k is a fixed constant and the following steps should be repeated six times as in Algorithm II). Take the initial value

$\beta_1^{(0)}, \beta_2^{(0)}$ from Algorithm II, and $\rho^{(0)}$ is the absolute value of spatial Moran's I = 0.4480. The initial values $\eta^{(0)} = (I - \rho^{(0)}C)^{-1}X\beta^{(0)}, \mu_{ik}^{(0)} = g^{-1}(\eta_{ik}^{(0)})$, and

$$Z_{ik}^{(0)} = g(\mu_{ik}^{(0)}) + g'(\mu_{ik}^{(0)})(y_{ik} - \mu_{ik}^{(0)}), t_{ik}^{(0)} = \frac{1}{a_{ik}V(\mu_{ik}^{(0)})(g'(\mu_{ik}^{(0)}))^2}, i = 1, 2, \dots, 10.$$

The $(m + 1)$ th iterative estimation $\hat{\beta}^{(m+1)}(u_{0k}, v_{0k})$ and $\hat{\rho}^{(m+1)}$ is

$$\begin{aligned} \hat{\beta}^{(m+1)}(u_{0k}, v_{0k}) &= ((A^{(m)}X)'T^{(m)}W(A^{(m)}X))^{-1}(A^{(m)}X)'T^{(m)}WZ^{(m)}, \\ \hat{\rho}^{(m+1)} &= \rho^{(m)} + ((A^{(m)}C\eta^{(m)})'T^{(m)}(A^{(m)}C\eta^{(m)}))^{-1}(A^{(m)}C\eta^{(m)})'T^{(m)}(Z^{(m)} - \eta^{(m)}). \end{aligned}$$

If (u_{0k}, v_{0k}) takes all the locations (u_{ik}, v_{ik}) , the estimate $\hat{\beta}^{(m+1)}$ can be given. When all estimated values arrive to converge, we will get $\hat{\beta} = \hat{\beta}^{(m)}(u_{ik}, v_{ik})$ and $\hat{\rho} = \hat{\rho}^{(m)}$ in a fixed the k th year.

Algorithm IV: Estimate the unknown coefficients in the GTWPAR model. Take the initial values $\beta_{1k}^{(0)}(u_{ik}, v_{ik}, t_k), \beta_{2k}^{(0)}(u_{ik}, v_{ik}, t_k), i = 1, 2, \dots, 10, k = 1, 2, \dots, 6$ from Algorithm III, and $\rho^{(0)}$ is the absolute value of spatiotemporal Moran's I = 0.2143. The initial value vector $\eta^{(0)} = (I - \rho^{(0)}C)^{-1}X\beta^{(0)}, \mu_{ik}^{(0)} = g^{-1}(\eta_{ik}^{(0)}), Z^{(0)} = (Z_1^{(0)}, Z_2^{(0)}, \dots, Z_{60}^{(0)})', T^{(0)} = \text{diag}(T_1^{(0)}, T_2^{(0)}, \dots, T_{60}^{(0)})$ and

$$Z_{ik}^{(0)} = g(\mu_{ik}^{(0)}) + g'(\mu_{ik}^{(0)})(y_{ik} - \mu_{ik}^{(0)}), T_{ik}^{(0)} = \frac{1}{a_{ik}V(\mu_{ik}^{(0)})(g'(\mu_{ik}^{(0)}))^2}.$$

The $(m + 1)$ th iterative estimations $\hat{\beta}(u_{00}, v_{00}, t_0)$ and $\hat{\rho}$ are

$$\begin{aligned} \hat{\beta}^{(m+1)}(u_{00}, v_{00}, t_0) &= ((A^{(m)}X)'T^{(0)}W(A^{(m)}X))^{-1} \\ &\quad \times (A^{(m)}X)'T^{(0)}WZ^{(m)}, \\ \hat{\rho}^{(m+1)} &= \rho^{(m)} + ((A^{(m)}C\eta^{(m)})'T^{(m)}(A^{(m)}C\eta^{(m)}))^{-1} \\ &\quad \times (A^{(m)}C\eta^{(m)})'T^{(m)}(Z^{(m)} - \eta^{(m)}). \end{aligned}$$

where $W = \{w_{ik}(u_{00}, v_{00}, t_0)\}$ and

$$w_{ik}(u_{00}, v_{00}, t_0) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(u_{00} - u_{ik})^2 + (v_{00} - v_{ik})^2}{h_{\xi}^2} + \frac{(t_0 - t_k)^2}{\tau h_{\xi}^2} \right) \right\}.$$

A detailed definition is given in Section 3.1. If (u_{00}, v_{00}, t_0) takes all the locations (u_{ik}, v_{ik}, t_k) and all estimations converge, we will get $\hat{\beta} = \hat{\beta}^{(m)}(u_{ik}, v_{ik}, t_k)$ and $\hat{\rho} = \hat{\rho}^{(m)}$.

It is worth noting that we use the parameter estimates of the previous model as the initial values of the next model to reduce the number of iterations and improve the operational efficiency. For example, the estimations of the GWPR model are selected as the initial values of the GWPAR model.

Appendix A.5. The Model Comparison Indicators (See Table 3)

(1) The coefficient of determination is defined by

$$R^2 = 1 - \frac{\sum(\eta - \hat{\eta})^2}{\sum(\eta - \bar{\eta})^2},$$

where η is a set of vectors $\{\eta_{ik}\}$, $\hat{\eta}$ and $\bar{\eta}$ are the parameter estimate and the mean value of η , respectively.

(2) Deviation can be defined as

$$D = \sum(y \ln(\frac{\hat{\mu}}{y}) + (y - \hat{\mu})),$$

where y is a set of response variables, and $\hat{\mu}$ is the estimation of $\mu=E(y)$.

(3) The corrected Akaike information criterion is

$$AICc = D + 2P + 2 \frac{P(P + 1)}{N - P - 1},$$

where D , P and N are the deviation, the number of parameters and the number of samples, respectively.

(4) Mean-square error is given by

$$MSE = \frac{1}{N - P} \sum (\hat{\eta} - \eta)^2,$$

where the parameter settings are the same as above.

References

- Hlongwa, P. Current ethical issues in HIV/AIDS research and HIV/AIDS care. *Oral Dis.* **2016**, *22*, 61–65. [[CrossRef](#)] [[PubMed](#)]
- Buve, A.; Bishikwabo-Nsarhaza, K.; Mutangadura, G. The spread and effect of HIV-1 infection in sub-Saharan Africa. *Lancet* **2002**, *359*, 2011–2017. [[CrossRef](#)]
- Vandormael, A.; Oliveira, T.D.; Tanser, F.; Till, B.; Herbeck, J.T. High percentage of undiagnosed HIV cases within a hyperendemic South African community: A population-based study. *J. Epidemiol Community Health* **2018**, *72*, 168–172. [[CrossRef](#)] [[PubMed](#)]
- Abbas, U.L.; Robert, G.; Anuj, M.; Gregory, H.; Mellors, J.W. Antiretroviral therapy and pre-exposure prophylaxis: Combined impact on HIV transmission and drug resistance in South Africa. *J. Infect. Dis.* **2013**, *208*, 224–234. [[CrossRef](#)] [[PubMed](#)]
- Wojcicki, J.M. Socioeconomic status as a risk factor for HIV infection in women in East, Central and Southern Africa: A systematic review. *J. Biosoc. Sci.* **2005**, *37*, 1–9. [[CrossRef](#)] [[PubMed](#)]
- Hallman, K. Gendered socioeconomic conditions and HIV risk behaviours among young people in South Africa. *Afr. J. AIDS Res.* **2005**, *4*, 37–50. [[CrossRef](#)]
- Chris, K.; Sizwe, Z.; Robert, C.; Siphon, D. Why have socioeconomic explanations been favoured over cultural ones in explaining the extensive spread of hiv in south africa? *S. Afr. J. HIV Med.* **2012**, *13*, 14–16.
- Stover, J. Projecting the demographic consequences of adult hiv prevalence trends: The spectrum projection package. *Sex. Transm. Infect.* **2004**, *80*, 14–18. [[CrossRef](#)]
- Shiboski, S.C.; Jewell, N.P. Statistical analysis of the time dependence of HIV infectivity based on partner study data. *J. Am. Stat. Assoc.* **1992**, *87*, 360–372. [[CrossRef](#)]
- Marsh, K.; Mahy, M.; Salomon, J.A.; Hogan, D.R. Assessing and adjusting for differences between hiv prevalence estimates derived from national population-based surveys and antenatal care surveillance, with applications for spectrum 2013. *AIDS* **2014**, *28*, S497–S505. [[CrossRef](#)]
- Palk, L.; Blower, S. Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the HIV epidemic in Malawi. *BMC Med.* **2018**, *16*, 22–31. [[CrossRef](#)] [[PubMed](#)]
- Nelder, J.A.; Wedderburn, R.W.M. Generalized linear models. *J. R. Stat. Soc.* **1972**, *135*, 370–384.
- Myers, R.H.; Montgomery, D.C. A tutorial on generalized linear models. *J. Qual. Technol.* **1997**, *29*, 274–275. [[CrossRef](#)]
- Madden, L.V.; Boudreau, M.A. Effect of strawberry density on the spread of anthracnose caused by colletotrichum acutatum. *Phytopathology* **1997**, *87*, 828–838. [[CrossRef](#)]
- Guisan, A.; Edwards, T.C.; Hastie, T. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecol. Model.* **2002**, *157*, 89–100. [[CrossRef](#)]
- Marshall, R.J. A review of methods for the statistical analysis of spatial patterns of disease. *J. R. Stat. Soc.* **1991**, *154*, 421–441. [[CrossRef](#)]
- Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [[CrossRef](#)]
- Nakaya, T.; Fotheringham, A.S.; Brunsdon, C.; Charlton, M.E. Geographically weighted Poisson regression for disease association mapping. *Stat. Med.* **2005**, *24*, 2695–2717. [[CrossRef](#)]

19. Hadayeghi, A.; Shalaby, A.S.; Persaud, B.N. Development of planning level transportation safety tools using geographically weighted Poisson regression. *Accid. Anal. Prev.* **2010**, *42*, 676–688. [[CrossRef](#)]
20. Li, Y.; Jiao, Y.; Browder, J.A. Modeling spatially-varying ecological relationships using geographically weighted generalized linear model: A simulation study based on longline seabird bycatch. *Fish. Res.* **2016**, *181*, 14–24. [[CrossRef](#)]
21. Anselin, L. Spatial econometrics: Methods and models. *J. Am. Stat. Assoc.* **1990**, *85*, 160–161.
22. Lichstein, J.W.; Simons, T.R.; Franzreb, S.K.E. Spatial Autocorrelation and Autoregressive Models in Ecology. *Ecol. Monogr.* **2002**, *72*, 445–463. [[CrossRef](#)]
23. Lee, L.F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **2004**, *72*, 1899–1925. [[CrossRef](#)]
24. Brunson, C.; Fotheringham, A.S.; Charlton, M.E. Some notes on parametric significance tests for geographically weighted regression. *J. Reg. Sci.* **2010**, *39*, 497–524. [[CrossRef](#)]
25. Zhang, Y.H. Generalized Space Variable Coefficient Autoregressive Model and Its Application in the Impact of Macroscopic-Factors on the HIV/AIDS Incidence. Master's Dissertation, Xinjiang University, Urumqi, China, 2017.
26. Yan, N.; Mei, C.L. A two-step local smoothing approach for exploring spatio-temporal patterns with application to the analysis of precipitation in the mainland of China during 1986–2005. *Environ. Ecol. Stat.* **2014**, *21*, 373–390. [[CrossRef](#)]
27. Fotheringham, A.S.; Crespo, R.; Yao, J. Geographical and temporal weighted regression (GTWR). *Geogr. Anal.* **2015**, *47*, 431–452. [[CrossRef](#)]
28. Huang, B.; Wu, B.; Michael, B. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 383–401. [[CrossRef](#)]
29. Pace, R.K.; Barry, R.; Clapp, J.M.; Rodriguez, M. Spatiotemporal autoregressive models of neighborhood effects. *J. R. Estate Financ. Econ.* **1998**, *17*, 15–33.
30. Wu, B.; Li, R.; Huang, B. A geographically and temporally weighted autoregressive model with application to housing prices. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1186–1204. [[CrossRef](#)]
31. Yan, N.; Mei, C.L.; Wang, N. A unified bootstrap test for local patterns of spatio-temporal association. *Environ. Plan. A* **2015**, *47*, 227–242.
32. Wu, X.Z. Multivariate analysis. In *Complex Data Statistical Methods Application Based on R*; Wu, X.Z., Ed.; China Renmin University Press: Beijing, China, 2013; pp. 116–144.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).