

บทความวิชาการ

การประยุกต์ใช้วิธีการวิเคราะห์เชิงพื้นที่สำหรับการศึกษาทางระบาดวิทยา:
การพัฒนางานวิจัยด้านสาธารณสุขด้วย Geographically weighted Poisson
regression (GWPR) โดยใช้ R

ศักดิ์สิน สิมสินธุ์^{1,*} ส.ม.

Received: March 17, 2025

Revised: May 15, 2025

Accepted: May 15, 2025

บทคัดย่อ

การวิจัยด้านสาธารณสุขมีแนวโน้มเกี่ยวข้องกับการวิเคราะห์เชิงพื้นที่เพิ่มขึ้น เพื่อระบุรูปแบบการกระจายของโรค ประเมินปัจจัยเสี่ยงและสนับสนุนการตัดสินใจเชิงนโยบายสำหรับพื้นที่โดยอิงหลักฐานอย่างเหมาะสม อย่างไรก็ตาม แบบจำลองสถิติดั้งเดิม เช่น การถดถอยแบบปัวซอง (Poisson regression) มีการประมาณค่าสัมประสิทธิ์การถดถอยของตัวแบบเป็นค่าเดียวทั้งพื้นที่ศึกษา (Spatial stationarity) จำกัดความสามารถในการอธิบายความผันแปรของความสัมพันธ์ระหว่างผลลัพธ์ด้านสุขภาพกับปัจจัยเสี่ยงที่เปลี่ยนแปลงตามพื้นที่ บทความนี้เสนอการวิเคราะห์เชิงพื้นที่ด้วยเทคนิคการถดถอยแบบปัวซองถ่วงน้ำหนักตามภูมิศาสตร์ (Geographically weighted Poisson regression: GWPR) ซึ่งเป็นวิธีการทางสถิติที่ให้ค่าสัมประสิทธิ์การถดถอยที่ผันแปรในแต่ละหน่วยพื้นที่ศึกษา (Spatial unit) ทำให้สามารถระบุรูปแบบเฉพาะเชิงพื้นที่ของการเกิดโรคและปัจจัยเสี่ยงที่มีอิทธิพลได้ โดยการบูรณาการค่าถ่วงน้ำหนักเชิงพื้นที่ (Spatial weighting) ด้วยฟังก์ชันเคอร์เนลที่เฉพาะ (Specified kernel functions) และประมาณค่าสัมประสิทธิ์การถดถอยเฉพาะพื้นที่ ช่วยให้ GWPR สามารถสร้างแบบจำลองทางระบาดวิทยาที่ระบุความสัมพันธ์เฉพาะพื้นที่ได้อย่างแม่นยำ บทความนี้อธิบายหลักทฤษฎีของ GWPR กรอบวิธีการทางสถิติ และการประยุกต์ใช้ในงานวิจัยด้านสาธารณสุข นอกจากนี้ยังนำเสนอขั้นตอนการวิเคราะห์ด้วยโปรแกรม R ครอบคลุมการเตรียมข้อมูล การหาค่าพารามิเตอร์ที่เหมาะสม การตรวจสอบสมมติฐาน และการแสดงผลค่าสัมประสิทธิ์การถดถอยที่แปรตามพื้นที่ โดยการทำให้ GWPR เข้าใจและใช้ได้ง่ายสำหรับนักวิชาการและบุคลากรด้านสาธารณสุขทำให้สามารถบูรณาการวิธีการวิเคราะห์กับข้อมูลสาธารณสุขจริง ซึ่งท้ายที่สุดจะช่วยสนับสนุนการพัฒนากลยุทธ์ด้านสาธารณสุขที่ขับเคลื่อนด้วยข้อมูลและมุ่งเน้นไปที่พื้นที่อย่างมีประสิทธิภาพบนพื้นฐานของหลักฐานเชิงประจักษ์

คำสำคัญ: การถดถอยปัวซองถ่วงน้ำหนักตามภูมิศาสตร์ การวิเคราะห์ข้อมูลสุขภาพเชิงพื้นที่ การวิจัยด้านสาธารณสุข
วิธีการทางสถิติเชิงพื้นที่

¹ อาจารย์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยพะเยา

* ผู้รับผิดชอบบทความ: saksin.si@up.ac.th

**Applying a spatial analysis method for epidemiological studies:
Advancing public health research with geographically weighted Poisson
regression (GWPR) using R**

Saksin Simsin^{1,*} M.P.H.

ABSTRACT

Public health research increasingly involves spatial analysis to identify disease patterns, assess context-specific risk factors, and inform evidence-based decision-making for specific areas. However, conventional statistical models such as Poisson regression assume spatial stationarity, limiting their ability to compute for geographic heterogeneity in the relationships between health outcomes and explanatory variables. This paper introduces geographically weighted Poisson regression (GWPR) as an advanced spatial analytical technique that addresses these limitations by allowing regression coefficients to vary across each spatial unit of the study region, enabling the identification of localized patterns of disease occurrence and influencing risk factors. By incorporating spatial weighting based on specified kernel functions and localized parameter estimation, GWPR enables the identification of region-specific associations and enhances the accuracy of epidemiological modeling. This review article elucidates the theoretical foundation of GWPR, its methodological framework, and its application in public health research, along with a comprehensive, step-by-step guide for implementing GWPR in R. The guide includes data preparation, bandwidth optimization, model fitting, diagnostics, and visualization of spatially varying coefficients. By making GWPR more accessible to researchers and public health practitioners, this method can be effectively integrated into real-world spatial health data analysis. Ultimately, it supports the development of targeted, data-driven public health strategies based on robust, location-specific evidence.

Keywords: Geographically weighted Poisson regression, Spatial health data analysis, Public health research, Spatial statistical method

¹ Lecturer, Faculty of Public Health, University of Phayao

* Corresponding author: saksin.si@up.ac.th

Introduction

Public health research increasingly relies on spatially distributed health data to understand disease patterns, assess risk factors, develop effective health interventions, and guide evidence-based policymaking. However, traditional global regression models, such as Poisson regression, face significant limitations when applied to spatial health data. One of the major drawbacks of Poisson regression is its assumption of spatial stationarity, which presumes that the relationships between health outcomes and explanatory variables remain constant across geographic space, overlooking spatial heterogeneity, potentially obscuring meaningful local variations in disease risk (Nakaya, Fotheringham, Brunson, & Charlton, 2005; Murakami, Tsutsumida, Yoshida, Nakaya, & Harris, 2023). This assumption is unrealistic in public health studies, where factors influencing disease incidence and mortality, such as environmental conditions, healthcare access, and socio-economic disparities, often vary significantly across regions (Nakaya et al., 2005; Murakami et al., 2023). Additionally, Poisson regression does not account for spatial autocorrelation, the tendency of geographically proximate observations to exhibit similar health outcomes. It neglects spatial weighting and neighborhood effects, where nearby locations tend to exhibit similar health outcomes. These are crucial for precisely determining the spatial diffusion of health risks and ensuring robust epidemiological analysis. This results in biased estimations and decreased model reliability for spatial autocorrelation, where neighboring locations typically exhibit similar health outcomes (Li, Wang, Liu, Bigham & Ragland, 2013; Munira & Sener, 2020). Another key issue is overdispersion, where the variance of the count data exceeds the mean, making

parameter estimates unreliable (Saefuddin, Saepudin & Kusumaningrum, 2013; Brunson, Fotheringham & Charlton, 1996; Wheeler, 2021).

To address the limitations of traditional global models, Geographically weighted poisson regression (GWPR) introduces “a powerful advancement by enabling regression coefficients to vary spatially across geographic units. This means that, rather than estimating a single, fixed relationship between predictors and outcomes across an entire study area, GWPR generates location-specific coefficients that reflect the unique conditions of each locality.” Such an approach is particularly valuable in public health, where the distribution of health determinants often differs significantly between regions (Kulldorff, 1997; Kirby, Delmelle & Eberth, 2017). Unlike global Poisson regression, which applies “the same model coefficients uniformly across all locations,” GWPR adopts a spatially adaptive framework by incorporating spatial weighting functions to produce estimates that respond to local variations in the data, enhances the detection of localized health risks, and enables a more detailed understanding of how socio-environmental determinants influence health outcomes across space (Fotheringham, Brunson & Charlton, 2002; Nakaya et al., 2005; Tobler, 1970). This local modeling approach improves accuracy and helps uncover geographic disparities that might be masked by global models, empowering policymakers and public health professionals to devise geographically targeted interventions that address the specific needs of diverse populations (Murakami et al., 2023; Saefuddin et al., 2013; Comber, Brunson, Charlton, Dong, Harris, Lu et al., 2020).

Furthermore, GWPR is particularly useful in studying healthcare accessibility and disparities,

helping researchers understand how factors such as distance to medical facilities, urban-rural differences, and economic status impact health outcomes across different regions. By identifying spatially varying relationships, GWPR supports the development of region-specific public health strategies aimed at reducing health inequities and improving overall population well-being (Comber et al., 2020; Li et al., 2013). This makes GWPR a valuable tool for epidemiologists and public health planners seeking to develop evidence-based, geographically targeted interventions that are tailored to the specific needs of different communities.

A salient illustration of the profound impact that statistical methods have on data interpretation and research conclusions can be seen in the study by Fundisi, Dlamini, Mokhele, Weir-Smith & Motolwana (2023), which explored determinants of HIV/AIDS self-testing uptake in South Africa using both generalized linear Poisson regression (GLPR) and geographically weighted Poisson regression (GWPR). The GLPR model, as a global approach, identified broad associations between HIV self-testing uptake and predictors such as gender, education and media exposure. However, it failed to capture local variations in these relationships across the

country's diverse socio-geographic landscape. In contrast, the GWPR model revealed substantial spatial heterogeneity, showing that the influence of predictors varied significantly across provinces. For instance, education and access to information had a stronger positive effect on self-testing uptake in urbanized provinces than in rural areas. This localized insight, unattainable through the global model, emphasized the importance of context-specific interventions. The study demonstrated that relying solely on a global model may obscure critical regional disparities, whereas GWPR offers a more nuanced understanding essential for designing targeted public health strategies. Crucially, the comparative evaluation of model performance, as presented in Table 1 (Fundisi et al., 2023), further substantiates the analytical superiority of the GWPR approach. Using the Akaike information criterion (AIC) as an indicator of model adequacy, the GWPR model exhibited a significantly lower AIC value compared to the GLPR model, suggesting a superior fit to the data. Moreover, the pseudo R^2 value, reflecting the proportion of variance explained, was markedly higher for the GWPR model (pseudo $R^2 = 0.91$) than for the GLPR model (pseudo $R^2 = 0.88$).

Table 1. Comparison of model performance between generalized linear Poisson regression (GLPR) and geographically weighted Poisson regression (GWPR).

Model Type	Pseudo R^2	AICc
GLRP	0.88	2552
GWPR	0.91	390

Note: Adapted from Fundisi et al. (2023).

This substantial enhancement in explanatory power highlights the statistical and practical advantages of GWPR in capturing complex spatial dynamics. These findings demonstrate that conventional global models may obscure critical localized disparities, whereas GWPR enables the identification of spatially heterogeneous determinants, thereby offering a robust foundation for the formulation of geographically tailored and equitable public health interventions (Fundisi et al., 2023).

A deeper understanding of the theoretical and practical differences is essential for selecting appropriate analytical methods. To further underscore the conceptual and methodological distinctions between traditional Poisson regression and geographically weighted Poisson regression (GWPR), Table 2 presents a comparative overview of their core characteristics in spatial epidemiological analysis (Bui, Mor, Chemtob, Ha & Levine, 2018;

Fundisi et al., 2023; Comber et al., 2020; Li et al., 2013). The comparison places particular emphasis on the assumptions underlying each model, their capacity to account for spatial dependencies, and their interpretative utility. Notably, GWPR addresses key limitations of Poisson regression by allowing for spatial non-stationarity, incorporating spatial autocorrelation, and providing localized interpretations that better reflect the realities of geographically diverse health data. However, despite its advantages, GWPR also introduces certain methodological complexities and computational demands. The table below offers a concise synthesis of these distinctions, helping to clarify when and why a geographically weighted approach may be preferable (Nakaya et al., 2005; Poliard, Kirakoya-Samadoulougou, Ouédraogo, Collart Dubourg & Samadoulougou, 2021; Tyas et al., 2023).

Table 2. The Comparative overview of Poisson regression and geographically weighted Poisson regression (GWPR) in spatial epidemiological analysis.

Aspect	Poisson regression	Geographically weighted Poisson regression (GWPR)
Spatial assumption	Assumes spatial stationarity—relationships between variables are uniform across all locations.	Assumes spatial non-stationarity, allowing relationships to vary geographically based on local data context.
Spatial autocorrelation	Does not consider that nearby areas often exhibit similar values, leading to biased or inefficient estimates.	Explicitly accounts for spatial autocorrelation by incorporating geographic location and proximity, improving estimation accuracy.
Interpretability	Provides a single, global interpretation that may obscure regional variation.	Provides location-specific interpretations, reflecting how relationships change across space.
Accuracy for spatial data	Often oversimplifies complex spatial variation, reducing accuracy in heterogeneous regions.	Enhances model precision by capturing localized patterns and spatial heterogeneity.
Use case	Best suited for homogeneous or small-scale spatial contexts with minimal regional variation.	Ideal for spatially diverse or large regions where relationships vary by location.
Output	Generates one set of regression coefficients for the entire dataset.	Generates different sets of coefficients for each geographic location, providing detailed local insight.
Policy application	Offers generalized insights, but may lack geographic specificity for targeted interventions.	Facilitates location-specific planning and policy decisions, improving resource allocation.
Advantages	• Simple to implement • Computationally efficient • Easy to interpret results • Ignores spatial variation and dependencies • May yield biased estimates in spatially complex settings	• Captures spatial heterogeneity • Supports localized health analysis • Improves policy relevance • Requires spatial data expertise • Computationally intensive • Sensitive to bandwidth and kernel choice
Limitations	• May lead to biased inferences • Lacks geographic specificity	• Risk of model overfitting if not properly calibrated

Note: Adapted and summarized from Nakaya et al. (2005), Bui et al. (2018), Comber et al. (2020), Fundisi et al. (2023), Poliard et al. (2021) and Tyas et al. (2023).

The subsequent section delineates the conceptual foundations and methodological application of geographically weighted Poisson regression (GWPR), situating it within the broader paradigm of spatial epidemiological modeling. It further illustrates the analytical workflow through an overview of implementation steps using R scripting, highlighting its applicability in public health research and spatial epidemiological analysis (Nakaya et al., 2005; Li et al., 2013; Comber et al., 2020).

The methodological framework of GWPR

The following section outlines the methodological framework of GWPR, detailing its formulation, spatial weighting techniques, and parameter estimation processes essential for analyzing spatial health data effectively. Considering a study region N comprising n separated sub-areas, the GWPR describes the number of case incidences and deaths as count data collected from the i -th area following a Poisson distribution.

The form of a geographically weighted Poisson regression (GWPR) model is as follows (Helmy, Kamaluddin, Iskandar & Suheryanto, 2022; McMillen, 2004).

$$\ln(E(y_i)) = \beta_0(u_{xi}, v_{yi}) + \beta_1(u_{xi}, v_{yi})x_{i1} + \beta_2(u_{xi}, v_{yi})x_{i2} + \dots + \beta_n(u_{xi}, v_{yi})x_{in} + \varepsilon_i$$

where y_i is the observed count data at area i . $\ln(E(y_i))$ is the natural log of the expected number of cancer incidents and deaths per area. $(\beta_0(u_{xi}, v_{yi}), \dots, \beta_n(u_{xi}, v_{yi})x_{in})$ are $n + 1$ continuous functions of the area (u_{xi}, v_{yi}) , denoting the two-dimensional coordinates of the i -th point in the geographical study region, describing the area of i with x and y coordinates, x_{in} denotes the independent variables, where coefficients $\beta_n(u_{xi}, v_{yi})$ are varying conditionals on the separated areas;

(u_{xi}, v_{yi}) is the i -th separated area; and ε_i is the random error term. Thus, in the GWPR method, the parameter can be expressed in the following matrix form (Li, Wang, Liu, Bigham & Ragland, 2013):

$$\beta_n = \begin{bmatrix} \beta_0(u_{x1}, v_{y1}) & \beta_1(u_{x1}, v_{y1}) & \dots & \beta_n(u_{x1}, v_{y1}) \\ \beta_0(u_{x2}, v_{y2}) & \beta_1(u_{x2}, v_{y2}) & \dots & \beta_n(u_{x2}, v_{y2}) \\ \dots & \dots & \dots & \dots \\ \beta_0(u_{xn}, v_{yn}) & \beta_1(u_{xn}, v_{yn}) & \dots & \beta_n(u_{xn}, v_{yn}) \end{bmatrix}$$

The estimated coefficient, β_n at the i -th separated area, can be calculated by calibrating a Poisson regression model with a likelihood that is spatially weighted, with the weights being a kernel function centered on u_i , u_i is a vector of location coordinates (u_{xi}, v_{yi}) (Saefuddin et al., 2013; Wheeler, 2021) as the following equation (Saefuddin et al., 2013; Fotheringham, Brunsdon & Charlton, 2009; Munira & Sener, 2020).

$$\hat{\beta}(i) = (X^T W(u_{xi}, v_{yi}) X)^{-1} X^T W(u_{xi}, v_{yi}) Y$$

$W(u_{xi}, v_{yi})$ represents a spatial weight matrix n by n , which can be expressed by $W(i)$:

$$W_i = \begin{bmatrix} W_{i1} & 0 & \dots & 0 \\ 0 & W_{i2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & W_{in} \end{bmatrix}$$

Where W_{ij} ($j = 1, 2, 3, \dots, n$) is the weight value of observation given to the area j in the model adjustment for the area i , and off-diagonal elements are 0. In this modeling conception, a parameter in the model is estimated for each area based on the observational data at nearby areas, weighting all neighboring observations by a function of distance from that subject, which means the observations closer to area i have more of an impact on the parameter estimation of i than the areas far from it. To estimate the smoothed geographical variations in the parameters with a distance-based weighting scheme, GWPR applies the spatial kernel method (Fotheringham et al. 2009; Munira & Sener, 2020; Li et al., 2013). There are two spatial kernel functions commonly used, (Hadayeghi, Shalaby & Persaud,

2010; Munira & Sener, 2020; Li et al., 2013) the Gaussian and bi-square functions, to produce the weighting scheme as follows:

Gaussian:
$$w_{ij} = \exp -\frac{1}{2} \times \left(\frac{\|u_i - u_j\|}{G_i} \right)^2$$

Bi-square:
$$w_{ij} = \begin{cases} \left[1 - \left(\frac{\|u_i - u_j\|}{G_i} \right)^2 \right]^2 & \text{if } \|u_i - u_j\| < G_i \\ 0 & \text{otherwise} \end{cases}$$

Where $\|u_i - u_j\|$ is the distance between two observations i and j , and G_i is the maximum distance to assign nonzero weights. The parameter G_i represents a quantity called the bandwidth controlled the weighting function's distance decay around each subject point. Observations located beyond bandwidth G_i are assigned a weight of zero (Hadayeghi et al., 2010; Munira & Sener, 2020; Li et al., 2013) as shown in the bi-square function (adaptive kernel). The fixed kernel of the Gaussian function maintains a constant bandwidth and sets the weighting function's magnitude constant to be the same for every area. The results of GWPR models are sensitive to the bandwidth of a given weighting function (Hadayeghi et al., 2010). The Corrected Akaike Information Criterion (AICc) is generally utilized to determine the bandwidth, in which the model with the lowest AICc is identified as the best model performance (Li et al., 2013). A recent application of the Geographically Weighted Generalized Poisson Regression (GWGPR) model underscores its analytical utility in spatial epidemiology. By systematically comparing kernel functions and employing cross-validation for bandwidth optimization, the study identified the fixed bisquare kernel as providing the optimal model fit, as evidenced by the lowest corrected Akaike Information Criterion (AICc) (Tyas, S. W., Gunardi & Puspitasari, L. A., 2023).

Overview of Steps for Estimating GWPR Coefficients (Conceição, 2020):

Step 1: Variable selection

Choosing relevant possible explanatory variables for modeling is the initial step, based on the literature study, expert advice and general knowledge.

Step 2: Multicollinearity check

The Pearson correlation coefficient (r), Spearman's correlation rank, or Pseudo R^2 , along with the variance inflation factor (VIF) and tolerance test, were used to evaluate the multicollinearity among explanatory variables.

Step 3: Estimating the coefficients of the GWPR model

1. Make a likelihood function of Poisson distribution from the n -numbers of response variable

$$L(\beta) = \prod_{i=1}^n \frac{\exp(-\mu(x_i, \beta)) (\mu(x_i, \beta))^{y_i}}{y_i!} \dots \dots \dots (1)$$

Where $\mu(x_i, \beta) = \exp(\mu(x'_i, \beta))$

2. Maximize the log-likelihood function in equation (1) were calculated:

$$\ln L(\beta) = -\sum_{i=1}^n \exp(x'_i, \beta) + \sum_{i=1}^n y_i (x'_i, \beta) - \sum_{i=1}^n \ln y_i! \dots \dots \dots (2)$$

An observation in GWPR is weighted in accordance with its proximity to area i -th, equation (2) can be expressed as (Fotheringham et al. 2009):

$$\begin{aligned} \ln^* L(\beta(u_i, v_i)) &= \left[-\sum_{i=1}^n \exp(x'_i, \beta(u_i, v_i)) \right. \\ &+ \sum_{i=1}^n y_i (x'_i, \beta(u_i, v_i)) \\ &\left. - \sum_{i=1}^n \ln y_i! \right] w_{ij}(u_i, v_i) \end{aligned}$$

3. Make a partial derivative of equation above with respect to the parameters in $\beta'(u_i, v_i)$ and its result must be equal to zero:

However, this equation is iterative.

$$\begin{aligned} \frac{\beta'(u_i, v_i)}{\ln^* L(\beta(u_i, v_i))} &= \left[-\sum_{i=1}^n x_i \exp(x'_i, \beta(u_i, v_i)) + \sum_{i=1}^n x_i y_i \right] w_{ij}(u_i, v_i) = 0 \dots \dots (3) \end{aligned}$$

4. Solving the maximization problem in equation (3) by a modified local Fisher scoring procedure, a form of iteratively reweighted least squares (IRLS), is necessary except for the special case of kernel mapping. In this local scoring procedure, the following matrix computation of weighted least squares should be repeated to update parameter estimates until they converge (Nakaya et al., 2005; Saefuddin et al., 2013):

$$\begin{aligned} \beta^{l+1}(u_i, v_i) \\ = \frac{(X'W(u_i, v_i)A(u_i, v_i)^{(l)}X)^{-1}X'W(u_i, v_i)A(u_i, v_i)^{(l)}Z(u_i, v_i)^{(l)}}{\beta^{l+1}} \\ l+1 \end{aligned}$$

Where β^{l+1} is a vector of local parameter estimates specific to location i and $l+1$ indicates the number of iterations. The vector of local parameter estimates specific to area i when l^{th} iteration is defined as:

$$\beta^l(u_i, v_i) = (\beta_0^{(l)}(u_i, v_i), \beta_1^{(l)}(u_i, v_i), \dots, \beta_n^{(l)}(u_i, v_i))'$$

X is a design matrix and X' denotes the transpose of X

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ 1 & x_{31} & x_{32} & \dots & x_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}$$

$W(u_i, v_i)$ denotes the diagonal spatial weights matrix for i -th area

$$W(u_i, v_i) = \text{diag}[w_{i1}, w_{i2}, \dots, w_{in}]$$

$A(u_i, v_i)^{(l)}$ denotes the variance weights matrix associated with the Fisher scoring for each i -th area.

$$\begin{aligned} A(u_i, v_i)^{(l)} \\ = \text{diag}[\hat{y}_1(\beta^{(l)}(u_i, v_i)), \hat{y}_2(\beta^{(l)}(u_i, v_i)), \dots, \hat{y}_n(\beta^{(l)}(u_i, v_i))] \\ Z(u_i, v_i)^{(l)} \text{ is a vector of adjusted dependent variables defined as:} \\ Z(u_i, v_i)^{(l)} = [z_1^{(l)}(u_i, v_i), z_2^{(l)}(u_i, v_i), \dots, z_n^{(l)}(u_i, v_i)]' \end{aligned}$$

By repeating the iterative procedure for every regression point i , sets of local parameter estimates are obtained.

5. Standard error of parameters in GWPR estimating

The standard Error of Parameters in GWPR estimates is given by

$$SE(\beta(u_i, v_i)) = \sqrt{\text{cov}(\hat{\beta}(u_i, v_i))k}$$

Where

$\text{cov}(\hat{\beta}(u_i, v_i))$ is the variance-covariance matrix of regression parameters estimates and $\text{cov}(\hat{\beta}(u_i, v_i))$ k is the k -th diagonal element of $\text{cov}(\hat{\beta}(u_i, v_i))$

$$\begin{aligned} \text{cov}(\hat{\beta}(u_i, v_i)) \\ = \text{Cov}(u_i, v_i)A(u_i, v_i)^{-1}[\text{Cov}(u_i, v_i)]' \\ \text{Cov}(u_i, v_i) \\ = (X'W(u_i, v_i)A(u_i, v_i)X)^{-1}X'W(u_i, v_i)A(u_i, v_i)X \end{aligned}$$

6. Parameters significance in GWPR testing

The following hypothesis was investigated to determine the significant local k -th parameter estimated:

$$\begin{aligned} H_o : \beta_k(u_i, v_i) &= 0 \\ H_o : \exists k \text{ where } \beta_k(u_i, v_i) &\neq 0; k = 1, 2, 3, \dots, n-1 \end{aligned}$$

The H_a following formula is used to determine the local pseudo t-statistic for the local estimation of the k -th parameter (Nakaya et al., 2005; Saefuddin et al., 2013):

$$t_k(u_i, v_i) = \frac{\beta_k(u_i, v_i)}{SE(\beta_k(u_i, v_i))}$$

The standard threshold of P-values for a significance test is effectively $|t_k(u_i, v_i)| > 1.96$.

Step 4: Model optimisation with stepwise-AICc procedure

An evaluation of a statistical model on a dataset is conducted using the Akaike Information Criterion (AIC) and the Corrected Akaike Information Criterion (AICc). The AICc values of two models calibrated on the same dataset were compared to optimize the final model's quality; the best fitting model with the smallest AICc value of GWR was kept (Fotheringham et al., 2009).

A practical guide for analyzing GWPR in R

This section outlines a systematic approach to implementing geographically weighted Poisson regression (GWPR) in R, focusing on data preparation, model estimation and interpretation of

spatial variations in health outcomes using the capabilities of the `spdep`, `GWmodel`, `sf`, `rgdal`, and raster packages. The first step involves preparing the dataset by ensuring that geographic coordinates are correctly formatted for spatial modeling. To establish a baseline for comparison, a global Poisson regression model is initially applied, highlighting the limitations of traditional models, which assume that relationships between variables remain constant across all locations (Nakaya et al., 2005; Murakami et al., 2023; Fotheringham et al., 2002). To address spatial heterogeneity, the optimal bandwidth is selected using AICc minimization, a crucial step that balances capturing local variations while preventing excessive smoothing (Brunsdon et al., 1996; Hadayeghi et al., 2010). Once the GWPR model is estimated, it generates spatially varying coefficients, providing deeper insights into how relationships between independent variables and health outcomes differ across regions (Li et al., 2013; Munira & Sener, 2020). The final stage involves visualizing the results using spatial maps, which illustrate geographic disparities in risk factors. These visualizations enable researchers and policymakers to identify location-specific public health challenges, facilitating more targeted and effective interventions (Comber et al., 2020; Nakaya et al., 2005).

The accompanying R script provides a structured, step-by-step guide to conducting GWPR analysis, offering a practical framework for modeling spatially heterogeneous health data in epidemiological research (Saefuddin et al., 2013; Kirby et al., 2017).

Practical Steps in the R Script for GWPR Analysis

1. Load necessary libraries-Install and load required R packages.

2. Import and preprocess spatial data-Ensure the dataset contains geographical coordinates and convert it into a spatial object.

3. Fit a global Poisson regression model-establish a baseline model for comparison.

4. Determine the optimal bandwidth-Use AICc to select an adaptive spatial kernel bandwidth.

5. Fit the GWPR model-perform spatially varying Poisson regression.

6. Extract and interpret GWPR results-Investigate spatial variation in coefficients.

7. Visualize spatially varying coefficients-map localized relationships.

8. Interpretation and Policy Implications

Step 1: Install and load required libraries (updated for 2025 best practices)

The following packages are essential:

GWmodel: Provides tools for geographically weighted regression (including Poisson regression).

sp: Legacy package for handling spatial objects, still used by older functions like `GWmodel`.

sf: A modern and preferred package for working with spatial data (supports Simple Features format).

terra: A newer and actively developed package for handling raster and vector spatial data (replacement for `rgdal` and `raster`).

ggplot2: For creating high-quality plots, which will be useful for visualizing regression results and spatial relationships.

Step 1.1: Install required packages

Install spatial and modeling libraries (run once per system)

```
install.packages("GWmodel")
```

For Geographically Weighted Regression (Poisson models included)

```
install.packages ("sp")
# Legacy spatial data classes (required by GWmodel)
install.packages ("sf")
# Modern spatial data handling (Simple Features)
install.packages ("terra")
# Raster/vector spatial data (modern alternative
to raster)
install.packages ("ggplot2")
# General plotting and scatter maps
install.packages ("spdep")
# For spatial weights and neighborhood structures
install.packages ("tmap")
# Thematic mapping for spatial data
install.packages ("car")
# Multicollinearity check
```

Step 1.2: Load required libraries

```
# Load libraries into your R session
library (GWmodel)
library (sp)
library (sf)
library (terra)
library (ggplot2)
library (spdep)
library (tmap)
library (car)
```

#Step 2: Load and prepare spatial data (modern sf approach)

A critical prerequisite for conducting geographically weighted Poisson regression (GWPR) is the preparation of an analytically suitable and spatially referenced dataset. This section outlines the essential data components and formats required for a robust GWPR analysis.

2.1 Understanding spatial data requirements

The dependent variable must adhere to the presumptions of the Poisson distribution, namely, non-negative integer values, and should represent

a count-based health outcome, such as the number of disease cases, mortality occurrences, or hospital admissions. To allow for localized modeling, every observation in the dataset needs to be spatially referenced, either using point coordinates or spatial regions denoted by a common geographic code. Point-based observations are regarded as point-referenced datasets that include latitude and longitude coordinates for each observation (e.g., households, individuals, health facilities). These can be converted into spatial objects using the `sf` package in R. Areal or polygon-based observations in studies aggregated at administrative levels (e.g., districts, provinces, sub-districts) each unit of observation corresponds to a geographic region. The dataset ought to have an area identifier (`district_id`, for example), which can be connected to a spatial polygon file like a GeoJSON or Shapefile.

Independent variables may include socioeconomic indicators, environmental exposures, demographic characteristics, or healthcare access measures and should be selected based on their theoretical relevance and spatial variability. The data may be organized as point-level observations or aggregated by administrative units.

In terms of format, a variety of spatial data formats are compatible with GWPR analysis in R. Spatial information can be provided as shapefiles or GeoJSON files for areal data, which can be imported using the `sf` package in R, or as a CSV file with coordinates, which can be transformed into spatial objects using `st_as_sf()`. When integrating non-spatial health data with spatial boundaries, a shared identifier must be used for accurate merging. Additionally, spatial datasets must be projected into an appropriate coordinate reference system (CRS), such as UTM, to ensure meaningful distance calculations required by the GWPR model.

2.2 R code example: Load and prepare spatial data

```
# Load the required library for spatial data
manipulation
library(sf)

# --- Case 1: Point-level data (CSV file with
latitude and longitude) ---

# Read the CSV file containing health data and
coordinates health_data <- read.csv("your_dataset.
csv")

# Replace with the actual path to CSV file

# Convert the data frame into a spatial object
using Simple Features (sf)

# Convert the tabular data to a spatial object
(sf format)

# Specify the coordinate columns and assign
the WGS84 geographic CRS (EPSG:4326) health_
sf <- st_as_sf(health_data, coords = c("longitude",
"latitude"), crs = 4326)

# Examine the structure of the spatial object

# Print the structure of the spatial object to
verify successful conversion print(health_sf)

# Plot the spatial distribution of health-related
observations

# Visual check: spatial points plot (st_geome-
try(health_sf), main = "Spatial Distribution of Health
Data")

# --- Case 2: Areal-level data (e.g., administrative
boundaries in a shapefile) ----

# --- If data is stored as a polygon shapefile ----

# Import a polygon shapefile representing
geographic units (e.g., districts or provinces)

# Read polygon shapefile representing admin-
istrative boundaries

# Read the shapefile using sf (Simple Features)
admin_sf <- st_read("path/to/your_shapefile.shp")

# Replace with the actual path to your
shapefile
```

```
# Inspect the structure of the spatial polygon
object print(admin_sf)

# --- Optional: Merging health data with spatial
polygons using a common geographic identifier ---

# ---- Merging health data (in CSV) with spatial
polygons (e.g., by district or region) ----

# Load a non-spatial health dataset to be
merged with the shapefile

# Read the non-spatial health data health_data
<- read.csv("your_health_data.csv")

# Replace with the actual path to your health
dataset

# Merge the spatial polygons with health
attributes using a shared column (e.g., 'district_id')
merged_data <- merge(admin_sf, health_data, by
= "district_id")

# The 'merged_data' object is now a spatial
data frame containing both geometry and attribute
information
```

Step 3: Fit a standard Poisson regression (baseline model)

The standard Poisson regression model serves as a foundational step in the analytical workflow, functioning as a global baseline against which the performance of the Geographically Weighted Poisson Regression (GWPR) model can be compared. This model assumes spatial stationarity, where the relationship between predictors and the outcome is constant across the study area.

Multicollinearity check

Before fitting the Poisson regression model, multicollinearity among explanatory variables was examined to ensure the stability and reliability of parameter estimates. The Pearson correlation coefficient (r), Spearman's rank correlation, and Pseudo R^2 were calculated to identify potential pairwise correlations among predictors. Furthermore,

the Variance Inflation Factor (VIF) and tolerance values were assessed, with a VIF greater than 10 or a tolerance below 0.1 indicating possible multicollinearity requiring model refinement.

```
# --- Multicollinearity diagnostics --- library (car)
# Compute Pearson correlation matrix among
predictors cor_matrix <- cor(merged_data[,
c ("predictor 1", "predictor 2", "predictor 3")],
method = "pearson") print (cor_matrix)
```

```
# Compute Variance Inflation Factors (VIF)
vif_model <- lm(case ~ predictor 1 + predictor 2 +
predictor 3, data = merged_data) vif (vif_model)
```

R code: Estimating and evaluating the standard Poisson model

```
# Estimate a global Poisson regression model
using the generalized linear model (GLM) framework
# Replace 'count' with the dependent variable and
specify relevant covariates
```

```
# Fit a standard Poisson regression model to
all data (ignoring location) poisson_model <- glm
(case ~ predictor 1 + predictor 2 + predictor 3,
family = poisson(link = "log"), data = merged_data)
```

```
# Review model coefficients and significance
```

```
# Generate a model summary to examine
global parameter estimates and diagnostics
summary(poisson_model)
```

```
# Optional diagnostic: assess overdispersion
in the Poisson model
```

Assessing the global model facilitates initial insights into the nature of the data, including potential overdispersion and the existence of spatial dependence in the residuals. If evidence suggests a lack of spatial homogeneity or poor model fit (e.g., elevated AIC values, spatial autocorrelation in residuals), the application of a spatially adaptive method such as GWPR becomes methodologically justified.

```
# Optional: Check for overdispersion (goodness-
of-fit)
```

```
# Compute the dispersion ratio (Pearson
chi-square statistic divided by degrees of freedom)
dispersion_ratio <- sum (residuals (poisson_model,
type = "pearson") ^2) / df.residual (poisson_model)
cat ("Dispersion Ratio:", round (dispersion_ratio,
2), "\n")
```

Interpretation of the findings: the regression coefficients indicate the overall associations between predictors and the health outcome across the entire study area. For instance, a positive and significant coefficient for predictor1 suggests that higher predictor1 levels are globally associated with increased disease counts. A better-fitting model is indicated by lower values of the Akaike information criterion (AIC); higher AIC values might be the result of model misspecification or unexplained spatial variation. Spatial clustering observed in deviance or Pearson residuals signals potential spatial non-stationarity, suggesting that the assumption of uniform effects is violated and that a spatially varying model, such as GWPR, may be more appropriate.

Step 4: Handling spatial weights and determine the optimal bandwidth for GWPR

Although spatial weight matrices (e.g., from contiguity) are standard in global spatial models, GWPR implemented in the GWmodel package utilizes internal weighting based on specified kernel functions and optimized bandwidths. Spatial weights can be helpful for initial diagnostics like checking for spatial autocorrelation in residuals, even if they are not required for GWPR.

4.1 Create a spatial weights matrix (optional for some GWPR models)

Optional: Create spatial weights for diagnostic analysis (e.g., Moran's I)

Create neighbor list from polygons

Ensure spatial_data is an sf polygon object

Generate neighbors list from a spatial polygon object (e.g., administrative boundaries)

`spatial\_data` must be an sf object with POLYGON geometries
neighbors <- poly2nb(spatial_data) # Create neighborhood list based on shared boundaries

Create a spatial weights matrix from the neighbor list
weights_matrix <- nb2listw(neighbors, style = "W") # "W" = row-standardized weights

View structure

summary(weights_matrix)

4.2 Determine the optimal bandwidth for GWPR

The bandwidth defines the spatial extent over which observations influence local model estimates. Selecting an optimal bandwidth is critical-small bandwidths may overfit, while large ones may oversmooth. The corrected Akaike information criterion (AICc) is often used to identify the bandwidth that offers the best trade-off.

Load GWmodel and convert sf to Spatial* format

Convert sf to Spatial object for compatibility with GWmodel
data_sp <- as(merged_data, "Spatial")

GWmodel requires Spatial* objects

Select optimal bandwidth using AICc

Choose optimal bandwidth for geographically weighted Poisson regression
bw_gwpr <- bw.gwr(formula = case ~ predictor 1 + predictor 2 + predictor3,

Define the model

data = data_sp,

approach = "AICc",

kernel = "bisquare", # Spatial weighting function

adaptive = TRUE, # Adaptive bandwidth adjusts to spatial density

family = "poisson" # Specify Poisson distribution for count data)

Output optimal bandwidth

print(bw_gwpr)

Step 5: Fit the GWPR model using the selected bandwidth

This step involves fitting a geographically weighted Poisson regression (GWPR) model to account for spatial heterogeneity in the relationship between health outcomes and covariates. The model is estimated using a spatial kernel and a previously selected optimal bandwidth. The kernel function defines how neighboring spatial units influence each other. The GWmodel package provides functions for this purpose, including gwpoisson() and gwr.basic(). gwpoisson() fits a Poisson model with spatially varying coefficients, using cross-validation for bandwidth selection; gwr.basic() is an alternative that applies a user-defined bandwidth (e.g., derived from AICc).

Load the required package

library(GWmodel)

Define the model formula

gwpr_formula <- cases ~ pollution + income # Replace with actual variables

Option 1: Fit GWPR model using gwpoisson() with bandwidth selected by cross-validation

gwpr_model <- gwpoisson(

formula = case ~ predictor 1 + predictor 2 + predictor 3,

data = merged_data, # Spatial object in sf format

```
listw = weights_matrix, # Optional spatial weights
for diagnostics
bw = "CV", # Bandwidth selected via cross-
validation
kernel = "bisquare" # Spatial weighting kernel
)
# Option 2: Fit GWPR model using gwr.basic
() with pre-determined optimal bandwidth
data_sp <- as(data_sf, "Spatial") # Convert to
Spatial* format
gwpr_model_alt <- gwr.basic (
formula = case ~ predictor 1 + predictor 2 +
predictor 3,
data = data_sp,
bw = bw_gwpr, # Bandwidth obtained in Step 4
family = "poisson",
kernel = "bisquare"
)
# Summarize model output
summary (gwpr_model)
summary (gwpr_model_alt)
```

Step 6: Model diagnostics and evaluation and analyze GWPR results

After fitting the geographically weighted Poisson regression (GWPR) model, it is crucial to evaluate its performance. Following GWPR model estimation, it is essential to assess model adequacy, identify potential misspecification, and interpret spatially varying relationships. This involves checking residuals for model fit, model fit evaluation, and analyzing the spatial variation in coefficients to detect any spatial non-stationarity.

6.1 Residual Diagnostics

Residual analysis helps determine the model's adequacy. Randomly distributed residuals suggest a good fit, while spatial patterns may indicate unaccounted heterogeneity.

```
# Extract and plot residuals from the GWPR
model
```

```
residuals <- gwpr_model$residuals
plot (residuals, main = "GWPR Residuals",
xlab = "Index", ylab = "Residuals")
```

6.2 Global vs local model comparison

Model performance can be compared using AICc, which penalizes model complexity. A lower AICc indicates better fit.

```
# Print model summary including AICc
# Compare AIC between global and local models
AIC(poisson_model)
# GWPR model AICc is included in its summary
print(gwpr_model)
# AICc is reported directly in the GWPR
model summary
# Lower = better fit
```

If GWPR yields a substantially lower AICc than the global model, this supports the inclusion of spatially varying coefficients.

6.3 Analyzing spatially varying coefficients

GWPR outputs localized coefficient estimates, enabling analysis of spatial non-stationarity.

```
# Extract spatially varying coefficients
# Convert spatial results to a data frame
gwpr_results <- as.data.frame(gwpr_model$SDF)
head(gwpr_results)
# Example: Plot spatial variation of the 'pollution'
coefficient
```

```
plot (gwpr_results$pollution, main = "Spatial
Variation in Pollution Coefficient",
xlab = "Location", ylab = "Coefficient Value")
```

Substantial variation across space indicates spatial non-stationarity. Areas with large or significant coefficients highlight locations where covariates exert stronger influence.

Step 7: Visualization of spatially varying coefficients

Visualizing the spatially varying coefficients and local model fit allows for intuitive interpretation of where covariate effects are strongest or weakest, offering valuable insights for spatial decision-making.

7.1 Mapping spatial coefficients

Mapping local coefficients provides a visual assessment of how the effect of each predictor (e.g., pollution, income) varies across space.

```
# Load required package library (tmap)
```

```
# Extract spatial object with coefficients from the GWPR model
```

```
spatial_results <- gwpr_model$SDF # SDF is a SpatialDataFrame containing coordinates and local coefficients
```

```
# Example: Visualize the spatial variation in the 'pollution' coefficient
```

```
tm_shape(spatial_results) + tm_borders() +  
tm_fill("predictor1", palette = "RdBu", style =  
"quantile", title = "Predictor1 Coefficient") + tm_  
layout (main.title = "Spatial Variation in Predictor 1  
Effect", legend.outside = TRUE)
```

The color gradient reflects spatial variation in the effect of pollution. Areas with higher (or lower) coefficients identify where pollution has a greater (or lesser) association with the outcome.

7.2 Visualizing local R^2 (goodness-of-fit)

Local R^2 values indicate how well the GWPR model explains variation in the dependent variable at each location. Mapping these values can identify regions where the model performs better or worse.

```
# Extract local  $R^2$  values
```

```
spatial_results$r_squared <- gwpr_model$SD-  
F$R2
```

```
# Map local  $R^2$ 
```

```
tm_shape(spatial_results) +
```

```
tm_borders() +
```

```
tm_fill("r_squared", palette = "YlOrRd", style =  
"quantile", title = "Local  $R^2$ ") +
```

```
tm_layout(main.title = "Model Fit (Local  $R^2$ )",  
legend.outside = TRUE)
```

Regions with higher R^2 values (darker shades) indicate better local model performance. Areas with low R^2 may suggest model misspecification or the need for additional covariates.

7.3 Scatter plot of spatial variation in predictor effects

A scatter plot using ggplot 2 can further highlight how a specific coefficient (e.g., pollution) varies geographically, especially if coordinate columns are available.

```
# Add coordinates to the results data frame
```

```
gwpr_results <- as.data.frame (spatial_results)
```

```
gwpr_results$longitude <- coordinates  
(spatial_results) [,1]
```

```
gwpr_results$latitude <- coordinates (spatial_  
results) [,2]
```

```
# Scatter plot of pollution effect
```

```
ggplot (gwpr_results, aes (x = longitude,  
y = latitude, color = predictor1)) + geom_point  
(size = 3) + scale_color_gradient 2 (low = "blue",  
mid = "white", high = "red",
```

```
midpoint = mean (gwpr_results$predictor 1,  
na.rm = TRUE)) + labs(title = "Spatial Distribution  
of Predictor1 Effect",
```

```
x = "Longitude", y = "Latitude", color =  
"Coefficient") + theme_minimal ()
```

Blue points reflect weaker or negative effects of pollution, while red points represent stronger positive effects. This helps identify spatial clusters where the predictor's influence is more pronounced.

Step 8: Interpretation and policy implications

Localized coefficients derived from GWPR facilitate a more refined interpretation of public health data. These estimates allow for the identification of areas where predictor variables exert a stronger or weaker influence on the outcome. For example, the effect of environmental pollution on disease incidence may be substantially greater in densely populated urban centers compared to rural regions. Such variation underscores the importance of spatial context in understanding health determinants and supports the development of location-specific hypotheses for further investigation.

The implications of these spatially explicit findings extend directly to public health policy and practice. By revealing regional differences in the strength of associations, GWPR outputs can inform the design and implementation of targeted interventions. Policymakers can allocate resources more efficiently by prioritizing regions where risk factors have the most pronounced effects. Moreover, this approach contributes to the promotion of health equity by addressing place-based disparities in exposure, vulnerability, and health outcomes.

Nevertheless, the application of GWPR is not without challenges. The method can be computationally intensive, especially when applied to large datasets or models incorporating multiple covariates. Furthermore, model outputs are sensitive to the selection of bandwidth, kernel function, and the spatial configuration of the data. These methodological considerations must be carefully addressed to ensure robust and interpretable results. Interpretation of spatially varying coefficients must also account for the local socio-environmental context in which they are estimated, avoiding overgeneralization or misrepresentation of localized effects.

In conclusion, GWPR offers a powerful analytical framework for uncovering spatial variability in the relationships between health outcomes and their predictors. By following the methodological steps outlined in this guide, researchers can implement GWPR in R to generate detailed spatial models that reveal localized associations. The insights gained through this process contribute to a more sophisticated understanding of public health phenomena and support the development of spatially responsive, evidence-based policies. While the technique demands careful consideration of computational and methodological issues, its capacity to reveal complex spatial patterns makes it an invaluable tool in spatial epidemiology and health geography.

References

- Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), 281–298.
- Bui, L. V., Mor, Z., Chemtob, D., Ha, S. T., & Levine, H. (2018). Use of geographically weighted Poisson regression to examine the effect of distance on tuberculosis incidence: A case study in Nam Dinh, Vietnam. *PLOS ONE*, 13(11), e0207068. <https://doi.org/10.1371/journal.pone.0207068>
- Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., Wang, Y., & Harris, P. (2020, April 13). The GWR route map: A guide to the informed application of geographically weighted regression. arXiv. Retrieved March 10, 2025 From <https://arxiv.org/abs/2004.06070>

- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Wiley.
- Fotheringham, A. S., Brunson, C., & Charlton, M. E. (2009). Geographically weighted regression. *The Sage handbook of spatial analysis*, 1, 243-254.
- Fundisi, E., Dlamini, S., Mokhele, T., Weir-Smith, G., & Motolwana, E. (2023). Exploring determinants of HIV/AIDS self-testing uptake in South Africa using generalized linear Poisson and geographically weighted Poisson regression. *Healthcare*, 11(6), 881. <https://doi.org/10.3390/healthcare11060881>
- Hadayeghi, A., Shalaby, A. S., & Persaud, B. N. (2010). Development of planning level transportation safety tools using geographically weighted Poisson regression. *Accident Analysis & Prevention*, 42(2), 676-688.
- Helmy, H., Kamaluddin, M. T., Iskandar, I., & Suheryanto. (2022). Investigating spatial patterns of pulmonary tuberculosis and main related factors in Bandar Lampung, Indonesia using geographically weighted Poisson regression. *Tropical Medicine and Infectious Disease*, 7(9), 212.
- Kirby, R. S., Delmelle, E., & Eberth, J. M. (2017). Advances in spatial epidemiology and geographic information systems. *Annals of Epidemiology*, 27(1), 1-9. <https://doi.org/10.1016/j.annepidem.2016.12.001>
- Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics-Theory and Methods*, 26(6), 1481-1496.
- Li, G., Wang, S., Liu, X., Bigham, J. M., & Ragland, D. R. (2013). Using geographically weighted Poisson regression for county-level crash modeling in California. *Safety Science*, 58, 89-97.
- McMillen, D. P. (2004). Geographically weighted regression: The analysis of spatially varying relationships. *Journal of Regional Science*, 44(1), 1-22.
- Munira, S., & Sener, I. N. (2020). A geographically weighted regression model to examine the spatial variation of the socioeconomic and land-use factors associated with Strava bike activity in Austin, Texas. *Journal of Transport Geography*, 88, 102865. <https://doi.org/10.1016/j.jtrangeo.2020.102865>
- Murakami, D., Tsutsumida, N., Yoshida, T., Nakaya, T., Lu, B., & Harris, P. (2023). A linearization for stable and fast geographically weighted Poisson regression. *International Journal of Geographical Information Science*, 37(8), 1818-1839.
- Nakaya, T., Fotheringham, A. S., Brunson, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, 24(17), 2695-2717.
- Poliart, A., Kirakoya-Samadoulougou, F., Ouédraogo, M., Collart, P., Dubourg, D., & Samadoulougou, S. (2021). Using geographically weighted Poisson regression to examine the association between socioeconomic factors and hysterectomy incidence in Wallonia, Belgium. *BMC Women's Health*, 21(1), 373.

- Saefuddin, A., Saepudin, D., & Kusumaningrum, D. (2013). Geographically weighted Poisson regression (GWPR) for analyzing the malnutrition data in Java-Indonesia. *Procedia Social and Behavioral Sciences*, 101, 100-108.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(sup1), 234-240.
- Tyas, S. W., Gunardi, & Puspitasari, L. A. (2023). Geographically weighted generalized Poisson regression model with the best kernel function in the case of the number of postpartum maternal mortality in East Java. *Methods X*, 10, 102002. <https://doi.org/10.1016/j.mex.2023.102002>
- Wheeler, D. C. (2021). Geographically weighted regression. In M. M. Fischer & P. Nijkamp (Eds.), *Handbook of regional science* (pp.1895-1921). Springer.