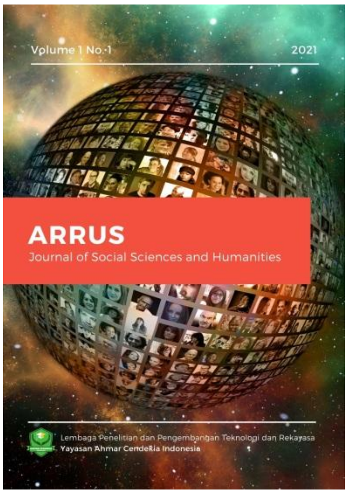
OPEN ACCESS

ISSN: 2776-7930 (Print) / 2807-3010 (Online)



RESEARCH ARTICLE

Geographically Weighted Poisson Regression (GWPR) Model with Fixed Gaussian Kernel and Fixed Bi-square Kernel Weights

Sitti Masyitah Meliyana*, Ansari Saleh Ahmar, Siti Nurazizah Auliah

Statistics Study Program, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Makassar, Indonesia.

*Corresponding author: Sitti Masyitah Meliyana, Statistics Study Program, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Makassar, Indonesia

E-mail: sittimasyitahmr@unm.ac.id

Abstract: This study aims to model the spatial distribution of tuberculosis (TB) cases in Makassar City in 2022 using the Geographically Weighted Poisson Regression (GWPR) approach. This method extends Poisson regression by incorporating spatial heterogeneity, weighting each location based on its geographical proximity. Two types of kernel weighting functions, fixed Gaussian kernel and fixed bi-square kernel, were used to determine the most effective model for identifying key factors influencing TB case numbers. The parameter estimation results indicate that the GWPR model with fixed bi-square kernel performs better than the global Poisson regression model, achieving an Akaike's Information Criterion (AIC) value of 97.69 and a coefficient of determination (R^2) of 99.93%. The findings reveal that the relationship between predictor variables and TB cases varies across districts, with the percentage of the productive-age population and population density emerging as dominant factors. These results highlight the advantages of the GWPR approach in capturing spatial dynamics more effectively than conventional regression models, making it a powerful analytical tool for designing targeted, region-specific public health interventions.

Keywords: Geographically Weighted Poisson Regression, Fixed Gaussian Kernel, Fixed Bi-square Kernel, Spatial Analysis, Tuberculosis.

1. Introduction

Regression analysis is a statistical method used to examine the relationship between one or more independent variables and a dependent variable (Astriawati, 2016). The Ordinary Least Squares (OLS) method is commonly used in regression analysis, but it has limitations as it does not account for the spatial structure of data. To address this issue, spatial regression was introduced, incorporating spatial aspects to capture relationships between neighboring observational units (Sejati et al., 2022). However, spatial regression still struggles to handle local variations (Nurmasari, 2016).

Geographically Weighted Regression (GWR) provides a more adaptable method by assigning unique weights to each data point according to its spatial proximity to the location under analysis (Widyaningsih & Fitrianingrum, 2022). This method incorporates the idea that the impact of independent variables can differ across geographic locations, thereby capturing dynamic spatial patterns. Building on this concept, Geographically Weighted Poisson Regression (GWPR) extends the principles of GWR to model dependent variables that follow a Poisson distribution, which is especially beneficial for modeling event frequencies in

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

specific regions (Arini & Nanih, 2022). GWPR is particularly effective in addressing intricate spatial variations in data distributions (Ananda et al., 2023).

Conventional Poisson regression models do not account for spatial variability, despite the fact that spatial heterogeneity often exists, where a predictor's influence may differ from one location to another. Geographically Weighted Poisson Regression (GWPR) addresses this limitation by incorporating spatial context through a kernel-based weighting mechanism, either fixed or adaptive (Tizona et al., 2017), enabling the model to better capture location-specific effects in Poisson-distributed outcome variables.

Tuberculosis (TB) remains one of the leading infectious diseases globally, with mortality and case rates continuing to rise (Supriyanti, 2022). In Makassar City, the burden of TB is particularly severe, with 5,444 cases reported out of an estimated 14,000 in 2022. Applying Geographically Weighted Poisson Regression (GWPR) to TB data is essential, as it accommodates complex spatial heterogeneity and allows for the identification of localized transmission patterns. This spatially sensitive approach facilitates the development of more targeted public health strategies. Moreover, GWPR improves the precision of regression modeling by detecting high-risk clusters, thereby supporting more effective disease control efforts (Tuasikal, 2018; Septiani, 2021; Helmy et al., 2022).

This study focuses on the Geographically Weighted Poisson Regression (GWPR) model using fixed Gaussian kernel and fixed bisquare kernel weighting, applied to TB case data in Makassar City.

2. Literature Review

2.1. Poisson Distribution

The Poisson distribution is a probability model used to describe events that occur with a low probability, where the events are influenced by a specific time frame or geographical region, and the outcomes are represented by discrete variables (Otaya, 2016). It is one of the simplest models for analyzing count data, which consists of non-negative integer values (Budiharti, 2021).

The probability mass function (PMF) of the Poisson distribution is expressed as follows (Walpole, 1995):

$$p(y;\mu) = \frac{e^{-\mu}\mu^{y}}{y!} \quad (y = 0, 1, 2, \dots.)$$

Where μ denotes the expected number of occurrences in a given time period or area, and y represents the observed number of events during that period or in that region.

2.2. Poisson Regression

The Poisson regression model is a type of regression analysis based on the Poisson distribution, often utilized for examining data where the response variable is discrete. It is classified as a nonlinear regression model (Kusuma et al., 2013). A defining feature of Poisson regression is equidispersion, which indicates that the mean and variance are equal. The relationship between the response variable (Y) and the predictor variable (X) is represented as follows (Esra et al., 2023):

$$E\ (((Y_i|X_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon_1$$

Thus, the Poisson model can be written as follows:

$$E\ (((Y_i|X_i) = \mu i = \exp(\ x_i^T \beta)\,, i = 1,2, \dots, n$$

where β represents the unknown parameter that needs to be estimated, and x_i is the independent variable.

https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

2.2.1. Parameter Estimation in Poisson Regression

The *Maximum Likelihood Estimation* (MLE) method is commonly used to estimate the parameters of the Poisson regression model (Fauziah, 2021). The log-likelihood function, maximized using the MLE method, is given as follows:

$$\frac{\partial \ln L\ (\beta)}{\partial\ \beta^T} = \sum_{i=1}^{n} (-x_i \exp(\ x_i^T \beta) +\ y_i x_i$$

2.2.2. Partial Test

The parameters estimated in the model may not always exhibit a substantial effect. Hence, it is necessary to conduct a partial or individual test to evaluate the significance of each parameter. The hypotheses for this test are outlined as follows:

H_0: $\beta_k = 0$ (The k-th variable has no significant effect)

H_1: $\beta_k \neq 0$ (The k-th variable has a significant effect)

The test statistic is given as:

$$z = \frac{\hat{\beta}_k}{se\ (\hat{\beta}_k)}\ ;$$

2.2.3. Spatial Heterogeneity Test

Spatial heterogeneity testing is conducted to assess whether the data of the response variable show point-based spatial heterogeneity. This can be evaluated using the Breusch-Pagan (BP) test.

2.3. Geographically Weighted Regression (GWR)

Geographically Weighted Regression (GWR) is an advanced method used to model spatial heterogeneity. Spatial heterogeneity refers to the variations or differences in characteristics across various geographic areas (Soraya et al., 2016). The spatial heterogeneity model in GWR can be represented as follows (Wang et al., 2014):

$$y_i = \beta_0 (u_i\ , v_i\ ) + \sum_{j=1}^{k}\ \beta_j\ (u_i\ , v_i\ )\ x_{ij} + \varepsilon_i,\ \ (i = 1, 2, \dots n)$$

In this model, yi represents the observed value of the response variable at observation i, while xij refers to the observed value of the predictor variable j at observation i. The term $\beta_0(ui,\ vi)$ denotes the intercept of the regression model, and βj$(ui,\ vi)$ indicates the regression coefficient for j = 0, 1, 2, ..., k. Furthermore, ui and vi represent the spatial coordinates of observation i, and εi is the error term associated with observation i.

2.4. Geographically Weighted Poisson Regression (GWPR)

Geographically Weighted Poisson Regression (GWPR) is an extension of the Poisson regression model that incorporates spatial variations within the data. Unlike traditional global Poisson regression, where the parameter estimates remain constant across all locations, GWPR enables these estimates to vary according to geographic position (Millah, 2015). This characteristic makes GWPR a more adaptable and region-specific method, allowing it to capture spatial heterogeneity with greater precision (Qomariyah et al., 2013).

For a given location i, with coordinates (ui,vi), the GWPR model is expressed as:

$$\mu_i\ (u_i, v_i) = \exp(\ X_i^T \hat{\beta}(\ u_i, v_i)),\ i = 1,2, \dots, n$$

With

$$\hat{\beta} = (u_i, v_i) = [\hat{\beta}(u_{i,} v_i) \beta_1(u_{i,} v_i)\ \beta_2(u_{i,} v_i) \dots\ \beta_p(u_{i,} v_i)]^T$$

https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

2.5. Geographically Weighted Poisson Regression (GWPR)

In GWPR, the kernel weighting function incorporates a parameter known as bandwidth, which can be interpreted as the radius (**b**) of a circle that determines the influence of nearby data points. To find the optimal bandwidth, one widely applied approach is Cross-Validation (CV), which is computed using the following formula (Sogen et al., 2023):

$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(\boldsymbol{b})]^2$$

This research employs two different types of spatial weighting functions: the fixed Gaussian kernel and the fixed bi-square kernel.

- Fixed gaussian kernel: $w_{ij}(u_{i,}v_i) = \exp\left(-\left(\frac{d_{ij}}{b}\right)^2\right)$
- Fixed bi-square kernel : $w_{ij}(u_{i,}v_i) = \begin{cases} \left(1-\left(\frac{d_{ij}}{b}\right)^2\right)^2 & , untuk\ d_{ij} \leq b \\ 0 & , untuk\ d_{ij} > b \end{cases}$

where b is a non-negative parameter referred to as the bandwidth, b_i represents the bandwidth at location i, and d_{ij} denotes the Euclidean distance between locations i and j.

2.6. Selecting the Best Model

2.6.1. Akaike Information Criterion (AIC)

In spatial modeling, the Akaike Information Criterion (AIC) is a commonly used tool for selecting the most appropriate regression model. The optimal model is identified by the lowest AIC value, as a lower AIC signifies a better trade-off between model complexity and goodness of fit. Developed by Akaike, this method relies on the Maximum Likelihood Estimation (MLE) approach. The AIC formula is given by the following equation (Sogen et al., 2023):

$$AIC = 2k - 2\ \ln(likelihood)$$

where k denotes the number of estimated parameters, and the likelihood refers to the maximum likelihood value of the model.

2.6.2. Coefficient of Determination (R^2)

The coefficient of determination (R^2), commonly referred to as R-squared, assesses the extent to which a regression model accounts for the variation in the observed data. It offers an indication of the model's predictive accuracy, with higher values suggesting a closer fit to the data. The formula for R^2 is expressed as:

$$R^2 = \frac{SSR}{SST}$$

where SSR (Sum of Squares for Regression) refers to the portion of variation explained by the model, and SST (Total Sum of Squares) represents the overall variation in the data.

2.6.3. Tuberculosis Cases

Spatial statistical methods have become crucial tools in public health and epidemiology, especially for detecting geographical differences in disease patterns and evaluating spatial risk factors. Traditional regression models, such as Poisson regression, have been commonly used to model count data, including disease incidence. However, these models typically assume spatial stationarity, which may not be applicable in real-world situations where the relationship between predictors and outcomes changes across different locations (Fotheringham et al., 2002).

Geographically Weighted Poisson Regression (GWPR) is an advanced version of Poisson regression that accounts for spatial heterogeneity by allowing the regression coefficients to vary geographically (Nakaya et al., 2005). This method provides more localized insights into

https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

how explanatory variables influence outcomes in various areas, making it especially valuable in public health research, where intervention strategies may need to be tailored to specific locations.

Recent studies have highlighted the effectiveness of GWPR in analyzing spatial patterns of different public health outcomes. For example, Lin et al. (2020) used GWPR to examine the spatial determinants of dengue fever in Taiwan, revealing that the effects of climatic and demographic factors varied across regions. Likewise, Huang et al. (2021) applied GWPR to assess COVID-19 transmission risks in urban China, identifying neighborhood level variations that global models would have missed.

In Indonesia, spatial epidemiology has gained traction, with GWPR being increasingly used to understand the spread of infectious diseases like tuberculosis (TB), malaria, and dengue. A study by Sari et al. (2022) used GWPR to analyze TB incidence in Central Java, finding that population density and access to healthcare facilities had regionally varying effects.

Despite its growing application, GWPR is still underused in health studies at the local level, particularly in urban areas like Makassar City. This study contributes to the existing literature by applying GWPR to TB incidence data in Makassar, comparing the performance of models using both fixed Gaussian and bi-square kernels, and investigating local variations in associated risk factors. By integrating spatial methods with epidemiological modelling, this research aims to provide evidence-based insights to help develop targeted TB control strategies.

3. Research Method and Materials

This research employs a quantitative approach, concentrating on the systematic gathering of objective data and subsequent numerical analysis. To investigate spatial patterns in tuberculosis (TB) cases, the study utilizes Geographically Weighted Poisson Regression (GWPR) as the main analytical technique. The data for this study were obtained from the Health Office of South Sulawesi Province and the South Sulawesi branch of Statistics Indonesia (BPS), available through their official website (https://www.bps.go.id).

To develop the GWPR model for TB cases, the analysis was conducted in the following steps:

(a). Descriptive Statistical Analysis: Summarizing key characteristics of the dataset.
(b). Multicollinearity Detection: Assessing correlations between predictor variables using the Variance Inflation Factor (VIF) test.
(c). Poisson Regression Analysis: Estimating Poisson regression model parameters, Evaluating parameter significance through both simultaneous and partial tests. Calculating Akaike Information Criterion (AIC) and R^2 to assess model performance.
(d). Spatial Heterogeneity Testing: Applying the Breusch-Pagan (BP) test to examine variations in spatial relationships.
(e). GWPR Model Development: Computing Euclidean distances between observation locations based on latitude and longitude.
    (1). Identifying location-specific GWPR estimations.
    (2). Determining the bandwidth selection range based on maximum Euclidean distance per location.
    (3). Testing various bandwidth values within the defined range.
    (4). Applying fixed Gaussian kernel and fixed bi-square kernel spatial weighting methods.
    (5). Estimating GWPR model parameters using the selected bandwidth.
    (6). Identifying the optimal bandwidth for GWPR analysis.
    (7). Performing significance tests on GWPR model parameters.
    (8). Calculating AIC and R^2 to evaluate model accuracy.
    (9). Interpreting the GWPR results for specific locations.

https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

(f). Model Selection: Evaluating and comparing the Poisson regression model with the GWPR models (employing fixed Gaussian and fixed bi-square kernels) using the AIC and R^2 criteria.
(g). Identifying Key Determinants: Analyzing which predictor variables significantly influence TB incidence in various regions.

4. Results and Discussion

4.1. Descriptive Analysis

This study relies on secondary data collected from the Makassar City Statistics Agency (BPS) and the Makassar City Health Office. The dataset includes both response and predictor variables, which help analyze the distribution of tuberculosis (TB) cases in Makassar City. The response variable (Y) represents the total number of TB cases recorded in 2022. for selecting the predictor variables by referring to relevant public health and epidemiological studies on tuberculosis (TB). The productive age population (X1) is included because individuals in this age group are not only more active socially and economically but also have a higher risk of exposure and transmission, making them a key group in TB control efforts. The availability of healthcare facilities (X2) reflects access to diagnostic services and treatment, which is crucial for early detection and successful TB management.

The percentage of the population living below the poverty line (X3) is an important socioeconomic indicator, as poverty is closely linked to poor living conditions, undernutrition, and limited access to healthcare—factors that significantly contribute to TB vulnerability. Meanwhile, the number of households practicing clean and healthy living behaviors (PHBS) (X4) captures the extent of preventive practices at the community level. These behaviors, such as maintaining ventilation and hygiene, can reduce the risk of TB transmission in residential settings.

Lastly, population density (X5) is considered due to its role in facilitating the spread of airborne diseases. Areas with high population density may create conditions that increase the likelihood of transmission through close and frequent interpersonal contact. The predictor variables include:
X1: Percentage of the population in the productive age group
X2: Number of available healthcare facilities
X3: Percentage of the population living below the poverty line
X4: Number of households practicing clean and healthy living behaviors (PHBS)
X5: Population density
The following section presents a descriptive analysis of the study data, which consists of one response variable and five predictor variables.

Table 1: Descriptive Statistics

Variable	Min	Max	Mean	Standard Deviation
Y	38.0	484.0	281.7	155.09
X1	1.04	12.35	6.66	4.04
X2	2.00	12.00	6.46	3.33
X3	0.54	6.48	2.94	1.86
X4	1169	32561	12913	9204.38
X5	3245	32645	15285	10144.52

Meliyana, et.al., *ARRUS Journal of Social Sciences and Humanities*, Vol. 5, No. 2 (2025)
https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

To incorporate spatial analysis, the study also records the geographic coordinates (latitude and longitude) of 15 districts in Makassar City, allowing for a more detailed examination of spatial patterns in TB case distribution.

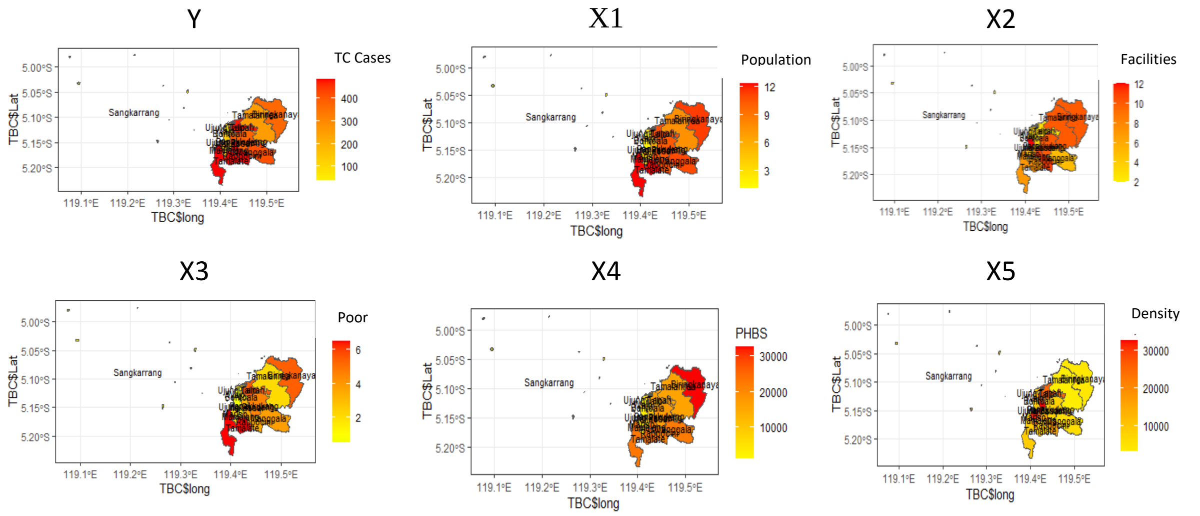


Figure 1: Map of Predictor and Response Variable Distribution in Makassar City

Figure 1 highlights five districts with the highest tuberculosis cases: Tamalate, Rappocini, Tallo, Manggala, and Makassar. Kepulauan Sangkarrang has the lowest percentage of the productive-age population, while Bontoala and Tallo have the fewest healthcare facilities. The highest percentage of impoverished populations is found in Tamalate, Tallo, Biringkanaya, Panakkukang, and Manggala. Additionally, Kepulauan Sangkarrang has the lowest number of households practicing clean and healthy living behaviors (PHBS).

4.2. Multicollinearity Detection

Prior to developing the Poisson Regression and Geographically Weighted Poisson Regression (GWPR) models, it is essential to assess the presence of multicollinearity among the predictor variables.

Table 2: VIF values for each predictor variableorrelation Analysis

Variable	VIF Value	Multicollinearity Indication
Percentage of Productive-Age Population (X_1)	7.945	No Multicollinearity
Number of Healthcare Facilities (X_2)	1.775	No Multicollinearity
Percentage of Impoverished Population (X_3)	9.443	No Multicollinearity
Number of Households with PHBS (X_4)	9.686	No Multicollinearity
Population Density (X_5)	1.413	No Multicollinearity

Table 2 indicates that the Variance Inflation Factor (VIF) values for all predictor variables are below 10, suggesting that there is no multicollinearity present. As a result, all variables are suitable for inclusion in the Poisson Regression modelling process.

4.3. Poisson Regression Model

In Poisson regression modeling, the initial step involves performing a simultaneous parameter test to assess whether the predictor variables have a collective impact on the response variable. The hypotheses for this test are outlined as follows:

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

H_1: At least one $\beta_k \neq 0$, for k = 1, 2, 3, 4, 5

Table 3: Simultaneous Hypothesis Testing of Poisson Regression Parameters

	Loglike	Df	x^2	p-Value
Parameter Response	-121.14	5	1207.5	2.2×10^{-16}

Based on the results presented in Table 3, the test decision is to reject H_0, as indicated by the chi-square value $\chi^2 = 1207.5$ with a p-value of 2.2×10^{-16}. This indicates that at least one of the predictor variables has a statistically significant effect on the response variable. Consequently, the next step is to perform a partial test on the Poisson regression model.

Table 4: Partial Hypothesis Testing of Poisson Regression Parameters

Variable and Parameter	Estimator ($\hat{\beta_k}$)	Standard Error	Z-statistic	p-Value	Test Decision
β_0	$0.399\text{x}10^1$	8.019×10^{-2}	49.760	$2\text{x}10^{-16}$	Reject H_0
X1(β_1)	1.816×10^{-1}	1.503×10^{-2}	12.083	$2\text{x}10^{-16}$	Reject H_0
$X2(\beta_2)$	1.807×10^{-2}	6.968×10^{-3}	2.594	9.49×10^{-3}	Reject H_0
$X3(\beta_3)$	-4.844×10^{-2}	2.338×10^{-2}	-2.071	3.832×10^{-2}	Reject H_0
$X4(\beta_4)$	-2.99×10^{-6}	4.933×10^{-6}	-0.607	5.4354×10^{-1}	Fail to Reject H_0
$X5(\beta_5)$	2.225×10^{-5}	2.106×10^{-6}	-2.469	$2\text{x}10^{-16}$	Reject H_0

Based on the calculations in Table 4, β_0 is significant. The variables Percentage of Productive-Age Population (X1), Number of Healthcare Facilities (X2), Percentage of Impoverished Population (X3), and Population Density (X5) each have a significant individual effect on the number of TB cases in Makassar City. This is indicated by their respective Z-statistic values and p-values, all of which are below 0.05.

On the other hand, the variable Number of Households Practicing Clean and Healthy Living Behaviors (PHBS) (X4) does not have a significant individual effect on TB cases in Makassar City, as its Z-statistic p-value is greater than 0.05.

Thus, the Poisson Regression model for TB cases in Makassar City is formulated as follows:

$$\mu = \exp\,(3.990 + \ 0.816x_1 + 0.018x_2 - \ 0.048x_3 + 0.000022x_5)$$

Based on calculations using R software, the Akaike Information Criterion (AIC) for the Poisson Regression model is 254.29, with an R^2 value of 90.06%, indicating a high level of model accuracy.

The next step is to conduct a spatial heterogeneity test to determine whether spatial heterogeneity exists in the response variable data. This test is performed using the Breusch-Pagan method.

Table 5: Spatial Heterogeneity Test

Breusch-Pagan	p-value	Test Decision
10.882	0.04377	5

Based on the calculations in Table 4.6, the test decision is to reject H_0 at a 0.05 significance level. This is indicated by a p-value of 0.04377, which is less than $\alpha = 0.05$. This result confirms the presence of spatial heterogeneity in the response variable data, meaning that there are differences in characteristics between regions. Therefore, the data can be appropriately modeled using Geographically Weighted Poisson Regression (GWPR).

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

4.4. GWPR Modelling

The initial step in estimating the GWPR model is to determine the locations where the model will be applied. Makassar City consists of 15 districts, meaning that the Euclidean distances between districts form a 15 × 15 matrix. The second step is to determine the optimal bandwidth for each observation location using the Cross-Validation (CV) criterion, ensuring appropriate spatial coverage between regions. In this process, two weighting criteria are applied: the fixed Gaussian kernel and the fixed bi-square kernel.

Table 6: Bandwidth values for the fixed gaussian kernel and fixed bisquare kernel at each location.

Location	Bandwidth	
	fixed gaussian kernel	fixed bisquare kernel
Mamajang	0.038183	0.030632
Manggala	0.075556	0.068678
Mariso	0.036245	0.031777
Sangkarrang	0.347720	0.343458
Rappocini	0.044063	0.042230
Tamalate	0.062677	0.052403
Makassar	0.029003	0.024790
Ujung Pandang	0.035822	0.024577
Panakkukang	0.043282	0.041501
Bontoala	0.033325	0.025937
Wajo	0.037588	0.032674
Ujung Tanah	0.045141	0.032576
Tallo	0.038920	0.034597
Tamalanrea	0.064466	0.062882
Biringkanaya	0.100553	0.098703

The bandwidth value differs at each observation location, indicating that each observation area is associated with a unique bandwidth.

4.4.1. Estimation of GWPR Model Parameters

The GWPR model with fixed bi-square kernel weighting demonstrates the relationship between the number of TB cases across 15 districts in Makassar City and the predictor variables. The estimated models for Location 1 (Mamajang District), Location 2 (Manggala District), and Location 3 (Mariso District) are presented as follows:

$$\hat{\mu}\,(u_1 v_1) = exp(268.32 - 1037.2X_1 + 562.03X_3 + 0.553X_4 + 0.0128X_5)$$

$$\hat{\mu}\,(u_2 v_2) = exp(-133.63+0.006X_5)$$

$$\hat{\mu}\,(u_3 v_3) = exp(524.92 - 1030.62X_1 - 28.14X_2 + 443.43X_3 + 0.43X_4 + 0.004X_5)$$

The GWPR model with fixed Gaussian kernel weighting illustrates the relationship between the number of TB cases across 15 districts in Makassar City and the predictor variables. The estimated models for Location 2 (Mangala District), Location 3 (Mariso District), and Location 15 (Biringkanaya District) are presented as follows:

$$\hat{\mu}\,(u_2 v_2) = exp(-44.56 + 0.003X_5)$$

https://doi.org/10.35877/soshum3812

OPEN ACCESS

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

$\hat{\mu}\ (u_3 v_3) = exp(-31.41 + 0.002X_5)$

$\hat{\mu}\ (u_{15} v_{15}) = exp(-35.69 + 43.99X_1)$

A comparison was made between the Poisson regression model and the GWPR models utilizing Fixed Bi-square Kernel and Fixed Gaussian Kernel weighting to identify the most appropriate model for analyzing TB cases in Makassar City in 2022.

4.4.2. *The Best Model Selection*

The criteria for model selection were based on Akaike's Information Criterion (AIC) and the Coefficient of Determination (R^2), with a lower AIC value and a higher R^2 value indicating a model that fits the data better. The outcomes of the model selection process are shown in **Table 7**.

Table 7: The best model selection

Model Prediction	AIC	R^2
Poisson Regression	254.29	90.06%
GWPR Fixed Bi-square Kernel	97.69	99.93%
GWPR Fixed Gaussian Kernel	153.24	95.99%

According to Table 7, the optimal model for forecasting the number of TB cases in Makassar City is the GWPR model with Fixed Bi-square Kernel weighting. This is evidenced by the model's lowest AIC value of 97.69 and the highest Coefficient of Determination (R^2) of 99.93%. Therefore, it can be concluded that the GWPR model with Fixed Bi-square Kernel weighting is the most appropriate model for analyzing TB cases in Makassar City.

4.5. *Interpretation of the Best Model*

The GWPR model with Fixed Bi-square Kernel weighting can be divided into eight categories, determined by the predictor variables that significantly impact the number of TB cases. These categories are outlined in Table 8 below:

Table 8: Classification of Predictor Variables Significantly Influencing TB Cases

Location	Significant Variable
Bontoala, Tamalanrea, Biringkanaya	(X1)
Manggala, ujung Pandang	(X5)
Sangkarrang, Wajo	(X1) and (X4)
Panakkukang	(X2) and (X5)
Tallo	(X3) and (X5)
Rappocini, Ujung Tanah	(X1), (X2) and (X3)
Mamajang, Tammalate, Makassar	(X1), (X3), (X4) and (X5)
Mariso	(X1), (X2), (X3), (X4) and (X5)

According to Table 8, the classification of significant variables using the GWPR method with Fixed Bi-square Kernel weighting indicates that all predictor variables significantly affect the response variable.

The significance of each variable is assessed by comparing the t-statistic at each observation point with the critical t-value at a 0.05 significance level for 14 degrees of freedom, which is 1.77. The decision rule is to reject H_0 if t-statistic > t-value (1.77) at a given observation location. As an example, the t-statistic values for each observation location in Mariso District are presented in Table 9 below.

https://doi.org/10.35877/soshum3812

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

Table 9: t-Statistic Values for Partial Testing of GWPR Fixed Bisquare Kernel Model Parameters

Variable	t-Statistic
X1	5.741
X2	3.934
X3	5.655
X4	5.703
X5	3.318

The map in Figure 2 displays the classification of the GWPR model with fixed bi-square kernel weighting, highlighting the predictor variables that have a significant impact on the number of TB cases across 15 districts of Makassar City.

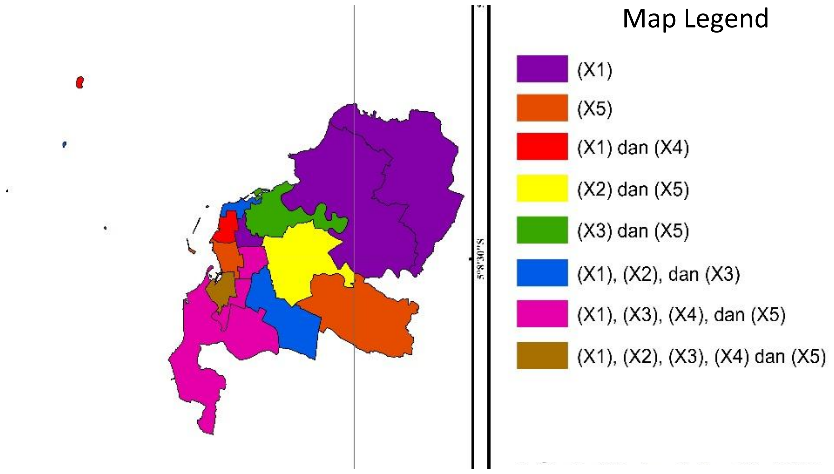


Figure 2: The classification of the GWPR model with fixed bi-square kernel weighting

Based on Figure 2, the map of Makassar City highlights different regions with distinct factors influencing TB cases: Purple-colored areas represent districts Bontoala, Tamalanrea, and Biringkanaya, where the percentage of the productive-age population (X1) is the primary influencing factor. Orange-colored areas indicate that in Manggala and Ujung Pandang districts, population density (X5) is the significant factor affecting TB cases. Blue-colored areas show that in Rappocini and Ujung Tanah districts, the influencing factors include the percentage of the productive-age population (X1), number of healthcare facilities (X2), and percentage of the impoverished population (X3). Similar classifications apply to other district groups as represented on the map.

Thus, the GWPR model with Fixed Bisquare Kernel weighting varies across districts. For example, in Mariso District, the estimated model is as follows:

$$\hat{\mu}(u_3 v_3) = \exp(524{,}92 - 1030{,}62X1 - 28{,}14X2 + 443{,}43X3 + 0{,}43X4 + 0{,}0047X5)$$

Based on this model, the following interpretations can be made: A one-unit increase in the percentage of the productive-age population (X1) is estimated to decrease the number of TB cases by a factor of exp (-1030.62) = 2.55×10^{-4}. A one-unit increase in the number of healthcare facilities (X2) is estimated to decrease TB cases by a factor of exp (-28.14) = 6.01×10^{-13}. A one-unit increase in the percentage of the impoverished population (X3) is estimated to increase TB cases by a factor of exp (443.43) = 3.79×10^{19}. A one-unit increase in the number of households practicing PHBS (X4) is estimated to increase TB cases by a factor of exp (0.43) = 1.537. A one-unit increase in population density (X5) is estimated to increase TB cases by a factor of exp (0.0047) = 1.0047. However, the result for PHBS households (X4) appears unrealistic, as an increase in PHBS households should theoretically reduce TB cases Several contextual factors may help explain the unexpected results regarding

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

PHBS (Clean and Healthy Living Behavior). Although PHBS programs are widely promoted across Indonesia, the reported number of households categorized as compliant may not accurately reflect actual behavioral practices. This discrepancy may stem from the reliance on self-reported data or limited observational verification, potentially introducing bias and overestimating adherence levels (Ministry of Health RI, 2021).

Furthermore, environmental conditions commonly found in densely populated urban areas, such as household crowding, inadequate ventilation, and insufficient natural lighting, can limit the effectiveness of individual healthy practices in preventing the airborne spread of tuberculosis. Thus, even with proper PHBS implementation at the household level, unfavorable living conditions may still facilitate TB transmission (Lönnroth et al., 2009). Another plausible explanation is that regions reporting higher PHBS rates may also possess stronger public health infrastructure and surveillance systems, resulting in increased case detection rather than a true elevation in TB incidence (Sharma & Mohan, 2020).

These considerations suggest that while PHBS remains a valuable public health initiative, its measurement alone may not sufficiently capture the multifaceted interaction between individual behavior, environmental exposure, and healthcare system capacity in influencing TB transmission risk.

5. Conclusion

Based on the conducted tests, the following conclusions can be drawn:

The models produced using GWPR with Fixed Bi-square Kernel weighting and GWPR with Fixed Gaussian Kernel weighting exhibit spatial variation across different districts in Makassar City. For instance, the GWPR model with Fixed Bi-square Kernel weighting for the Mariso District is represented as follows:

$$\hat{\mu}(u_3 v_3) = \exp(524.92 - 1030.62X1 - 28.14X2 + 443.43X3 + 0.43X4 + 0.0047X5)$$

The significant factors influencing TB cases in Mariso District, Makassar City, in 2022 include the percentage of the productive-age population (X1), number of healthcare facilities (X2), percentage of the impoverished population (X3), number of households practicing PHBS (X4), and population density (X5).

References

Aliu, M. A., Zubedi, F., Yahya, L., & Oroh, F. A. (2022). The Comparison of Kernel Weighting Functions in Geographically Weighted Logistic Regression in Modeling Poverty in Indonesia. Jurnal Matematika, Statistika dan Komputasi, 18(3), 362-384.

Ananda, N. M. S., Suyitno, S., & Siringoringo, M. (2023). Geographically Weighted Panel Regression Modelling of Human Development Index Data in East Kalimantan Province in 2017-2020. Jurnal Matematika, Statistika dan Komputasi, 19(2), 323-341.

Arini, R. W., & Nanih, S. (2022). Analisis Sistem Antrian Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan: Studi Kasus Puskesmas Margadadi. Jurnal Riset Rumpun Ilmu Teknik (JURRITEK), 1(1), 23-37.

Astriawati, N. (2016). Penerapan Analisis Regresi Linier Berganda Untuk Menentukan Pengaruh Pelayanan Pendidikan Terhadap Efektifitas Belajar Taruna Di Akademi Maritim Yogyakarta. Majalah Ilmiah Bahari Jogja, 14(23), 22-37.

Budiharti, L. M. (2021). Pemodelan dan Pemetaan Jumlah Penderita Kusta di Jawa Barat dengan Regresi Binomial Negatif dan Flexibly Shaped Spatial Scan Statistic. Jurnal Riset Statistika, 99-106.

Esra, R., Nohe, D. A., & Fathurahman, M. (2023). Pemilihan Model Terbaik pada Generalized Poisson Regression Menggunakan Akaike Information Criterion. Statistika, 23(1), 73-87.

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

Fairuz, F. (2023). Identifikasi Mycobacterium Tuberculosis Dengan Uji xpert mtb/rif Pada Formalin-Fixed Paraffin Embedded (ffpe) Tuberkulosis Ekstraparu. Infokes: Jurnal Ilmiah Rekam Medis dan Informatika Kesehatan, 13(2), 78-82.

Fathurahman, M. (2010). Pemilihan model regresi terbaik menggunakan akaike's information criterion. J. EKSPONENSIAL, 1(2), 26-33.

Fauziah, G. (2021). Estimasi Pseudo Poisson Maximum Likelihood Untuk Mengatasi Masalah dalam Model Log-Linear pada Kasus Kusta di Jawa Barat Tahun 2018. Jurnal Riset Statistika, 57-62.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley.

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.

uang, J., Lu, X., & Wang, Z. (2021). Spatial heterogeneity of COVID-19 transmission and its determinants using GWPR. Science of the Total Environment, 758, 144279.

Kementerian Kesehatan Republik Indonesia. (2021). Profil Kesehatan Indonesia Tahun 2020. Jakarta: Pusat Data dan Informasi Kemenkes RI.

Kusuma, W., Komalasari, D., & Hadijati, M. (2013). Model Regresi Zero Inflated Poisson Pada Data Overdispersion. Jurnal Matematika, 3(2), 71-85.

Lin, H. C., Wen, T. H., & Chen, Y. C. (2020). Spatial modeling of dengue fever risk with Geographically Weighted Poisson Regression. International Journal of Environmental Research and Public Health, 17(3), 1236.

Lönnroth, K., Jaramillo, E., Williams, B. G., Dye, C., & Raviglione, M. (2009). Drivers of tuberculosis epidemics: The role of risk factors and social determinants. Social Science & Medicine, 68(12), 2240–2246. https://doi.org/10.1016/j.socscimed.2009.03.041

Mei, R. W. (2018). Pemodelan Geographically Weighted Binary Logistic Regression (GWBLR) dengan menggunakan pembobot Adaptive Gaussian Kernel (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).

Millah, U. H. (2015). Estimasi Parameter Model Geographically Weighted Poisson Regression (GWPR) Pada Data Yang Mengandung Outlier (Doctoral Dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).

Nakaya, T., Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. Statistics in Medicine, 24(17), 2695–2717.

Nurmasari, R. (2016). Prediksi Jumlah Penderita Demam Berdarah Di Kabupaten Malang Berdasarkan Kondisi Cuaca Menggunakan Metode Regresi Dengan Transformasi Natural Logaritmik Pada Variabel Respon. Departemen Sistem Informasi, Surabaya.

Otaya, L. G. (2016). Distribusi Probabilitas Weibull dan Aplikasinya (Pada Persoalan Keandalan (Reliability) dan Analisis Rawatan (Mantainability). Tadbir: Jurnal Manajemen Pendidikan Islam, 4(2), 44-66.

Qomariyah, N., Purnami, S. W., & Purnomo, S. (2013). Pemodelan Faktor-Faktor Yang Mempengaruhi Jumlah Kematian Ibu Di Jatim Dengan Pendekatan GWPR (Geographically Weighted Poisson Regression). Jurnal Sains dan Seni ITS, 2(2), D311-D316.

Sari, D. P., Rahmawati, Y., & Nugroho, H. (2022). Analisis spasial kejadian tuberkulosis dengan GWPR di Jawa Tengah. Jurnal Epidemiologi Kesehatan Indonesia, 6(2), 45–55.

Saudah, A., Baharuddin, B., Abapihi, B., Ruslan, R., & Agusrawati, A. (2022). Model Jumlah Balita Pneumonia Kota Kendari dan Kabupaten Konawe Dengan Generalized Poisson Regresi: Model Jumlah Balita Pneumonia Kota Kendari dan Kabupaten Konawe. Jurnal Matematika Komputasi dan Statistika, 2(3), 202-207.

ISSN: 2776-7930 (Print) / 2807-3010 (Online)

Sejati, A. E., Sugiarto, A., Anasi, P. T., Utaya, S., & Bachri, S. (2022). Tantangan Filsafat Geografi Dalam Perkembangan Geografi Terkini: Kajian Ontologi, Epistemologi, Aksiologi, dan Etika. Majalah Geografi Indonesia, 36(2).

Septiani, I. N. (2021). Pemodelan Jumlah Kematian Bayi dengan Pendekatan Geographically Weighted Poisson Regression (GWPR) Studi Kasus di Provinsi Kalimantan Barat. Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya, 10(1).

Sharma, S. K., & Mohan, A. (2020). Tuberculosis: From an incurable scourge to a curable disease – journey over a millennium. The Indian Journal of Medical Research, 151(5), 431–441. https://doi.org/10.4103/ijmr.IJMR_436_20

Soraya, E., Wardhana, W., & Sadono, R. (2016). Pemodelan spasial resiliensi ekosistem Gunungapi Merapi pasca erupsi. Jurnal Ilmu Kehutanan, 10(2), 86-97.

Sogen, D. L., Arum, P. R., & Wasono, R. (2023, November). Pemodelan Geographically Weighted Poisson Regression (GWPR) Dengan Pembobot Fixed Bisquare Kernel Dan Adaptive Bisquare Kernel Pada Kasus DBD Di Riau. In Prosiding Seminar Nasional Unimus (Vol. 6).

Supriyanti, E. (2022). Bab 3 Penerapan Kesehatan Masyarakat di Indonesia. Kesehatan Masyarakat Dalam Kebidanan, 21.

Tizona, A. R., Goejantoro, R., & Wasono, W. (2017). Pemodelan Geographically Weighted Regression (Gwr) Dengan Fungsi Pembobot Adaptive Kernel Bisquare Untuk Angka Kesakitan Demam Berdarah di Kalimantan Timur Tahun 2015. Ekponensial , 8(1), 87-94.

Tuasikal, N. J. (2018). Implementasi Webgis, Geographically Weighted Poisson Regression (GWPR), dan Clustering K-Means Untuk Analisis Data Tuberkulosis Dunia (Studi Kasus: Kasus Tuberkulosis Dunia Tahun 2015).

Walpole, Ronald. E. 1995. Pengantar Statistika, Edisi ke-3. PT Gramedia Pustaka Utama, Jakarta.

Wang, S., Fang, C., Ma, H., Wang, Y., & Qin, J. (2014). Spatial Differences and Multi-Mechanism of Carbon Footprint Based on GWR Model in Provincial China. Journal of Geographical Sciences, 24, 612-630.

Widyaningsih, Y., & Fitrianingrum, M. R. (2022). Pemodelan Spasial pada Data Produk Domestik Regional Bruto di Pulau Jawa Sebelum dan Ketika Pandemi. Jurnal Statistika dan Aplikasinya, 6(1), 12-25.

Wijaya, M. S., Mantik, M. F., & Rampengan, N. H. (2021). Faktor Risiko Tuberkulosis pada Anak. e-CliniC, 9(1).