

Pemodelan Jumlah Kasus Tuberkulosis Paru di Indonesia dengan *Geographically Weighted Negative Binomial Regression*

Modeling The Number of Pulmonary Tuberculosis Cases in Indonesia using Geographically Weighted Negative Binomial Regression

Dhita Reginna Putri¹, M. Fathurahman^{2a)}, and Suyitno³

^{1,2,3} Program Studi Statistika, Jurusan Matematika FMIPA Universitas Mulawarman, Indonesia

^{a)}Corresponding author: fathur@fmipa.unmul.ac.id

ABSTRACT

Geographically Weighted Negative Binomial Regression (GWNBR) model is the development of Negative Binomial Regression (NBR) model applied to spatial data. The parameter estimation of GWNBR model is performed at each observation location using spatial weighting. The purpose of this study is to determine the GWNBR model of the number of pulmonary tuberculosis cases in Indonesia in 2021 and determine the factors that influence pulmonary tuberculosis cases in Indonesia in 2021. The research data are secondary data obtained from the Indonesian Ministry of Health and Indonesian Central Agency on Statistics with seven predictor variables. Parameter estimation method is Maximum Likelihood Estimation (MLE). Spatial weighting is calculated by using the Adaptive bi-square weighting function and the optimum bandwidth is determined by using the Cross-Validation (CV). The research results showed that the exact Maximum Likelihood (ML) estimator could not be obtained analytically and the approximation of ML estimator was obtained by using the Newton-Raphson iterative method. Based on the results of the parameter testing of GWNBR model, it was concluded that the factors affecting the number of tuberculosis cases were local and varied in 34 provinces. The factor affecting locally are population density, the percentage of districts/cities implementing GERMAS, and number of hospitals.

Keywords: *gwnbr, mle, pulmonary tuberculosis*

1. Introduction

Tuberkulosis merupakan penyakit menular yang disebabkan oleh bakteri TB yaitu *Myobacterium Tuberculosis*. Sebagian besar bakteri ini menyerang organ paru-paru, tetapi bakteri ini dapat juga menyerang orang tubuh lainnya (World Health Organization, 2020). Pada tahun 2015, 10,4 juta kasus tuberkulosis paru baru ditemukan di dunia dan 10 persen diantaranya telah meninggal. Oleh karena itu, tuberkulosis paru dianggap sebagai penyakit menular mematikan di dunia setelah HIV. Tuberkulosis paru banyak ditemukan di negara dengan populasi yang besar, salah satunya adalah Indonesia (World Health Organization, 2016).

Indonesia merupakan negara yang menempati peringkat kedua di dunia dengan penderita penyakit tuberkulosis paru tertinggi setelah India (World Health Organization, 2020). Karena penyakit tuberkulosis ini sangat bahaya, maka perlu dilakukan pencegahan dengan cara melakukan tindakan seperti mencari tahu faktor-faktor apa saja yang mempengaruhi tuberkulosis paru di Indonesia dengan menggunakan regresi Poisson. Regresi Poisson merupakan model regresi nonlinier yang sering digunakan untuk memodelkan data *count* (Agresti, 2007). Pada penerapannya, model regresi Poisson harus memenuhi asumsi equidispersi, yang berarti asumsi nilai varians dari variabel responnya sama dengan nilai rata-ratanya. Namun pada kenyataannya, data *count* sering dijumpai kondisi dimana nilai variansnya berbeda dengan nilai rata-ratanya. Perbedaan tersebut mengakibatkan data *count* mengalami kondisi overdispersi.

Overdispersi adalah kondisi dimana nilai variansnya lebih besar dari nilai rata-ratanya (Keswari, dk., 2014). Jika data *count* mengalami overdispersi, maka data tersebut tidak tepat lagi dianalisis dengan regresi Poisson. Terdapat banyak metode regresi yang dapat digunakan untuk mengatasi kasus overdispersi, salah satunya adalah *Negative Binomial Regression (NBR)*. NBR merupakan model yang digunakan sebagai alternatif dari model regresi Poisson yang mengalami overdispersi tanpa harus menghilangkan kondisi overdispersi dari data tersebut (Fatma, dkk., 2016), karena model ini menyertakan parameter regresi sekaligus parameter dispersi. Terdapat salah satu metode lagi yang dapat mengatasi masalah overdispersi namun tetap memperhatikan aspek spasial. Aspek spasial perlu diperhatikan karena setiap lokasi memiliki kondisi geografis yang berbeda-beda. Sehingga metode yang tepat untuk mengatasi itu semua adalah model *Geographically Weighted Negative Binomial Regression (GWNBR)*.

Model GWNBR merupakan pengembangan dari model NBR dalam menduga data yang memiliki aspek spasial untuk data *count* yang mengalami overdispersi. Model GWNBR akan menghasilkan parameter lokal dengan masing-masing lokasi memiliki parameter yang berbeda-beda dan untuk mendapatkan model GWNBR diperlukan pembobotan spasial berdasarkan jarak satu lokasi pengamatan dengan lokasi pengamatan yang lainnya (Fotheringham, dkk., 2002). Pembobot spasial yang paling sering digunakan untuk memodelkan

GWNBR adalah *adaptive bandwidth* dengan fungsi pembobot *bi-square*.

Berdasarkan uraian di atas, kasus tuberkulosis paru masih menjadi penyebab kematian yang tinggi di Indonesia dan metode yang tepat digunakan untuk mengetahui faktor-faktor yang menyebabkan kasus tuberkulosis paru adalah GWNBR. Oleh karena itu, penelitian ini akan menerapkan model GWNBR untuk mengetahui faktor-faktor yang mempengaruhi kasus tuberkulosis paru di Indonesia tahun 2021.

2. Tinjauan Pustaka

2.1 Regresi Poisson

Model regresi Poisson merupakan penerapan dari *Generalized Linear Models* (GLM) dan dapat digunakan untuk memodelkan data respon berdistribusi Poisson yang merupakan keluarga eksponensial (Simarmata & Ispriyanti, 2011). Regresi Poisson adalah model regresi nonlinier yang sering digunakan untuk memodelkan data *count*. Jika variabel acak (Y) merupakan distribusi Poisson dengan parameter μ , maka fungsi kepadatan peluang dari distribusi Poisson dinyatakan sebagai berikut (Suryani, dkk., 2021):

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

dimana

μ : Rata-rata banyaknya kejadian sukses dalam satuan unit tertentu yaitu $E(Y) = \mu$ dan $Var(Y) = \mu$.
Secara matematis, model regresi Poisson diformulasikan sebagai berikut (Hilbe, 2014):

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}^*), i = 1, 2, \dots, n \tag{2}$$

dimana:

\mathbf{x}_i^T : Vektor variabel prediktor untuk pengamatan ke- i , dengan $\mathbf{x}_i^T = [1 \ X_{1i} \ X_{2i} \ \dots \ X_{pi}]$
 $\boldsymbol{\beta}^*$: Vektor parameter, dengan $\boldsymbol{\beta}^* = [\beta^* \ \beta_1^* \ \beta_2^* \ \dots \ \beta_p^*]^T$.

Model regresi Poisson pada persamaan (2) didapatkan dengan cara melakukan penaksiran parameter modelnya menggunakan metode *Maximum Likelihood Estimation* (MLE). Penaksiran parameter model regresi Poisson diawali dengan membentuk fungsi *likelihood* dan *log-likelihood*. Berdasarkan persamaan (2), maka bentuk fungsi *likelihood* dan *ln-likelihood* sebagai berikut:

$$L(\boldsymbol{\beta}^*) = \prod_{i=1}^n \frac{\exp[y_i \mathbf{x}_i^T \boldsymbol{\beta}^* - \exp(\mathbf{x}_i^T \boldsymbol{\beta}^*)]}{y_i!} \tag{3}$$

$$\ell(\boldsymbol{\beta}^*) = \sum_{i=1}^n (y_i \mathbf{x}_i^T \boldsymbol{\beta}^* - \exp(\mathbf{x}_i^T \boldsymbol{\beta}^*)) - \ln y_i! \tag{4}$$

dimana:

$L(\boldsymbol{\beta}^*)$: Fungsi *likelihood*
 $\ell(\boldsymbol{\beta}^*)$: Fungsi *log-likelihood*.

Selanjutnya adalah memaksimumkan fungsi *log-likelihood* pada persamaan (4) dengan cara menentukan turunan parsial pertama fungsi *log-likelihood* terhadap parameter $\boldsymbol{\beta}^*$ yang kemudian disamakan dengan nol, sebagai berikut:

$$\frac{\partial \ell(\boldsymbol{\beta}^*)}{\partial \beta_k^*} = \sum_{i=1}^n x_{ik} (y_i - \exp[\mathbf{x}_i^T \boldsymbol{\beta}^*]) = 0. \tag{5}$$

Berdasarkan persamaan (5) diketahui bahwa bentuk fungsi *log-likelihood* merupakan sistem persamaan non linier, sehingga solusi eksak untuk mendapatkan penaksir ML tidak dapat ditemukan secara analitik. Metode alternatif untuk menyelesaikan ini adalah dengan menggunakan metode iterasi Newton-Raphson. Setelah didapatkan penaksir parameter model regresi Poisson, maka hal selanjutnya yang dilakukan adalah melakukan pengujian signifikansi parameter yang bertujuan untuk mengetahui adanya pengaruh dari variabel prediktor terhadap variabel respon. Uji signifikansi parameter model regresi Poisson terdiri dari uji serentak dan uji parsial (Hilbe, 2014).

Pengujian serentak dilakukan untuk mengetahui adanya pengaruh variabel prediktor terhadap variabel respon secara serentak. Hipotesis yang digunakan untuk pengujian serentak adalah

$$H_0 : \beta_1^* = \beta_2^* = \dots = \beta_p^* = 0 \tag{6}$$

$$H_1 : \text{Minimal terdapat satu } \beta_k^* \neq 0, k = 1, 2, \dots, p.$$

Statistik uji yang digunakan untuk persamaan (6) adalah metode *Maximum Likelihood Ratio Test* (MLRT) yang didefinisikan sebagai berikut:

$$G_1 = 2 \left(L(\widehat{\Omega}_1) - L(\widehat{\omega}_1) \right) \tag{7}$$

dimana:

$L(\widehat{\Omega}_1)$: Nilai maksimum *log-likelihood* untuk model regresi Poisson dengan melibatkan variabel prediktor
 $L(\widehat{\omega}_1)$: Nilai maksimum *log-likelihood* untuk model regresi Poisson tanpa melibatkan variabel prediktor.

Statistik uji MLRT pada persamaan (7) berdistribusi *chi-square* dengan derajat bebasnya adalah p . Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $G_1 \geq \chi^2_{(\alpha;p)}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$, dengan $p_{value} = P(G > G_1)$, dimana G adalah variabel acak berdistribusi χ^2_p dan G_1 adalah nilai statistik uji G , serta p merupakan banyaknya variabel prediktor (Karima, 2021).

Selanjutnya adalah pengujian parameter model regresi Poisson secara parsial yang bertujuan untuk mengetahui adanya pengaruh dari masing-masing variabel prediktor terhadap variabel respon secara individu. Hipotesis yang digunakan untuk pengujian parsial adalah

$$\begin{aligned} H_0 : \beta_k^* &= 0 \\ H_1 : \beta_k^* &\neq 0, k = 1, 2, \dots, p. \end{aligned} \tag{8}$$

Statistik uji yang digunakan untuk persamaan (8) adalah uji Wald yang didefinisikan sebagai berikut

$$Z_1 = \frac{\hat{\beta}_k^*}{\sqrt{Var(\hat{\beta}_k^*)}} \tag{9}$$

dimana:

$$\begin{aligned} \hat{\beta}_k^* &: \text{Penaksir parameter model regresi Poisson} \\ \sqrt{Var(\hat{\beta}_k^*)} &: \text{Standard Error untuk penaksir parameter model regresi Poisson.} \end{aligned}$$

Statistik uji Wald pada persamaan (9) berdistribusi normal baku. Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $|Z_1| > Z_{\alpha/2}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$, dengan $p_{value} = 2(1 - P(Z \leq |Z_1|))$, dimana Z adalah variabel acak berdistribusi normal baku dan Z_1 adalah nilai statistik uji Z (Karima, 2021).

2.2 Overdispersi Regresi Poisson

Overdispersi merupakan kondisi dimana nilai varians lebih besar dari nilai rata-rata, hal tersebut dapat dijelaskan sebagai berikut (Suryani, dkk., 2021):

$$Var(Y) > E(Y). \tag{10}$$

Overdispersi juga dapat diperiksa dengan menggunakan statistic uji yang ditemukan pada *package* AER dari *software* R, karena uji ini mengimplementasikan uji yang dikembangkan oleh Cameron dan Trivedi.

2.3 Negative Binomial Regression (NBR)

Model NBR merupakan model yang berasal dari campuran Poisson-Gamma serta penerapan dari GLM yang menggambarkan hubungan antara variabel respon dengan variabel prediktor. NBR digunakan sebagai alternative dari model regresi Poisson yang mengalami overdispersi (Hilbe, 2011). Misalkan diberikan sampel acak variabel respon berdistribusi NB dengan parameter disperse η , maka bentuk persamaan fungsi kepadatan peluang NB dinyatakan sebagai berikut:

$$f(y_i, \mu, \eta) = \frac{\Gamma(y_i + \frac{1}{\eta})}{y_i! \Gamma(\frac{1}{\eta})} \left(\frac{1}{1 + \eta\mu}\right)^{\frac{1}{\eta}} \left(\frac{\eta\mu}{1 + \eta\mu}\right)^{y_i}; y_i = 0, 1, 2, \dots, \tag{11}$$

dimana:

$$\begin{aligned} \mu &: \text{Parameter lokasi dengan } \mu > 0 \\ \eta &: \text{Parameter disperse dengan } \eta > 0 \end{aligned}$$

Secara matematis, model NBR diformulasikan sebagai berikut (Hilbe, 2014):

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}), i = 1, 2, \dots, n \tag{12}$$

dimana:

$$\begin{aligned} \mu_i &: \text{Rata-rata variabel respon untuk pengamatan ke-}i \\ \mathbf{x}_i^T &: \text{Vektor variabel prediktor untuk pengamatan ke-}i, \text{ dengan } \mathbf{x}_i^T = [1 \ X_{1i} \ X_{2i} \ \dots \ X_{pi}] \\ \boldsymbol{\beta} &: \text{Vektor parameter, dengan } \boldsymbol{\beta} = [\beta \ \beta_1 \ \beta_2 \ \dots \ \beta_p]^T. \end{aligned}$$

Model NBR pada persamaan (12) didapatkan dengan cara melakukan penaksiran parameter model NBR menggunakan metode *Maximum Likelihood Estimation* (MLE). Penaksiran parameter model NBR diawali dengan membentuk fungsi *likelihood* dan *log-likelihood*. Berdasarkan persamaan (12), maka bentuk fungsi *likelihood* dan *ln-likelihood* sebagai berikut:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \left[\frac{\Gamma(y_i + \frac{1}{\eta})}{\Gamma(\frac{1}{\eta})} \left(\frac{1}{y_i!}\right) \left(\frac{1}{1 + \eta \exp(\mathbf{x}_i^T \boldsymbol{\beta})}\right)^{\frac{1}{\eta}} \left(\frac{\eta \exp(\mathbf{x}_i^T \boldsymbol{\beta})}{1 + \eta \exp(\mathbf{x}_i^T \boldsymbol{\beta})}\right)^{y_i} \right] \tag{13}$$

$$\ell(\theta) = \sum_{i=1}^n \left[\ln \left(y_i + \frac{1}{\eta} \right) - \ln \left(\frac{1}{\eta} \right) - \ln(y_i!) + y_i \ln(\eta + \mathbf{x}_i^T \boldsymbol{\beta}) - \left(\frac{1}{\eta} + y_i \right) \ln \left(1 + \eta \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \right) \right] \quad (14)$$

dimana:

$L(\boldsymbol{\beta}^*)$: Fungsi *likelihood*

$\ell(\boldsymbol{\beta}^*)$: Fungsi *log-likelihood*

dengan $\theta = [\eta \ \boldsymbol{\beta}^T]^T$ dan $\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_p]^T$

Selanjutnya adalah memaksimumkan fungsi *log-likelihood* pada persamaan (14) dengan cara menentukan turunan parsial pertama fungsi *log-likelihood* terhadap parameter θ yang kemudian disama dengankan nol. Namun, fungsi *log-likelihood* NBR yang terbentuk merupakan sistem persamaan non linier, sehingga solusi eksak untuk mendapatkan penaksir ML tidak dapat ditemukan secara analitik. Metode alternatif untuk menyelesaikan ini adalah dengan menggunakan metode iterasi Newton-Raphson. Setelah didapatkan penaksir parameter model NBR, maka hal selanjutnya yang dilakukan adalah melakukan pengujian signifikansi parameter model NBR terdiri dari uji serentak dan uji parsial (Hilbe, 2014).

Pengujian serentak dilakukan untuk mengetahui adanya pengaruh variabel prediktor terhadap variabel respon secara serentak. Hipotesis yang digunakan untuk pengujian serentak adalah

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 \quad (15)$$

$$H_1 : \text{Minimal terdapat satu } \beta_k \neq 0, k = 1, 2, \dots, p.$$

Statistik uji yang digunakan untuk persamaan (15) adalah metode *Maximum Likelihood Ratio Test* (MLRT) yang didefinisikan sebagai berikut:

$$G_2 = 2 \left(L(\hat{\Omega}_2) - L(\hat{\omega}_2) \right) \quad (16)$$

dimana:

$L(\hat{\Omega}_2)$: Nilai maksimum *log-likelihood* untuk model NBR dengan melibatkan variabel prediktor

$L(\hat{\omega}_2)$: Nilai maksimum *log-likelihood* untuk model NBR tanpa melibatkan variabel prediktor.

Statistik uji MLRT pada persamaan (16) berdistriburi *chi-square* dengan derajat bebasnya adalah p . Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $G_2 \geq \chi^2_{(\alpha;p)}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$, dengan $p_{value} = P(G > G_2)$, dimana G adalah variabel acak berdistriburi χ^2_p dan G_2 adalah nilai statistik uji G , serta p merupakan banyaknya variabel prediktor.

Selanjutnya adalah pengujian parameter model NBR secara parsial yang bertujuan untuk mengetahui adanya pengaruh dari masing-masing variabel prediktor terhadap variabel respon secara individu. Hipotesis yang digunakan untuk pengujian parsial adalah

$$H_0 : \beta_k = 0 \quad (17)$$

$$H_1 : \beta_k \neq 0, k = 1, 2, \dots, p.$$

Statistik uji yang digunakan untuk persamaan (17) adalah uji Wald yang didefinisikan sebagai berikut

$$Z_2 = \frac{\hat{\beta}_k}{\sqrt{\text{Var}(\hat{\beta}_k)}} \quad (18)$$

dimana:

$\hat{\beta}_k$: Penaksir parameter model NBR

$\sqrt{\text{Var}(\hat{\beta}_k)}$: *Standard Error* untuk penaksir parameter model NBR.

Statistik uji Wald pada persamaan (18) berdistriburi normal baku. Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $|Z_2| > Z_{\alpha/2}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$, dengan $p_{value} = 2(1 - P(Z \leq |Z_2|))$, dimana Z adalah variabel acak berdistriburi normal baku dan Z_2 adalah nilai statistik uji Z .

2.4 Heterogenitas Spasial

Heterogenitas spasial adalah fenomena dimana hubungan antarvariabel atau unsur galat bervariasi antarlokasi. Pengujian heterogenitas spasial dilakukan untuk melihat apakah terdapat kekhasan pada lokasi pengamatan, sehingga parameter regresi yang dihasilkan berbeda-beda secara spasial. Hipotesis yang digunakan untuk pengujian heterogenitas spasial adalah:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \quad (19)$$

$$H_1 : \text{Minimal terdapat satu } \sigma_i^2 \neq \sigma^2.$$

Statistik uji yang digunakan untuk persamaan (19) adalah uji Breusch-Pagan yang didefinisikan sebagai berikut

$$BP = \left(\frac{1}{2} \right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (20)$$

dimana:

Z : Matriks berukuran $n(p + 1)$ berisi vektor yang telah dinormal standarkan untuk setiap pengamatan dengan f adalah $f_i = \frac{\sigma_i^2}{\sigma^2}$ dimana $e_i = y_i - \hat{y}_i$.

Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $BP > \chi^2_{(\alpha;p)}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$.

2.5 Pembobot Spasial

Pembobot spasial merupakan pembobot yang menjelaskan letak data yang satu dengan yang lainnya (Simamora & Ratnasari, 2014). Terdapat beberapa literatur yang digunakan untuk menentukan besarnya pembobot pada masing-masing lokasi yang berbeda, diantaranya dengan menggunakan fungsi *adaptive kernel* (Yasin, 2011). Fungsi *adaptive kernel* yaitu fungsi kernel yang memiliki nilai *bandwidth* yang berbeda-beda pada masing masing lokasi pengamatan. Pembobot yang sering digunakan pada fungsi *adaptive kernel* adalah fungsi *adaptive bi-square* yang diformulasikan sebagai berikut

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b}\right)^2\right)^2, & \text{untuk } d_{ij} \leq b \\ 0, & \text{untuk } d_{ij} > b \end{cases} \quad (21)$$

dengan

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (22)$$

dimana:

d_{ij} : Jarak *Euclidean* antara lokasi ke- i dan lokasi ke- j

u_i : Koordinat *latitude* pada titik pengamatan ke- i

v_i : Koordinat *longitude* pada titik pengamatan ke- i

b : Nilai parameter penghalus fungsi yang nilainya selalu positif atau yang disebut sebagai *bandwidth*

Menurut Fotheringham dkk (2002) salah satu metode untuk mendapatkan *bandwidth* optimum adalah menggunakan pendekatan *Cross Validation* (CV) yang diformulasikan sebagai berikut (Widayaka, dkk., 2016)

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2 \quad (23)$$

dengan $\hat{y}_{\neq i}(b)$ adalah nilai penaksir dari y_i dimana titik pengamatan lokasi (u_i, v_i) tidak dilibatkan pada proses penaksiran.

2.6 Model Geographically Weighted Negative Binomial Regression (GWNBR)

Model GWNBR merupakan pengembangan dari model NBR dan salah satu metode yang cukup efektif untuk menduga data yang memiliki heterogenitas spasial pada data *count* yang mengalami overdispersi. Berdasarkan model NBR pada persamaan (12) dapat dikembangkan menjadi model GWNBR. Misalkan koordinat semua lokasi pengamatan diketahui, maka model GWNBR pada lokasi ke- i dengan koordinat (u_i, v_i) dinyatakan sebagai berikut (Ricardo & Carvalho, 2013):

$$\mu(u_i, v_i) = \exp(x_i^T \beta(u_i, v_i)), i = 1, 2, \dots, n \quad (24)$$

dimana:

x_i^T : Vektor variabel prediktor untuk pengamatan ke- i , dengan $x_i^T = [1 \ X_{1i} \ X_{2i} \ \dots \ X_{pi}]$

$\beta(u_i, v_i)$: Vektor parameter, dengan $\beta = [\beta_0(u_i, v_i) \ \beta_1(u_i, v_i) \ \beta_2(u_i, v_i) \ \dots \ \beta_p(u_i, v_i)]^T$.

Model NBR pada persamaan (24) didapatkan dengan cara melakukan penaksiran parameter model GWNBR menggunakan metode *Maximum Likelihood Estimation* (MLE). Penaksiran parameter model GWNBR diawali dengan membentuk fungsi *likelihood* dan *log-likelihood*. Berdasarkan persamaan (24), maka bentuk fungsi *likelihood* dan *ln-likelihood* sebagai berikut:

$$L(\theta(u_i, v_i)) = \prod_{j=1}^n w_{ij} \left[\frac{\Gamma\left(y_j + \frac{1}{\eta(u_i, v_i)}\right)}{\Gamma\left(\frac{1}{\eta(u_i, v_i)}\right)} \left(\frac{1}{y_j!}\right) \left(\frac{1}{1 + \eta(u_i, v_i) \exp(x_i^T \beta(u_i, v_i))}\right)^{\frac{1}{\eta(u_i, v_i)}} \Delta_1 \right] \quad (25)$$

$$\ell(\theta(u_i, v_i)) = \sum_{j=1}^n w_{ij} \left[\ln\left(y_j + \frac{1}{\eta(u_i, v_i)}\right) - \ln\left(\frac{1}{\eta(u_i, v_i)}\right) - \Delta_2 - \Delta_3 \right] \quad (26)$$

dengan

$\theta(u_i, v_i) = [\eta(u_i, v_i) \ \beta^T(u_i, v_i)]^T$

$\beta(u_i, v_i) = [\beta_1(u_i, v_i) \ \beta_2(u_i, v_i) \ \dots \ \beta_p(u_i, v_i)]^T$

$$\Delta_1 = \left(\frac{\eta(u_i, v_i) \exp(\mathbf{x}_j^T \boldsymbol{\beta}(u_i, v_i))}{1 + \eta(u_i, v_i) \exp(\mathbf{x}_j^T \boldsymbol{\beta}(u_i, v_i))} \right)^{y_j}$$

$$\Delta_2 = \ln(y_j!) + y_j \ln(\eta(u_i, v_i) + \mathbf{x}_j^T \boldsymbol{\beta}(u_i, v_i))$$

$$\Delta_3 = \left(\frac{1}{\eta(u_i, v_i)} + y_j \right) \ln(1 + \eta(u_i, v_i) \exp(\mathbf{x}_j^T \boldsymbol{\beta}(u_i, v_i)))$$

Selanjutnya adalah memaksimalkan fungsi *log-likelihood* pada persamaan (26) dengan cara menentukan turunan parsial pertama fungsi *log-likelihood* terhadap parameter θ yang kemudian disama dengankan nol. Namun, fungsi *log-likelihood* model GWNBR yang terbentuk merupakan sistem persamaan non linier, sehingga solusi eksak untuk mendapatkan penaksir ML tidak dapat ditemukan secara analitik. Metode alternatif untuk menyelesaikan ini adalah dengan menggunakan metode iterasi Newton-Raphson. Setelah didapatkan penaksir parameter model GWNBR, maka hal selanjutnya yang dilakukan adalah melakukan uji kesamaan model NBR dengan model GWNBR dan pengujian signifikansi parameter model GWNBR terdiri dari uji serentak dan uji parsial.

Uji kesamaan model NBR dengan model GWNBR digunakan untuk mengetahui adanya pengaruh faktor lokasi terhadap model GWNBR. Hipotesis yang digunakan untuk uji kesamaan model NBR dengan GWNBR adalah

$$H_0 : \beta_k(u_i, v_i) = \beta_k, i = 1, 2, \dots, n$$

$$H_1 : \beta_k(u_i, v_i) \neq \beta_k, k = 1, 2, \dots, p. \tag{27}$$

Statistik uji yang digunakan untuk persamaan (27) adalah metode devians yang didefinisikan sebagai berikut:

$$F^* = \frac{G_2/dfA}{G_3/dfB} \tag{28}$$

dimana:

- G_2 : Devians model NBR yang diberikan oleh persamaan (16)
- G_3 : Devians model GWNBR yang diformulasikan sebagai berikut

$$G_3 = 2 \left(L(\hat{\Omega}_3) - L(\hat{\omega}_3) \right) \tag{29}$$

Pengujian serentak dilakukan untuk mengetahui adanya pengaruh variabel prediktor terhadap variabel respon secara serentak. Hipotesis yang digunakan untuk pengujian serentak adalah

$$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_p(u_i, v_i) = 0, i = 1, 2, \dots, n$$

$$H_1 : \text{Minimal terdapat satu } \beta_k(u_i, v_i) \neq 0, k = 1, 2, \dots, p. \tag{30}$$

Statistik uji yang digunakan untuk persamaan (30) adalah metode *Maximum Likelihood Ratio Test* (MLRT) yang diberikan oleh persamaan (29). Statistik uji MLRT pada persamaan (29) berdistribusi *chi-square* dengan derajat bebasnya adalah p . Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $G_3 \geq \chi^2_{(\alpha; p)}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$, dengan $p_{value} = P(G > G_3)$, dimana G adalah variabel acak berdistribusi χ^2_p dan G_3 adalah nilai statistik uji G , serta p merupakan banyaknya variabel prediktor.

Selanjutnya adalah pengujian parameter model GWNBR secara parsial yang bertujuan untuk mengetahui adanya pengaruh dari masing-masing variabel prediktor terhadap variabel respon secara individu. Hipotesis yang digunakan untuk pengujian parsial adalah

$$H_0 : \beta_k(u_i, v_i) = 0$$

$$H_1 : \beta_k(u_i, v_i) \neq 0, k = 1, 2, \dots, p. \tag{31}$$

Statistik uji yang digunakan untuk persamaan (31) adalah uji Wald yang didefinisikan sebagai berikut

$$Z_3 = \frac{\hat{\beta}_k(u_i, v_i)}{\sqrt{\text{Var}(\hat{\beta}_k(u_i, v_i))}} \tag{32}$$

dimana:

- $\hat{\beta}_k(u_i, v_i)$: Penaksir parameter model GWNBR
- $\sqrt{\text{Var}(\hat{\beta}_k(u_i, v_i))}$: *Standard Error* untuk penaksir parameter model GWNBR.

Statistik uji Wald pada persamaan (32) berdistribusi normal baku. Kriteria pengujian adalah H_0 ditolak pada taraf signifikansi α jika nilai $|Z_3| > Z_{\alpha/2}$ atau dapat dinyatakan dalam bentuk lain yaitu H_0 ditolak jika $p_{value} < \alpha$, dengan $p_{value} = 2(1 - P(Z \leq |Z_3|))$, dimana Z adalah variabel acak berdistribusi normal baku dan Z_3 adalah nilai statistik uji Z .

2.7 Ukuran Kebaikan Model

Ukuran kebaikan model dapat menggunakan metode *Akaike Information Criterion* (AIC). AIC adalah

kriteria kesesuaian model dalam menduga secara statistik dan dilakukan dengan cara membandingkan semua kemungkinan model yang dapat terbentuk dari variabel prediktor yang membentuk model. Adapun perhitungan nilai AIC diformulasikan sebagai berikut:

$$AIC = -2\ell(\hat{\theta}^*) + 2K \quad (33)$$

dimana:

K : Banyaknya parameter

$\ell(\hat{\theta}^*)$: Nilai fungsi *log-likelihood*.

Berdasarkan persamaan (33), model terbaik adalah model yang memiliki nilai AIC terkecil diantara model yang lain (Sembiring, 2019).

2.8 Pendeteksian Multikolinieritas

Multikolinieritas adalah suatu kondisi dimana terdapat hubungan linier atau korelasi yang tinggi antara satu variabel prediktor dengan variabel prediktor yang lain (Setiawan, 2015). Jika terdapat multikolinieritas pada suatu model regresi, maka menyebabkan taksiran parameter model regresi yang dihasilkan akan mempunyai *error* yang sangat besar. Salah satu pendeteksian kasus multikolinieritas yaitu dengan menggunakan kriteria nilai *Variance Inflation Factor* (VIF). Nilai VIF dinyatakan sebagai berikut:

$$VIF_k = \frac{1}{1 - R_k^2}, k = 1, 2, \dots, p \quad (34)$$

dengan:

$$R_k^2 = \frac{\sum_{i=1}^n (X_{ki} - \hat{X}_{ki})^2}{\sum_{i=1}^n (X_{ki} - \bar{X}_{ki})^2} \quad (35)$$

dimana:

R_k^2 : Koefisien determinasi model regresi untuk prediktor ke- k

X_{ki} : Nilai pengamatan ke- i untuk prediktor ke- k

\hat{X}_{ki} : Nilai taksiran dari X_{ki}

2.9 Tuberkulosis Paru

Tuberkulosis paru adalah penyakit infeksius, yang terutama menyerang penyakit parenkim paru. Tuberkel merupakan tonjolan kecil dan keras yang terbentuk waktu sistem kekebalan membangun tembok mengelilingi bakteri dalam paru. TB paru ini bersifat menahun dan secara khas ditandai oleh pembentukan granuloma dan menimbulkan nekrosis jaringan. TB paru dapat menular melalui udara, waktu seseorang dengan TB aktif pada paru batuk, bersin, atau bicara (Werdhani, 2009).

3. Metode Penelitian

3.1. Sumber Data dan Variabel Penelitian

Data penelitian ini adalah data sekunder yang diperoleh dari Kementerian Kesehatan Indonesia dan Badan Pusat Statistik Indonesia. Sedangkan variabel penelitian terdiri dari variabel respon, variabel prediktor, dan koordinat lokasi pengamatan. Variabel respon yaitu jumlah kasus tuberkulosis paru di 34 provinsi di Indonesia tahun 2021 (Y). Data variabel prediktor terdiri dari kepadatan penduduk (X_1), persentase rumah layak huni (X_2), persentase penduduk yang merokok (X_3), persentase Tempat Fasilitas Umum (TFU) yang memenuhi syarat standar (X_4), persentase kabupaten/kota yang melaksanakan Gerakan Masyarakat Hidup Sehat (GERMAS) (X_5), jumlah rumah sakit (X_6), dan persentase rumah tangga yang memiliki akses sanitasi layak (X_7). Sedangkan data koordinat lokasi pengamatan adalah pasangan letak garis lintang dan bujur dari 34 provinsi di Indonesia.

3.2. Populasi, Sampel, dan Teknik Sampling

Populasi pada penelitian ini adalah jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia. Sampel yang digunakan adalah jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021. Sedangkan teknik *sampling* menggunakan *purposive sampling*, yaitu teknik penentuan sampel dengan pertimbangan. Pertimbangan peneliti dalam menentukan sampel adalah kasus tuberkulosis paru masih menjadi penyebab kematian tertinggi di Indonesia.

3.3. Teknik Analisis Data

Teknik analisis data dalam penelitian ini adalah sebagai berikut:

1. Mendeskripsikan gambaran umum dari data jumlah kasus tuberkulosis paru beserta variabel-variabel yang diduga mempengaruhinya.
2. Mendeteksi adanya multikolinieritas antar variabel prediktor dengan kriteria VIF.
3. Pemodelan regresi Poisson dengan langkah sebagai berikut:
 - a. Melakukan penaksiran parameter model regresi Poisson menggunakan metode *Maximum Likelihood Estimation* (MLE).

- b. Melakukan pengujian signifikansi parameter model regresi Poisson secara serentak menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) dan secara parsial dengan metode Wald.
4. Melakukan pengujian overdispersi.
5. Pemodelan *Negative Binomial Regression* (NBR) dengan langkah sebagai berikut:
 - a. Melakukan penaksiran parameter model NBR dengan metode MLE.
 - b. Melakukan pengujian signifikansi parameter model NBR secara serentak menggunakan metode MLRT dan secara parsial menggunakan metode Wald.
6. Pengujian heterogenitas spasial dengan menggunakan uji Breusch-Pagan.
7. Pemodelan *Geographically Weighted Negative Binomial Regression* (GWNBR) dengan langkah sebagai berikut:
 - a. Menghitung jarak *Euclidean* antar titik lokasi pengamatan berdasarkan posisi geografis (*latitude* dan *longitude*).
 - b. Menentukan nilai *bandwidth* optimum berdasarkan kriteria CV minimum.
 - c. Menghitung nilai pembobot fungsi *adaptive kernel bi-square*.
 - d. Melakukan penaksiran parameter model GWNBR menggunakan metode MLE.
 - e. Melakukan pengujian kesamaan model NBR dengan model GWNBR.
 - f. Melakukan pengujian signifikansi parameter model GWNBR secara serentak menggunakan metode MLRT dan secara parsial menggunakan metode Wald.
 - g. Pemilihan model terbaik dengan membandingkan model NBR dengan model GWNBR menggunakan metode AIC.
 - h. Melakukan interpretasi terhadap model GWNBR.
8. Menarik kesimpulan.

4. Hasil dan Pembahasan

4.1. Deskripsi Data Penelitian

Deskripsi data penelitian dilakukan untuk mengetahui gambaran umum dari jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021. Deskripsi data penelitian meliputi nilai maksimum, nilai minimum, rata-rata, dan standar deviasi dari variabel respon dan variabel prediktor yang digunakan. Hasil deskripsi data penelitian dapat dilihat pada Tabel 1.

Tabel 1. Deskripsi Data Penelitian

Variabel	Maksimum	Minimum	Rata-rata	Standar Deviasi
Y	35.245	678	5.902	7.584
X ₁	17.031	10	773	2.899,2670
X ₂	85,15	27,6	60,14	12,1622
X ₃	34,07	19,58	28	3,0637
X ₄	84,3	17	53,66	17,1634
X ₅	100	0	49,99	31,0971
X ₆	400	12	89,47	100
X ₇	97,12	40,81	80,97	9,9257

Berdasarkan Tabel 4, diketahui bahwa rata-rata kasus tuberkulosis paru di Indonesia tahun 2021 adalah sebanyak 5.902 kasus, dengan standar deviasi sebanyak 7.584 kasus. Jumlah kasus tuberkulosis paru tertinggi di Provinsi Jawa Barat, yaitu sebanyak 35.245 kasus dan terendah di Provinsi Kalimantan Utara, yaitu sebanyak 678 kasus.

Sementara itu, rata-rata kepadatan penduduk di Indonesia sebanyak 773 jiwa/km² dengan standar deviasi sebanyak 2.899,267 jiwa/km². Kepadatan penduduk tertinggi di Provinsi DKI Jakarta, yaitu sebanyak 17.031 jiwa/km² dan terendah di Provinsi Kalimantan Utara, yaitu sebanyak 10 jiwa/km². Rata-rata persentase rumah layak huni sebanyak 60,14% dengan standar deviasi sebanyak 12,1622%. Persentase rumah layak huni tertinggi di Provinsi DI Yogyakarta, yaitu sebanyak 85,15% dan terendah di Provinsi Kepulauan Bangka Belitung, yaitu sebanyak 27,6%. Rata-rata persentase penduduk merokok di Indonesia sebanyak 28% dengan standar deviasi sebanyak 3,0637%. Persentase penduduk merokok tertinggi di Provinsi Lampung, yaitu sebanyak 34,07% dan terendah di Provinsi Bali, yaitu sebanyak 19,58%. Rata-rata persentase TFU yang memenuhi syarat standar di Indonesia sebanyak 53,66% dengan standar deviasi sebanyak 31,0971%. Persentase TFU yang memenuhi syarat standar tertinggi di Provinsi Bengkulu, yaitu sebanyak 84,3% dan terendah di Provinsi DI Yogyakarta, yaitu sebanyak 17%. Rata-rata persentase kabupaten/kota yang melaksanakan GERMAS di Indonesia sebanyak 49,99% dengan standar deviasi 31,0971%. Persentase kabupaten/kota yang melaksanakan GERMAS tertinggi di Provinsi Bengkulu, yaitu sebanyak 100% dan terendah di Provinsi Papua, yaitu sebanyak 0%. Rata-rata jumlah rumah sakit di Indonesia sebanyak 89,47 unit

dengan standar deviasi 100,424 unit. Jumlah rumah sakit tertinggi di Provinsi Jawa Timur, yaitu sebanyak 400 unit dan terendah di Provinsi Kalimantan Utara, yaitu sebanyak 12 unit. Rata-rata persentase rumah tangga yang memiliki akses sanitasi layak di Indonesia sebanyak 80,97% dengan standar deviasi sebanyak 9,9257%. Persentase rumah tangga yang memiliki akses sanitasi layak tertinggi di Provinsi DI Yogyakarta, yaitu sebanyak 97,12% dan terendah di Provinsi Papua, yaitu sebanyak 40,81%.

4.2. Pemeriksaan Multikolinieritas

Pemeriksaan multikolinieritas bertujuan untuk mengetahui apakah terdapat hubungan linier antar variabel prediktor dalam suatu model regresi. Pemeriksaan multikolinieritas terhadap variabel prediktor menggunakan kriteria VIF. Hasil yang diperoleh dapat dilihat pada Tabel 2.

Tabel 2. Nilai VIF Variabel Prediktor

Variabel	VIF
X_1	1,6138
X_2	1,7519
X_3	1,3111
X_4	1,1731
X_5	1,3614
X_6	1,2036
X_7	1,8515

Berdasarkan Tabel 2, didapatkan nilai VIF setiap variabel prediktor kurang dari 10. Hasil ini dapat disimpulkan bahwa tidak terdapat multikolinieritas antar variabel prediktor. Sehingga, pemodelan regresi Poisson dan NBR dapat menggunakan semua variabel prediktor.

4.3. Pemodelan Jumlah Kasus Tuberkulosis Paru dengan Regresi Poisson

Pemodelan jumlah kasus tuberkulosis paru di Indonesia tahun 2021 dengan Regresi Poisson yang terdiri dari 2 tahapan, yaitu melakukan penaksiran parameter model regresi Poisson dan pengujian parameter secara serentak serta parsial. Hasil yang diperoleh dapat dilihat pada Tabel 3.

Tabel 3. Hasil Pengujian Parameter dan Uji Parsial Parameter Model Regresi Poisson

Parameter	Taksiran	Standard Error	$ Z_1 $	P_{value}
$\hat{\beta}_0^*$	4,894	0,0566	85,63	< 0,001*
$\hat{\beta}_1^*$	0,0001	< 0,001	49,52	< 0,001*
$\hat{\beta}_2^*$	0,0092	0,0004	24,79	< 0,001*
$\hat{\beta}_3^*$	0,0740	0,0011	66,66	< 0,001*
$\hat{\beta}_4^*$	0,0043	0,0002	23,79	< 0,001*
$\hat{\beta}_5^*$	-0,0072	0,0001	-66,57	< 0,001*
$\hat{\beta}_6^*$	0,0064	< 0,001	356,09	< 0,001*
$\hat{\beta}_7^*$	0,0040	0,0005	8,71	< 0,001*

*Signifikan pada $\alpha = 0,1$.

Berdasarkan Tabel 3, didapatkan model regresi Poisson sebagai berikut:

$$\hat{\mu} = \exp(4,894 + 0,0001X_1 + 0,0092X_2 + 0,0740X_3 + 0,0043X_4 - 0,0072X_5 + 0,0064X_6) + \exp(+0,0040X_7). \tag{36}$$

Tahapan selanjutnya yaitu melakukan pengujian parameter model regresi Poisson secara serentak dan secara parsial. Hasil pengujian parameter model regresi Poisson secara serentak diperoleh nilai statistik uji G_1 sebesar 1.344.419 dan p_{value} sebesar 0. Selain itu, berdasarkan tabel distribusi *chi-kuadrat* didapatkan nilai $\chi^2_{(0,1;7)}$ sebesar 12,017. Karena nilai $G_1 > \chi^2_{(0,1;7)}$ dan nilai $p_{value} < \alpha = 0,1$, maka diputuskan H_0 ditolak. Sehingga disimpulkan bahwa variabel kepadatan penduduk, persentase rumah layak huni, persentase penduduk merokok, persentase TFU yang memenuhi syarat standar, persentase kabupaten/kota yang melaksanakan GERMAS, jumlah rumah sakit, dan persentase rumah tangga yang memiliki akses sanitasi layak secara serentak mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021.

Berdasarkan hasil pengujian parsial pada Tabel 3, diketahui bahwa penaksir parameter $\hat{\beta}_1^*, \hat{\beta}_2^*, \hat{\beta}_3^*, \hat{\beta}_4^*, \hat{\beta}_5^*, \hat{\beta}_6^*, \hat{\beta}_7^*$ memiliki nilai statistik uji Wald (Z_1) lebih dari $Z_{\alpha/2}$ sebesar 1,65, maka diputuskan H_0 ditolak. Sehingga disimpulkan bahwa variabel kepadatan penduduk, persentase rumah layak huni, persentase penduduk merokok, persentase TFU yang memenuhi syarat standar, persentase kabupaten/kota yang

melaksanakan GERMAS, jumlah rumah sakit, dan persentase rumah tangga yang memiliki akses sanitasi layak secara signifikan mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021.

4.4. Pendeteksian Overdispersi pada Model Regresi Poisson

Setelah dilakukan pengujian parameter model regresi Poisson, dilakukan pendeteksian overdispersi model regresi poisson. Hasil yang diperoleh adalah nilai standar deviasi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021 sebanyak 7.584 lebih dari nilai rata-ratanya yaitu 5.902. Selain itu, nilai uji Z^* sebesar 4,9204 lebih dari $Z_{(1-0,1/2)}$ sebesar 2,94, maka diputuskan H_0 ditolak. Hal ini menunjukkan bahwa terdapat overdispersi pada model regresi Poisson, sehingga untuk menangani overdispersi ini digunakan model NBR.

4.5. Pemodelan Jumlah Kasus Tuberkulosis Paru dengan NBR

Setelah terdeteksi adanya overdispersi pada model regresi Poisson, maka tahapan selanjutnya dilakukan pemodelan jumlah kasus tuberkulosis paru di Indonesia tahun 2021 dengan NBR yang terdiri dari 2 tahapan, yaitu melakukan penaksiran parameter model NBR dan pengujian parameter secara serentak serta parsial. Hasil yang diperoleh dapat dilihat pada Tabel 4.

Tabel 4. Hasil Pengujian Parameter dan Uji Parsial Parameter Model NBR

Parameter	Taksiran	Standard Error	$ Z_2 $	P_{value}
$\hat{\beta}_0$	6,294	1,312	4,797	< 0,001*
$\hat{\beta}_1$	0,0001	< 0,001	1,672	0,0946
$\hat{\beta}_2$	0,0085	0,0082	1,038	0,2994
$\hat{\beta}_3$	0,0557	0,0280	1,987	0,0469*
$\hat{\beta}_4$	-0,0012	0,0047	-0,246	0,8057
$\hat{\beta}_5$	-0,0056	0,0028	-1,993	0,0463*
$\hat{\beta}_6$	0,0077	0,0008	9,433	< 0,001*
$\hat{\beta}_7$	-0,0063	0,0103	-0,612	0,5405

*Signifikan pada $\alpha = 0,1$.

Berdasarkan Tabel 4, didapatkan model NBR sebagai berikut:

$$\hat{\mu} = \exp(6,294 + 0,0001X_1 + 0,0085X_2 + 0,0557X_3 - 0,0012X_4 - 0,0056X_5 + 0,0077X_6) + \exp(-0,0063X_7). \tag{37}$$

Tahapan selanjutnya yaitu melakukan pengujian parameter model NBR secara serentak dan secara parsial. Hasil pengujian parameter model NBR secara serentak diperoleh nilai statistik uji G_2 sebesar 3.997.602 dan p_{value} sebesar 0. Selain itu, berdasarkan tabel distribusi *chi-kuadrat* didapatkan nilai $\chi^2_{(0,1;7)}$ sebesar 12,017. Karena nilai $G_2 > \chi^2_{(0,1;7)}$ dan nilai $p_{value} < \alpha = 0,1$, maka diputuskan H_0 ditolak. Sehingga disimpulkan bahwa variabel kepadatan penduduk, persentase rumah layak huni, persentase penduduk merokok, persentase TFU yang memenuhi syarat standar, persentase kabupaten/kota yang melaksanakan GERMAS, jumlah rumah sakit, dan persentase rumah tangga yang memiliki akses sanitasi layak secara serentak mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021.

Berdasarkan hasil pengujian parsial pada Tabel 4, diketahui bahwa penaksir parameter $\hat{\beta}_3, \hat{\beta}_5, \hat{\beta}_6$ memiliki nilai statistik uji Wald (Z_2) lebih dari $Z_{\alpha/2}$ sebesar 1,65, maka diputuskan H_0 ditolak. Sehingga disimpulkan bahwa variabel persentase penduduk merokok, persentase kabupaten/kota yang melaksanakan GERMAS, dan jumlah rumah sakit secara signifikan mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021. Sedangkan penaksir parameter $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_4, \hat{\beta}_7$ memiliki nilai statistik uji Wald (Z_2) kurang dari $Z_{\alpha/2}$ sebesar 1,65, maka diputuskan H_0 gagal ditolak. Sehingga disimpulkan bahwa variabel kepadatan penduduk, persentase rumah layak huni, persentase TFU yang memenuhi syarat standar, dan persentase rumah tangga yang memiliki akses sanitasi layak secara signifikan tidak mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021.

4.6. Pengujian Heterogenitas Spasial

Setelah dilakukan pengujian parameter model NBR, dilakukan pengujian heterogenitas spasial. Hasil yang diperoleh adalah nilai Breusch-Pagan sebesar 21,216 dan p_{value} sebesar 0,0035. Selain itu, berdasarkan tabel distribusi *chi-kuadrat* didapatkan nilai $\chi^2_{(0,1;7)}$ sebesar 12,017. Karena nilai Breusch-Pagan $> \chi^2_{(0,1;7)}$ dan nilai $p_{value} < \alpha = 0,1$, maka diputuskan H_0 ditolak. Hal ini menunjukkan bahwa terdapat heterogenitas spasial pada data respon.

4.7. Pemodelan Jumlah Kasus Tuberkulosis Paru dengan GWNBR

Setelah terdeteksi adanya heterogenitas spasial, maka tahapan selanjutnya dilakukan pemodelan jumlah kasus tuberkulosis paru di Indonesia tahun 2021 dengan GWNBR yang terdiri dari 5 tahapan, yaitu menghitung jarak *Euclidean*, menghitung pembobot spasial, menentukan *bandwidth* optimum, melakukan penaksiran parameter NBR, dan pengujian parameter. Adapun jarak *Euclidean* dan pembobot spasial yang diperoleh dapat dilihat pada Tabel 5.

Tabel 5. Jarak *Euclidean* dan Pembobot Spasial

Lokasi (<i>j</i>)	Lokasi ke-1		...	Lokasi ke-34	
	d_{1j}	w_{1j}	...	d_{34j}	w_{34j}
1	0,0000	1,0000	...	46,0927	0,0000
2	3,8801	0,9385	...	42,4857	0,0000
3	8,2188	0,6770	...	40,3879	0,0000
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮
32	32,6263	0,0000	...	13,5123	0,4191
33	39,2300	0,0000	...	6,8776	0,8879
34	46,0927	0,000	...	0,0000	1,0000

Tahapan selanjutnya yaitu melakukan perhitungan *bandwidth* optimum setiap lokasi pegamatan berdasarkan nilai CV minimum. Hasil yang diperoleh dapat dilihat pada Tabel 6.

Tabel 6. *Bandwidth* Optimum

Lokasi	<i>Bandwidth</i>
1	46,0741
2	42,4671
⋮	⋮
17	26,2036
⋮	⋮
33	39,2203
34	46,0830

Berdasarkan Tabel 6, jarak *Euclidean* antar lokasi pengamatan ke-17 (Provinsi Bali) ke lokasi pengamatan ke-33 (Provinsi Papua Barat) lebih dekat daripada jarak antara lokasi pengamatan ke-17 (Provinsi Bali) ke lokasi pengamatan ke-2 (Provinsi Sumatera Utara), sehingga pembobot spasial yang diberikan pada lokasi pengamatan ke-33 untuk model GWNBR lokasi pengamatan ke-17, lebih kecil daripada pembobot spasial yang diberikan pada lokasi pengamatan ke-2.

Setelah menentukan *bandwidth* optimum untuk setiap lokasi, maka tahapan selanjutnya dilakukan pemodelan jumlah kasus tuberkulosis paru di Indonesia tahun 2021 dengan GWNBR yang terdiri dari 2 tahapan, yaitu melakukan penaksiran parameter model GWNBR dan pengujian parameter. Pengujian parameter dan uji parsial parameter model GWNBR pada lokasi pengamatan ke-33 (Provinsi Papua Barat) dapat dilihat pada Tabel 7.

Tabel 7. Hasil Pengujian Parameter dan Uji Parsial Parameter GWNBR

Lokasi	Parameter	Taksiran	Standard Error	$ Z_3 $	<i>pvalue</i>
33	$\hat{\beta}_0(u_{12}, v_{12})$	6,2697	< 0,001	335.511,3179	< 0,001*
	$\hat{\beta}_1(u_{12}, v_{12})$	< 0,001	< 0,001	36,7877	< 0,001*
	$\hat{\beta}_2(u_{12}, v_{12})$	-0,0056	0,0355	-0,1571	0,8752
	$\hat{\beta}_3(u_{12}, v_{12})$	0,0461	0,0511	0,9022	0,3669
	$\hat{\beta}_4(u_{12}, v_{12})$	0,0023	0,0053	0,4367	0,6623
	$\hat{\beta}_5(u_{12}, v_{12})$	-0,0049	0,0015	-3,2107	0,0013*
	$\hat{\beta}_6(u_{12}, v_{12})$	0,0085	0,0003	28,8012	< 0,001*
	$\hat{\beta}_7(u_{12}, v_{12})$	0,0040	0,0277	0,1456	0,8843

*Signifikan pada $\alpha = 0,1$.

Berdasarkan Tabel 7, didapatkan model GWNBR sebagai berikut:

$$\hat{\mu}(u_{33}, v_{33}) = \exp(6,2697 + 0,0000X_{1,33} - 0,0056X_{2,33} + 0,0461X_{3,33} + 0,0023X_{4,33}) + \exp(-0,0049X_{5,33} + 0,0085X_{6,33} + 0,0040X_{7,33}). \tag{38}$$

Interpretasi dari model GWNBR pada persamaan (38) sebagai berikut:

1. Nilai taksiran kepadatan penduduk sebesar 0,0000 menyatakan bahwa setiap peningkatan kepadatan penduduk 1 jiwa/km² akan meningkatkan jumlah kasus tuberkulosis paru menjadi $\exp(0,0000) = 1$ kali.
2. Nilai taksiran TFU yang memenuhi syarat standar sebesar 0,0023 menyatakan bahwa setiap peningkatan TFU yang memenuhi syarat standar 1% akan menaikkan jumlah kasus tuberkulosis paru menjadi $\exp(0,0023) = 1,0023$ kali.
3. Nilai taksiran kabupaten/kota yang melaksanakan GERMAS 1% akan menurunkan jumlah kasus tuberkulosis paru menjadi $\exp(-0,0049) = 0,9951$ kali.

Tahapan selanjutnya yaitu melakukan pengujian kesamaan model NBR dengan model GWNBR dan pengujian parameter model GWNBR secara serentak serta secara parsial. Hasil pengujian kesamaan model NBR dengan model GWNBR diperoleh nilai statistik uji F^* sebesar 42.784,14 dan p_{value} sebesar 0. Selain itu, berdasarkan tabel distribusi F didapatkan nilai $F_{(0,1;7;238)}$ sebesar 1,7422. Karena nilai $F^* > F_{(0,1;7;238)}$ dan nilai $p_{value} < \alpha = 0,1$, maka diputuskan H_0 ditolak. Hal ini menunjukkan bahwa model NBR dan model GWNBR tidak identik

Sementara itu, berdasarkan Tabel 7, didapatkan hasil pengujian parameter model GWNBR secara serentak diperoleh nilai statistik uji G_3 sebesar 93,4365 dan p_{value} sebesar 0. Selain itu, berdasarkan tabel distribusi *chi-kuadrat* didapatkan nilai $\chi^2_{(0,1;7)}$ sebesar 12,017. Karena nilai $G_3 > \chi^2_{(0,1;7)}$ dan nilai $p_{value} < \alpha = 0,1$, maka diputuskan H_0 ditolak. Sehingga disimpulkan bahwa variabel kepadatan penduduk, persentase rumah layak huni, persentase penduduk merokok, persentase TFU yang memenuhi syarat standar, persentase kabupaten/kota yang melaksanakan GERMAS, jumlah rumah sakit, dan persentase rumah tangga yang memiliki akses sanitasi layak secara serentak mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021.

Berdasarkan hasil pengujian parsial pada Tabel 7, diketahui bahwa penaksir parameter $\hat{\beta}_1(u_{12}, v_{12}), \hat{\beta}_5(u_{12}, v_{12}), \hat{\beta}_6(u_{12}, v_{12})$ memiliki nilai statistik uji Wald (Z_3) lebih dari $Z_{\alpha/2}$ sebesar 1,65, maka diputuskan H_0 ditolak. Sehingga disimpulkan bahwa variabel kepadatan penduduk, persentase kabupaten/kota yang melaksanakan GERMAS, dan jumlah rumah sakit secara signifikan mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021. Sedangkan penaksir parameter $\hat{\beta}_2(u_{12}, v_{12}), \hat{\beta}_3(u_{12}, v_{12}), \hat{\beta}_4(u_{12}, v_{12}), \hat{\beta}_7(u_{12}, v_{12})$ memiliki nilai statistik uji Wald (Z_3) kurang dari $Z_{\alpha/2}$ sebesar 1,65, maka diputuskan H_0 gagal ditolak. Sehingga disimpulkan bahwa variabel persentase rumah layak huni, persentase penduduk merokok, persentase TFU yang memenuhi syarat standar, dan persentase rumah tangga yang memiliki akses sanitasi layak secara signifikan tidak mempengaruhi jumlah kasus tuberkulosis paru disetiap provinsi di Indonesia tahun 2021.

4.8. Pemilihan Model Terbaik

Setelah dilakukan pengujian parameter model NBR dan model GWNBR, selanjutnya ialah dilakukan pemilihan model terbaik antara model NBR dan model GWNBR didasarkan kriteria AIC. Hasil yang diperoleh dapat dilihat pada Tabel 8.

Tabel 8. Pemilihan Model Terbaik dengan AIC

Model	AIC
NBR	613,3387
GWNBR	573,2879

Berdasarkan Tabel 8, menunjukkan bahwa dari kedua model tersebut, model GWNBR memiliki nilai AIC paling kecil dibandingkan dengan model NBR. Sehingga dapat disimpulkan bahwa model GWNBR lebih baik dalam memodelkan jumlah kasus tuberkulosis paru disetiap provinsi yang ada di Indonesia tahun 2021.

5. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, maka kesimpulan yang dapat diambil adalah Provinsi Papua Barat menjadi salah satu model GWNBR yang mewakili Indonesia. Karena faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis paru di Provinsi Papua Barat juga mempengaruhi di provinsi lain yang ada di Indonesia. Faktor-faktor yang berpengaruh secara lokal terhadap jumlah kasus tuberkulosis paru di Provinsi Papua Barat adalah kepadatan penduduk, persentase kabupaten/kota yang melaksanakan GERMAS, dan jumlah rumah sakit.

6. Daftar Pustaka

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. New York: John Wiley.

- Fotheringham, A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. England: John Wiley & Sons, Ltd.
- Hilbe, J. M. (2011). *Negative Binomial Regression*. New York: Cambridge University Press.
- Hilbe, J. M. (2014). *Modeling Count Data*. New York: Cambridge University Press.
- Karima, A., Suyitno, & Hayati, M. (2021). Analisis Faktor-Faktor yang Mempengaruhi Jumlah Kasus Tuberkulosis di Indonesia Menggunakan Model Geographically Weighted Poisson Regression. *Jurnal Ekspansional*, 12(1), 7-16. <https://doi.org/10.30872/eksponensial.v12i1.754>.
- Keswari, N., Sumarjaya, I., & Suciptawati, N. (2014). Perbandingan Regresi Binomial Negatif dan Regresi Generalisasi Poisson dalam Mengatasi Overdispersi (Studi Kasus: Jumlah Tenaga Kerja Usaha Pencetak Genteng di Br. Dukuh, Desa Pejaten). *Jurnal Matematika*, 3(3), 107-115. <https://doi.org/10.24843/MTK.2014.v03.i03.p072>.
- Ricardo, A., & Carvalho, T. (2013). *Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion*. New York: Springer Science.
- Sembiring, A. P. (2019). Faktor-Faktor yang Mempengaruhi Besar Klaim Asuransi Jiwa dengan Menggunakan Model Regresi Tobit. *Jurnal Bimaster*, 8(4), 729-736. <https://dx.doi.org/10.26418/bbimst.v8i4.35996>.
- Setiawan, A. (2015). *Pengantar Statistika Matematika*. Semarang: Widya Sari Press Salatiga.
- Simarmata, R., & Ispriyanti, D. (2011). Penanganan Overdispersi pada Model Regresi Poisson Menggunakan Model Regresi Binomial Negatif. *Jurnal Media Statistika*, 4(2), 95-104. <https://doi.org/10.14710/medstat.4.2.95-104>.
- Suryani, I., Yasin, H., & Kartikasari, P. (2021). Pemodelan Jumlah Kasus Demam Berdarah Dengue (DBD) di Jawa Tengah dengan Geographically Weighted Negative Binomial Regression (GWNBR). *Jurnal Gaussian*, 10(1), 135-148. <https://doi.org/10.14710/j.gauss.10.1.136-148>.
- Werdhani, R. A. (2009). *Patofisiologi, Diagnosis, dan Klasifikasi Tuberkulosis*. Jakarta: FKUI.
- Widayaka, P., Mustafid., & Rahmawati, R. (2016). Pendekatan Mixed Geographically Weighted Regression untuk Pemodelan Pertumbuhan Ekonomi Menurut Kabupaten/Kota di Jawa Tengah. *Jurnal Gaussian*, 5(4), 727-736. <https://doi.org/10.14710/j.gauss.5.4.727-736>.
- World Health Organization. (2016). *Global Tuberculosis Report 2016*. Geneva: World Health Organization.
- World Health Organization. (2020). *Global Tuberculosis Report 2020*. Geneva: World Health Organization.
- Yasin, H. (2011). Pemilihan Variabel pada Model Geographically Weighted Regression. *Jurnal Media Statistika*, 4(2), 63-72. <https://doi.org/10.14710/medstat.4.2.63-72>.