VARIANSI: Journal of Statistics and Its application on Teaching and Research **ISSN 2684-7590 (Online)**
Vol. 5 No. 3 (2023) 118-131

DOI: 10.35580/variansiunm151

GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION (GWNBR) IN MODELING THE RISK FACTORS OF PNEUMONIA DISEASE AMONG TODDLERS IN THE CENTRAL SULAWESI PROVINCE

Zakiyah Mar'ah[a], Zulkifli Rais[a*], A. Sulfiana Haris[a],

(Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia)

Abstract

This research was conducted to map and model the number of Pneumonia cases in Central Sulawesi Province using the *Geographically Weighted Negative Binomial Regression* (GWNBR) approach. The data used were Pneumonia case data in Central Sulawesi Province obtained from the Health Publication of Central Sulawesi Province in 2021. The analysis results with the GWNBR method indicated that predictor variables significantly influencing the number of Pneumonia cases in each district/city of Central Sulawesi Province were Exclusive Breastfeeding Percentage ($\boldsymbol{X_1}$), Complete Basic Immunization Percentage ($\boldsymbol{X_2}$), Percentage of Toddlers Receiving Vitamin A ($\boldsymbol{X_3}$), and Percentage of Coverage of Toddler Services ($\boldsymbol{X_5}$). Meanwhile, the variable Low Birth Weight ($\boldsymbol{X_4}$) does not significantly affect the cases.

Keywords: Pneumonia, Poisson Regression, Overdispersion, Negative Binomial Regression, GWNBR.

1. Pendahuluan

Pneumonia merupakan kondisi yang sering terjadi pada sistem pernapasan dan menjadi penyebab utama kematian anak-anak secara global, khususnya pada anak usia balita. Popularitas penyakit ini semakin meningkat, terutama sejak munculnya pandemi Covid-19 (Kemenkes, 2020). Pneumonia merupakan infeksi pernapasan akut yang menjadi penyebab utama kematian pada balita di negara-negara berkembang. Penyakit ini bisa disebabkan oleh virus, bakteri, atau jamur dan sering disebut sebagai paru-paru basah (Josefa et al., 2019).

World Health Organization (WHO) menyatakan Pneumonia sebagai penyebab kematian tertinggi pada balita melebihi penyakit lainnya seperti campak, malaria, dan AIDS. Pada tahun 2019, Pneumonia menyebabkan kematian anak dibawah 5 tahun sebesar 14% dari seluruh kematian anak dengan total kematian 740.180 jiwa. Kasus Pneumonia banyak terjadi di negara-negara berkembang seperti Asia Tenggara sebesar 39% dan Afrika sebesar 30%. WHO menyebutkan Indonesia menduduki peringkat ke 8 dunia dari 15 negara yang memiliki angka kematian balita dan anak yang diakibatkan oleh Pneumonia (WHO, 2022).

Di Indonesia, Pneumonia menduduki urutan kedua penyebab kematian pada balita setelah diare, dimana pada tahun 2019 persentase kematian akibat Pneumonia pada balita sebesar 9,5 persen dari seluruh kematian balita (Kemenkes, 2020). Jumlah tersebut turun 10,19% dibandingkan pada tahun 2020 yaitu sebanyak 309.838 kasus (Kemenkes RI, 2022).

Salah satu provinsi dengan tingkat penemuan kasus Pneumonia tertinggi di pulau Sulawesi yaitu Provinsi Sulawesi Tengah dan menduduki posisi ke-8 sebagai provinsi dengan kasus Pneumonia tertinggi di Indonesia, dengan total kasus sebesar 6.273 balita (Kemenkes, 2022). Namun, cakupan penemuan kasus Pneumonia balita di provinsi Sulawesi Tengah Tahun 2022 hanya 39,8% sedangkan target cakupan penemuannya sebesar 70%. Jika ditinjau dari kabupaten/kotanya, dari 13 kabupaten/kota-nya hanya Kabupaten Banggai yang dapat mencapai target yang ditetapkan secara nasional.

Penyakit Pneumonia merupakan penyakit menular dimana penyebarannya dapat dipengaruhi oleh kondisi geografis

[1] Corresponding author.
E-mail address: rahmatwahyudi13@gmail.com

suatu wilayah. Setiap wilayah tentunya memiliki kondisi geografis, ekonomi, dan sosial budaya yang berbeda, sehingga faktor-faktor yang mempengaruhi terjadinya Pneumonia di suatu wilayah juga akan berbeda. Salah satu cara untuk mengetahui faktor yang berpengaruh secara signifikan terhadap penyakit Pneumonia adalah analisis regresi.

Analisis regresi merupakan suatu teknik untuk membangun persamaan linear dan menggunakan persamaan tersebut untuk melakukan estimasi (Mar'ah et al., 2023). Salah satu model regresi yang dapat digunakan untuk menjelaskan hubungan antara variabel respon yang berupa data diskrit dengan peluang kejadian yang jarang terjadi dalam selang waktu tertentu dan variabel bebas yang berupa kontinu atau diskrit adalah regresi Poisson.

Analisi regresi Poisson adalah salah satu regresi nonlinier yang sering digunakan untuk memodelkan hubungan antara variabel respon yang berupa data diskrit dengan variabel prediktor yang berupa data diskrit atau kontinu. Pada model regresi Poisson terdapat asumsi yang harus dipenuhi yaitu nilai variansi yang diperoleh sama dengan nilai rataannya, keadaan ini biasa disebut equidispersi. Namun pada kenyataannya asumsi ini sangat jarang terjadi. Biasanya data *count* memiliki variansi lebih besar dari rataan yang disebut overdispersi atau sebaliknya yaitu rataan lebih besar dari variansi atau biasa disebut undispersi (Safrida et al., 2013). Salah satu solusi untuk mengatasi kasus overdispersi yaitu dengan regresi Binomial Negatif.

Regresi Binomial Negatif dapat digunakan untuk memodelkan data Poisson yang mengalami overdispersi karena distribusi Binomial Negatif merupakan perluasan dari distribusi Poisson-Gamma yang memuat parameter dispersi (Widyaningsih et al., 2021). Pada pemodelan data, seringkali ditemukan adanya efek atau pola spasial pada data, sehingga perlu dilakukan uji efek spasial.

Uji efek spasial meliputi pengujian dependensi spasial dan heterogenitas spasial. Dependensi spasial menunjukkan bahwa pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang letaknya berdekatan. Sedangkan heterogenitas spasial adalah perbedaan kondisi antara suatu lokasi dengan lokasi lainnya, yang ditinjau dari segi geografis, keadaan sosial-budaya maupun hal lain yang dapat menimbulkan kondisi heterogenitas spasial (Kurniasari & Ariastita, 2014). Jika suatu data mengalami kasus overdispersi maka digunakan analisis regresi Binomial Negatif untuk mengatasinya. Dengan memperhatikan efek spasialnya, maka digunakan metode *Geographically Weighted Negative Binomial Regression* (GWNBR).

Geographically Weighted Negative Binomial Regression (GWNBR) merupakan salah satu metode untuk memodelkan data cacah yang mempunyai heterogenitas spasial dan overdispersi. Model ini akan menghasilkan estimasi parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda (Rini, 2018).

2. Kajian Pustaka

2.1 Analisis Regresi

Secara umum, analisis regresi adalah metode yang digunakan untuk memahami hubungan antara satu atau lebih variabel prediktor (independen) dengan variabel respon (dependen). Tujuan utama dari analisis regresi adalah untuk menemukan hubungan fungsional antara variabel-variabel tersebut sehingga dapat digunakan untuk memprediksi nilai variabel respon berdasarkan nilai-nilai variabel prediktor (Purba & Purba, 2022).

1) Regresi Linear Sederhana

Persamaan regresi linear sederhana secara umum yaitu:

$$\hat{Y} = a + bx \tag{2.1}$$

Keterangan:

$\hat{Y}$: Respon (Variabel Terikat/*dependent*)
a : Konstanta
b : Koefisien regresi variabel *independent*
x : Prediktor (variabel bebas/*independent*)

2) Regresi Linear Berganda

Regresi linier ganda berguna untuk meramal variabel *dependent* yang dipengaruhi oleh dua atau lebih variabel *independent*Adapun rumus yang dipakai disesuaikan dengan jumlah variabel yang diteliti, yaitu sebagai berikut:

$$\hat{Y} = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n \tag{2.2}$$

Keterangan :

$\hat{Y}$: Respon (Variabel Terikat/*dependent*)
a : Konstanta
$b_1, b_2, \dots, b_n$: Koefisien regresi variabel *independent*
$X_1, X_2, \dots, X_n$: Prediktor (variabel bebas/*independent*)

2.2 Multikolinearitas

Multikolinearitas adalah suatu kondisi dimana terjadi korelasi antara variabel bebas atau antar variabel bebas tidak bersifat saling bebas. Variabel $\boldsymbol{X_1}, \boldsymbol{X_2}, \dots, \boldsymbol{X_p}$ dikatakan bersifat saling bebas jika matriks korelasi antar variabel membentuk matriks indentitas. Dalam model regresi, adanya korelasi antar variabel prediktor menyebabkan taksiran parameter regresi yang dihasilkan akan memiliki error yang sangat besar.
Pendeteksian kasus multikolinearitas dapat dilihat melalui beberapa cara yaitu sebagai berikut:

a. Jika koefisien korelasi pearson (r_{ij}) antar variabel prediktor lebih dari 0,95 ($r_{ij} > 0{,}95$) maka terdapat korelasi antar variabel tersebut.
b. Nilai VIF (*Varian Inflation Factor*) lebih besar dari 10 menunjukkan adanya multikolinearitas antarvariabel prediktor. Nilai VIF dinyatakan sebagai berikut:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{2.3}$$

dengan R_j^2 adalah koefisien determinasi antar X_j dengan variabel prediktor lainnya (Dean & Hocking, 1997).

2.2 Regresi Poisson

Distribusi Poisson merupakan distribusi yang digunakan untuk data jumlah kejadian yang terjadi secara acak pada interval waktu tertentu dengan hasil pengamatan berupa variabel diskrit (Cahyandari, 2014). Regresi Poisson merupakan model regresi nonlinear yang sering digunakan untuk menganalisis suatu data *count* . Regresi Poisson adalah salah satu regresi yang digunakan untuk memodelkan antara variabel respon dan variabel prediktor dengan mengasumsikan variabel *Y* berdistribusi Poisson. Persamaan model regresi Poisson dapat ditulis sebagai berikut (Rahmadeni Feni Fatkhuli, 2019):

$$\hat{\mu}_i = \exp\,(\boldsymbol{X}_i^T\boldsymbol{\beta})$$
$$\hat{\mu}_i = \exp\,(\beta_0 + \beta_1X_{i1} + \beta_2X_{i2} + \cdots + \beta_pX_{ip})$$
$$\ln\,(\hat{\mu}_i) = \beta_0 + \beta_1X_{i1} + \beta_2X_{i2} + \cdots + \beta_pX_{ip} \tag{2.4}$$

dengan μ_i merupakan rata-rata jumlah kejadian yang terjadi dalam interval waktu tertentu.

2.3 Estimasi Parameter Model Regresi Poisson

Estimasi parameter model regresi Poisson dilakukan dengan metode *maximum likelihood estimation* (MLE) yaitu dengan cara memaksimumkan fungsi *likelihood*. Fungsi *likelihood* dari regresi Poisson adalah sebagai berikut (Cameron dan Trivedi, 1998):

a. Mengambil n data sampel random
b. Membentuk fungsi likelihood dari regresi Poisson, yaitu

$$lnL(\beta) = ln\left(\prod_{i=1}^{n} \frac{\exp(-\mu_i)\,\mu_i^{y_i}}{y_i!}\right)$$
$$= \sum_{i=1}^{n} \ln\left(\prod_{i=1}^{n} \frac{\exp(-\mu_i)\,\mu_i^{y_i}}{y_i!}\right)$$

$$= -\sum_{i=1}^{n} e^{x_i{}^T\beta} + \sum_{i=1}^{n} y_i x_i{}^T\beta - \sum_{i=1}^{n} ln(y_i!)$$ 2.5

Kemudian persamaan di atas diturunkan terhadap β^T yang merupakan bentuk vektor, menjadi

$$\frac{\partial lnL(\beta)}{\partial\beta^T} = -\sum_{i=1}^{n} X_i \exp(X_i\beta) + \sum_{i=1}^{n} y_i X_i$$ 2.6

2.5 Overdispersi

Regresi Poisson dikatakan overdispersi apabila nilai variansnya lebih besar dari nilai rata-ratanya. Jika pada data diskrit terjadi overdispersi dan tetap menggunakan regresi Poisson sebagai metode penyelesaiannya, maka akan diperoleh suatu kesimpulan yang tidak valid karena nilai *standart error* menjadi *under estimate*. Hal ini disebabkan karena parameter koefisien regresi yang dihasilkan dari regresi Poisson tidak efisien meskipun koefisien regresinya tetap konsisten (Damayanti CR & Yanti, 2022).
Overdispersi dapat diketahui melalui nilai parameter dispersi yang diperoleh melalui rumus:

$$\boldsymbol{\theta} = \frac{\boldsymbol{G^2}}{\boldsymbol{df}}$$ 2.7

dengan
$\boldsymbol{\theta}$: parameter dispersi
df : *degree of freedom*
$\boldsymbol{G^2}$: Nilai *devians*, didefinisikan sebagai

$$\boldsymbol{G^2 = 2\sum_{i=1}^{n} y_i ln\left(\frac{y_i}{\lambda_i}\right)}$$ 2.8

dimana,
$\boldsymbol{y_i}$: frekuensi observasi untuk kategori atau kelompok data ke-i
$\boldsymbol{\lambda_i}$: frekuensi yang diharapkan untuk kelompok data ke-i
Jika $\boldsymbol{\theta}$ > 1 artinya terjadi overdispersi pada regresi Poisson, jika $\boldsymbol{\phi}$< 1 artinya terjadi underdispersi dan jika $\boldsymbol{\theta}$: 1 berarti tidak terjadi kasus over/under dispersi yang disebut dengan equidispersi (Famoye, Wulu, dan Singh, 2004).

2.6 Regresi Binomial Negative

1) Model Binomial Negatif

$$\widehat{y_j} \sim NB\left[t_j exp\left(\sum_p \beta_p x_{jp}\right), \theta\right]; j = 1,2,3 \dots, n$$ 2.9

$\boldsymbol{E(Y) = \mu}$ dan $\boldsymbol{Var(Y) = \mu + \alpha\mu^2}$ dengan $\boldsymbol{\theta}$ adalah parameter dispersi. Kemudian fungsi probabilitas Binomial Negatif dapat dinyatakan sebagai berikut (Greene, 2008):

$$\boldsymbol{f(y;\mu,\theta) = \frac{\Gamma\left(y+\frac{1}{\theta}\right)}{y!\Gamma\left(\frac{1}{\theta}\right)}\left(\frac{1}{1+\theta\mu}\right)^{\frac{1}{\alpha}}\left(\frac{\theta\mu}{1+\theta\mu}\right)^{y};\ y = 0,1,2 \dots, n}$$ 2.10

Model regresi Binomial Negatif dinyatakan sebagai berikut.

$$\widehat{y_j} = \exp\left(\beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \cdots + \beta_p X_{pj}\right)$$ 2.11

2) Estimasi Parameter Model Regresi Binomial Negatif

Metode *Maximum Likelihood Estimation* (MLE) digunakan untuk estimasi parameter dalam regresi Binomial Negatif. Pada estimasi regresi Binomial Negatif, untuk memaksimumkan fungsi *likelihood* maka digunakan metode iterasi *Newton Raphson*. Fungsi *likelihood* dari regresi Binomial Negatif adalah sebagai berikut:

$$\boldsymbol{L(\beta,\theta) = \prod_{i=1}^{n} \frac{\Gamma\left(y+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y!}\left(\frac{1}{1+\theta\mu_i}\right)^{\frac{1}{\theta}}\left(\frac{\theta\mu_i}{1+\theta\mu_i}\right)^{y}}$$ 2.12

dengan

$\boldsymbol{L(\beta,\theta)}$: fungsi *likelihood*
y : variabel acak yang mengikuti distribusi Poisson
θ : parameter distribusi Invers-Gamma
$\boldsymbol{\mu_i}$: adalah parameter mean dari distribusi Poisson untuk observasi ke-i.

3) Pengujian Parameter Model

Pengujian parameter model regresi Poissson dan Binomial Negatif terdiri dari uji serentak dan uji parsial. Hipotesis untuk uji signifikansi secara serentak dengan *Maximum Likelihood Ratio Test* (MLRT), yaitu:

$\boldsymbol{H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0}$

$\boldsymbol{H_1}$: paling sedikit ada satu $\boldsymbol{\beta_j \neq 0; j = 1, 2, \dots, p}$

Statistik Uji:

$$\boldsymbol{D(\hat{\beta}) = -2ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right)} \quad (2.13)$$

dimana,

$\boldsymbol{\ln L(\hat{\omega})}$: fungsi *likelihood* untuk model tanpa melibatkan variable prediktor

$\boldsymbol{\ln L(\hat{\Omega})}$: fungsi *likelihood* untuk model yang melibatkan variable prediktor

Statistik uji yang digunakan mengikuti distribusi *chi-square* dengan derajat bebas *p* dengan keputusan tolak $\boldsymbol{H_0}$ jika $\boldsymbol{D(\hat{\beta}) > \chi^2_{(p;\alpha)}}$ yang berarti bahwa paling sedikit ada satu parameter yang berpengaruh secara signifikan terhadap variabel respon.

Setelah dilakukan dengan uji serentak dilanjutkan dengan uji signifikansi secara parsial untuk mengetahui parameter mana saja yang memberikan pengaruh yang signifikan terhadap model, hipotesis uji signifikansi secara parsial yaitu sebagai berikut:

$\boldsymbol{H_0 : \beta_j = 0}$

$\boldsymbol{H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, p}$

Statistik Uji:

$$\boldsymbol{Z_{hitung} = \frac{\widehat{\beta_J}}{se(\widehat{\beta_J})}} \quad (2.14)$$

dimana,

$\boldsymbol{\widehat{\beta_J}}$: koefisien model variable prediktor ke-j

$\boldsymbol{se\widehat{\beta_J}}$: standard error dari estimasi *maximum likelihood*

Tolak $\boldsymbol{H_0}$ jika statistik uji Z atau $\boldsymbol{|Z_{hitung}| > Z_{\left(\frac{\alpha}{2}\right)}}$ atau nilai *P-Value* < α. Tolak $\boldsymbol{H_0}$ artinya bahwa parameter ke-j signifikan terhadap model

2.7 Uji Efek Spasial

1) Dependensi Spasial

Dependensi spasial menunjukkan bahwa pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang letaknya berdekatan. Pengujian dependensi spasial dapat dilakukan dengan uji *Moran's I* (Anselin & Bera, 1998).

Hipotesis uji *Moran's I* adalah sebagai berikut:

H_0 : $\boldsymbol{\mu_i}$: 0 (tidak ada dependensi antar lokasi)

H_0 : $\boldsymbol{\mu_i} \neq 0$ (ada dependensi antar lokasi)

Statistik Uji

$$\boldsymbol{Z = \frac{I - E(I)}{\sqrt{var(I)}}} \quad (2.15)$$

dimana:

$$\boldsymbol{Var(I) = \frac{n^2.S_1 - n.S_2 + 3.S_0^2}{(n^2-1)S_0^2} - [E(I)]^2 I = \frac{n\sum_{i=1}^{n}\sum_{i=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{s_0 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

$$\boldsymbol{S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{ij} + w_{ji})^2 \;\; E(I) = -\frac{1}{n-1} S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

dengan:

$\boldsymbol{y_i}$: nilai pengamatan pada lokasi ke-i

$\boldsymbol{y_j}$: nilai pengamatan pada lokasi ke-i

$\boldsymbol{\bar{y}}$: rata-rata

$\boldsymbol{w_{ij}}$: matriks pembobot terstandarisasi antara lokasi i dan j

$\boldsymbol{n}$: Banyaknya lokasi pengamatan

Tolak H_0, jika $|Z_{hitung}| > Z_{\frac{\alpha}{2}}$ atau *p-value* < α, yang berarti terdapat dependensi spasial.

2) Heterogenitas Spasial

Pengujian heteroskedastisitas spasial dilakukan untuk melihat apakah terdapat kekhasan pada setiap lokasi pengamatan, sehingga parameter regresi yang dihasilkan berbeda-beda secara spasial. Pengujian heteroskedastisitas spasial dilakukan menggunakan statistik uji *Breusch-Pagan* (BP) dengan hipotesis sebagai berikut (Pratama & Wulandari, 2015):

$\boldsymbol{H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_p^2 = \sigma^2}$ (varians antar lokasi sama)

$\boldsymbol{H_1}$: paling sedikit ada satu $\boldsymbol{\sigma_i^2 \neq \sigma^2}$ (varians antar lokasi berbeda)

Statistik uji

$$BP = \left(\frac{1}{2}\right) f^T Z (Z^T Z)^{-1} Z^T f \sim \chi^2_{(p)} \qquad 2.16$$

dimana,

$\boldsymbol{p}$: banyaknya variabel prediktor

$\boldsymbol{f}$: $\boldsymbol{(f_1, f_2, \dots, f_n)^T}$ dengan $\boldsymbol{f_i = \frac{e_i^2}{\sigma^2} - 1}$

e_i : error untuk obvservasi ke-i

$\boldsymbol{\sigma^2}$: variansi dari $\boldsymbol{e_i}$

$\boldsymbol{\sigma_i^2}$: variansi dari variabel prediktor ke-i

Z : matriks berukuran n × (p+1) yang berisi vektor konstan

Tolak H_0 jika statistic uji $BP > \chi^2{}_{(p,\alpha)}$ atau *p-value* < α, yang Berarti variansi antar lokasi berbeda.

2.8 Geographically Weighted Negative Binomial Regression (GWNBR)

Geographically Weighted Negative Binomial Regression (GWNBR) merupakan salah satu metode yang cukup efektif untuk menduga parameter lokal dengan unit amatan berupa wilayah pada data *count* yang memiliki heterogenitas spasial dan mengalami kasus overdispersi. Model GWNBR akan menghasilkan parameter lokal yang masing-masing lokasi memiliki parameter yang berbeda-beda (Pratama & Wulandari, 2015). Model GWNBR dirumuskan sebagai berikut:

$$\widehat{h_j} \sim NB\left[t_j exp\left(\sum_k \beta_k(u_j, v_j)x_{jk}\right), \theta(u_j, v_j)\right]; j = 1, \dots, n \tag{2.17}$$

dimana,

$\widehat{h_j}$	: nilai observasi variabel respon ke-j
t_j	: *offset* variabel
x_{jk}	: nilai observasi variabel prediktor ke-k pada pengamatan lokasi (u_j, v_j)
$\beta_k(u_j, v_j)$	: koefisien regresi variabel prediktor ke-k untuk setiap lokasi (u_j, v_j)
$\theta(u_j, v_j)$	: parameter disperse untuk setiap lokasi (u_j, v_j)
u_j	: koordinat latitude titik pengamatan ke-j
v_j	: koordinat longitude titik pengamatan ke-j

Model GWNBR dapat ditulis sebagai berikut:

$$h_j = \exp\left(\beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \cdots + \beta_p X_{pj}\right) \tag{2.18}$$

1) Estimasi Parameter Model GWNBR

Estimasi parameter model GWNBR dilakukan dengan menggunakan metode *Maximum Likelihood Estimate* (MLE). Fungsi *likelihood* dituliskan sebagai berikut:

$$\boldsymbol{L(\beta_{(u_i,v_i)}, \theta_i | y_i, x_i) = \prod_{i=1}^{n}\left(\prod_{r=0}^{y_i-1}\left(\mathbf{r} + \frac{1}{\theta_i}\right)\right)\frac{1}{(y_i!)}\left(\frac{1}{1+\theta_i\mu_i}\right)^{\frac{1}{\theta}}\left(\frac{\theta_i\mu_i}{1+\theta_i\mu_i}\right)^{y_i}} \tag{2.19}$$

$\boldsymbol{L(\beta, \theta)}$: fungsi *likelihood*
$\boldsymbol{y}$: variabel acak yang mengikuti distribusi Poisson
$\boldsymbol{\theta_i}$: parameter distribusi Invers-Gamma
$\boldsymbol{\mu_i}$: adalah parameter mean dari distribusi Poisson untuk observasi ke-i.

2) Pengujian Parameter Model GWNBR

Pengujian signifikansi parameter model GWNBR terdiri dari uji serentak dan uji parsial. Uji signifikansi secara serentak dengan menggunakan *Maximum Likelihood Ratio Test* (MLRT), dengan hipotesis sebagai berikut:

$\boldsymbol{H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \cdots = \beta_p(u_i, v_i) = 0}$

$\boldsymbol{H_1}$: paling sedikit ada satu $\boldsymbol{\beta_j(u_i, v_i) \neq 0; j = 1, 2, \dots, p}$

Statistik Uji:

$$\boldsymbol{D(\widehat{\beta}) = -2ln\left(\frac{L(\widehat{\omega})}{L(\widehat{\Omega})}\right)} \tag{2.20}$$

dimana,

$\boldsymbol{\ln L(\widehat{\omega})}$: fungsi *likelihood* untuk model tanpa melibatkan variable prediktor

$\boldsymbol{\ln L(\widehat{\Omega})}$: fungsi *likelihood* untuk model yang melibatkan variable prediktor

Tolak $\boldsymbol{H_0}$ jika $\boldsymbol{D(\widehat{\beta}) > \chi^2_{(p;\alpha)}}$ yang berarti bahwa paling sedikit ada satu parameter yang berpengaruh secara signifikan terhadap variabel respon.

Kemudian dilanjutkan dengan uji signifikansi secara parsial dengan hipotesis uji signifikansi sebagai berikut:

$\boldsymbol{H_0 : \beta_j(u_i, v_i) = 0}$

$\boldsymbol{H_1 : \beta_j(u_i, v_i) \neq 0 ; j = 1, 2, \dots, p}$

Statistik Uji:

$$\boldsymbol{Z_{hitung} = \frac{\widehat{\beta_j}(u_i,v_i)}{se(\widehat{\beta_j}(u_i,v_i))}} \tag{2.21}$$

dimana,

$\widehat{\boldsymbol{\beta}_J}$: koefisien model variable prediktor ke-j

$\boldsymbol{se}\widehat{\boldsymbol{\beta}_J}$: standard error dari estimasi *maximum likelihood*

Tolak $\boldsymbol{H_0}$ jika statistik uji Z atau $\left|\boldsymbol{Z_{hitung}}\right| > \boldsymbol{Z}_{\left(\frac{\alpha}{2}\right)}$, yang Berarti bahwa parameter tersebut berpengaruh secara signifikan terhadap variable respon di tiap lokasi.

2.9 Pemilihan Bandwidth dan Pembobot Optimum

Pemilihan *bandwidth* optimum menjadi sangat penting karena akan mempengaruhi ketepatan model terhadap data, yaitu mengatur varians dan bias dari model. Penentuan *bandwidth* optimum dilakukan menggunakan metode *Cross Validation* (CV) sebagai berikut:

$$CV(k) = \sum_{i=1}^{n}\left(y_i - \hat{y}_{\neq i}(k)\right)^2 \qquad 2.22$$

dengan $\hat{y}_{\neq i}(k)$ merupakan nilai penaksir y_i dengan pengamatan lokasi (u_i, v_i) dihilangkan dari proses penaksiran.

Dalam penaksiran GWNBR di suatu titik (u_i, v_i) diperlukan adanya pembobot spasial. Peran pembobot dalam GWNBR sangat penting karena nilai pembobot mewakili letak data observasi antara satu dengan yang lainnya. Dimana pembobot yang digunakan adalah kernel *adaptive bisquare.* Fungsi kernel *adaptive bisquare* dirumuskan sebagai berikut:

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{k}\right)^2\right)^2, \text{untuk } d_{ij} \leq k \\ 0 \qquad , \text{untuk } d_{ij} > k \end{cases} \qquad 2.23$$

dengan

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i, v_j\right)^2} \qquad 2.24$$

d_{ij} adalah jarak *Euclidean* antara lokasi (u_i, v_i) ke lokasi (u_j, v_j) dan *k* adalah nilai *bandwidth* yang menunjukkan jumlah atau proporsi dari observasi untuk dimasukkan pada estimasi parameter pada lokasi pengamatan.

2.10 Pemilihan Model Terbaik

1) Akaike Information Criterion (AIC)

Rumus AIC adalah sebagai berikut:

$$\boldsymbol{AIC = 2p - 2ln\hat{L}(\hat{\theta})} \qquad 2.25$$

dimana,

$\boldsymbol{\hat{L}(\hat{\theta})}$: nilai likelihood yang didapat dari fungsi log likelihood model regresi

$\boldsymbol{p}$: jumlah parameter dalam model

2) McFadden's R-Squared

Pada pemodelan regresi Poisson dan Negatif Binomial tidak terdapat nilai R^2. Sebagai gantinya, metode yang dapat digunakan yaitu *McFadden's R-Squared*, yang berkisar antara 0 dan 1, dengan nilai tertinggi menunjukkan kecocokan model. Adapun rumus yang digunakan yaitu:

$$\boldsymbol{McFadden's\ R - Squared = 1 - \left(\frac{loglikelihood_{model}}{loglikelihood_{null}}\right)} \qquad 2.26$$

$\boldsymbol{loglikelihood_{model}}$: Nilai loglikelihood dari model yang dipasang

$loglikelihood_{null}$: Nilai loglikelihood dari model nol (model yang hanya melibatkan intercep)

3. Metode Penelitian

Adapun teknik analisis data yang dilakukan dalam penelitian ini yaitu sebagai berikut:

1. Mengambil data kasus Pneumonia pada balita di Provinsi Sulawesi Tengah yang diperoleh dari Profil Kesehatan Provinsi Sulawesi Tengah
2. Melakukan analisis deskriptif untuk melihat karakteristik variabel kasus Pneumonia dan faktor-faktor yang diduga mempengaruhinya di Provinsi Sulawesi Tengah
3. Pengujian kasus multikolinearitas dengan melihat koefisien korelasi pearson dan nilai VIF
4. Melakukan pemodelan regresi Poisson, dengan langkah-langkah sebagai berikut:
   a. Penaksiran parameter model regresi Poisson menggunakan metode MLE dan menyusunnya kedalam model
   b. Menguji signifikansi parameter model regresi Poisson secara serentak dan uji parsial
5. Pengujian asumsi dispersi regresi Poisson
6. Melakukan pemodelan regresi Binomial Negatif, dengan langkah-langkah sebagai berikut:
   a. Penaksiran parameter model regresi Poisson menggunakan metode MLE dan menyusunnya kedalam model
   b. Menguji signifikansi parameter model regresi Poisson secara serentak dan uji parsial
7. Melakukan uji efek spasial
   a. Menguji dependensi spasial dengan uji *Moran's I*
   b. Menguji heterogenitas spasial degan uji *Breusch-Pagan*
8. Pemodelan GWNBR untuk kasus Pneumonia di Provinsi Sulawesi Tengah, dengan langkah-langkah sebagai berikut:
   a. Menghitung jarak *Euclidean* antar lokasi pengamatan berdasarkan posisi geografis
   b. Memperoleh *bandwidth* optimal untuk setiap lokasi pengamatan dengan menggunakan *Cross Validation* (CV)
   c. Menghitung matriks pembobot dengan menggunakan fungsi kernel *Adaptive Bisquare*
   d. Pengujian parameter model GWNBR dengan uji signifikansi terhadap model yaitu uji serentak dan uji parsial
   e. Melakukan pemodelan dan interpretasi model GWNBR
   f. Pemilihan model terbaik dengan kriteria AIC dan *McFadden's R-Squared*

4. Hasil dan Pembahasan

4.1 Hasil

1) Multikolinearitas

Tabel 4.1 Koefisien Korelasi antara Variabel Prediktor

	X_1	X_2	X_3	X_4
X_2	0,2310			
X_3	0,4805	0,3167		
X_4	0,2057	-0,2292	0,3321	
X_5	-0,1783	0,3219	-0,1286	-0,27665

Berdasarkan Tabel 4.1 dapat diketahui bahwa semua variabel prediktor memiliki koefisien korelasi Pearson yang kurang dari 0,95 yang artinya tidak terdapat kasus multikolinieritas.

Tabel 4.2 Nilai VIF dari Variabel Prediktor

Variabel Prediktor	VIF

X_1	1,3671
X_2	1,4914
X_3	1,6162
X_4	1,3351
X_5	1,2516

Berdasarkan Tabel 4.2 menunjukkan nilai VIF dari masing-masing variabel prediktor memiliki nilai yang kurang dari 10, dapat disimpulkan bahwa tidak terdapat kasus multikolinieritas

2) Pemodelan Regresi Poisson

Tabel 4.3 Estimasi Parameter Model Regresi Poisson

	Estimasi	Std. Error	*P-Value*	Signifikansi
(Intercept)	0,8341	0,2163	0,0001	Signifikan
X_1	0,0650	0,0020	$< 2 \times 10^{-16}$	Signifikan
X_2	-0,0158	0,0011	$< 2 \times 10^{-16}$	Signifikan
X_3	0,0109	0,0021	$2,33 \times 10^{-7}$	Signifikan
X_4	-0,0408	0,0110	$0,0002$	Signifikan
X_5	0,0220	0,0006	$< 2 \times 10^{-16}$	Signifikan
Devians : 669,59				Df : 7
AIC : 776,65				

Berdasarkan Tabel 4.3, diperoleh model regresi Poisson sebagai berikut:

$$\hat{\mu}_{Sulteng} = \exp\left(0,8341 + 0,0650X_1 - 0,0158X_2 + 0,0109X_3 - 0,0408X_4 + 0,0220X_5\right)$$

3) Pemeriksaan Overdispersi

Tabel 4.4 Parameter Dispersi Regresi Poisson

$D(\hat{\beta})$	Df	θ
669,59	7	95,6557

Rasio nilai devians dengan derajat bebasnya bernilai 95,6557. Nilai tersebut lebih besar dari 1 yang artinya data jumlah kasus Penumonia mengalami kasus overdispersi.

4) Pemodelan Regresi Binomial Negatif

Tabel 4.5 Estimasi Parameter Model Regresi Binomial Negatif

	Estimasi	Std. Error	*P-Value*	Signifikansi
(Intercept)	2,7305	1,4767	0,0644	Signifikan
X_1	0,0610	0,0170	0,0003	Signifikan
X_2	-0,0124	0,0084	0,1405	Tidak Signifikan
X_3	-0,0103	0,0195	0,5954	Tidak Signifikan
X_4	-0,0293	0,0815	0,7190	Tidak Signifikan
X_5	0,0205	0,0042	$1,6 \times 10^{-6}$	Signifikan
Devians : 13,559				Df : 7
AIC : 175,98				

Berdasarkan Tabel 4.5, diperoleh model regresi Binomial Negatif sebagai berikut:

$$\hat{y}_{Sulteng} = \exp(2{,}7305 + 0{,}0610X_1 - 0{,}0124X_2 - 0{,}0103X_3 - 0{,}0293X_4 + 0{,}0205X_5)$$

Nilai devians ($D(\hat{\beta})$) pada Tabel 4.5 dibagi dengan derajat bebasnya sebesar 1,937. Nilai ini jauh lebih kecil jika dibandingkan dengan nilai rasio model regresi Poisson pada Tabel 4.4 yaitu 95,6557. Serta nilai AIC model regresi Binomial Negatif sebesar 175,98 juga lebih kecil dari nilai AIC model regresi Poisson yaitu 776,65. Hal ini menunjukkan bahwa regresi Binomial Negatif dapat mengatasi kasus overdispersi pada regresi Poisson, sehingga model regresi Binomial Negatif lebih baik daripada model regresi Poisson.

5) Pengujian Efek Spasial

a. Uji Dependensi Spasial

Tabel 4.6 Uji Moran's *I*

I	***E(I)***	***Var(I)***	***P-Value***
-0,0928	-0,0833	0,0732	0,8966

Berdasarkan Tabel 4.6, diperoleh nilai *p-value* sebesar 0,8966 sehingga dengan taraf nyata α = 10% didapatkan kesimpulan bahwa gagal tolak H_0 yang artinya tidak ada dependensi spasial

b. Uji Heterogenitas Spasial

Tabel 4.7 Uji Breusch-Pagan

BP	***DF***	***P-Value***
10,675	5	0,0582

Nilai BP diperoleh sebesar 10,675 dengan *p-value* sebesar 0,0582. Dengan menggunakan taraf nyata α = 10% dan derajat bebas sebesar 5, diperoleh nilai $\boldsymbol{\chi^2_{(5;0,1)}}$ = 9,236. Sehingga berdasarkan kedua kriteria (*p-value* dan *Breusch-Pagan*) dimana nilai BP = 10,675 > $\boldsymbol{\chi^2_{(5;0,1)}}$ = 9,236 dan *p-value* = 0,0582 < α = 10%, dapat disimpulkan bahwa tolak H_0 dimana variansi antar lokasi berbeda atau terdapat perbedaan karakteristik antara satu titik pengamatan dengan pengamatan lainnya.

6) Pemodelan Geographically Weighted Negative Binomial Regression (GWNBR)

a. Pengujian Parameter Model GWNBR

Tabel 4.8 Estimasi Parameter Model GWNBR

Kabupaten/Kota	**Variabel yang Signifikan**
Kabupaten Banggai	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Banggai Kepulauan	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Banggai Laut	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Buol	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Donggala	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kota Palu	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Morowali	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Morowali Utara	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Parigi Moutong	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Poso	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Sigi	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Tojo Una Una	$\boldsymbol{X_1, X_2, X_3, X_5}$
Kabupaten Toli-toli	$\boldsymbol{X_1, X_2, X_3, X_5}$

Berdasarkan Tabel 4.8, diperoleh hasil bahwa pada taraf signifikansi 5%, seluruh variabel berpengaruh signifikan terhadap kasus Pneumonia di tiap Kabupaten/kota Provinsi Sulawesi Tengah kecuali variabel BBLR $(\boldsymbol{X_4})$.

b. Pemilihan Model Terbaik

Tabel 4.9 Pemilihan Model Terbaik

Model	Nilai AIC	Nilai *McFadden's R-Squared*
Regresi Poisson	776,65	77,38%
Regresi Binomial Negatif	175,98	68,31%
GWNBR	**171,21**	**92,39%**

4.2 Pembahasan

Data jumlah kasus Pneumonia di Provinsi Sulawesi Tengah merupakan data *count* yang mengikuti distribusi Poisson. Namun, setelah pengecekan asumsi equidispersi pada regresi Poisson, model tersebut memiliki nilai parameter dispersi (Tabel 4.4) sebesar 95,6557 > 1, artinya terjadi overdispersi. Jika tetap dilanjutkan dengan menggunakan regresi Poisson sebagai metode penyelesaiannya, maka akan diperoleh suatu kesimpulan yang tidak valid karena nilai *standart error* menjadi *under estimate* sehingga berdampak pada suatu variabel ang bisa saja muncul sebagai parameter yang signifikan padahal variabel tersebut sebenarnya tidak signiifikan (Damayanti CR & Yanti, 2022). Oleh karena, solusi yang dapat dilakukan untuk mengatasi masalah tersebut yaitu menggunakan metode regresi Binomial Negatif.

Selanjutnya dilakukan pengujian efek spasial untuk melihat apakah terdapat perbedaan karakteristik pada setiap kabupaten/kota di Provinsi Sulawesi Tengah (Heterogenitas Spasial) atau pengamatan di suatu lokasi dipengaruhi oleh pengamatan di lokasi lainnya yang letaknya berdekatan (Dependensi Spasial). Berdasarkan uji dependensi spasial, diperoleh nilai *p-value* (0,8966) > α (0,1), sehingga disimpulkan bahwa tidak terdapat dependensi spasial pada data. Kemudian dilakukan pengujian heterogenitas spasial dan diperoleh nilai *p-value* (0,0582) < α (0.1), sehingga disimpulkan bahwa terdapat perbedaan karakteristik antara satu titik pengamatan dengan pengamatan lainnya. Maka dapat dilanjutkan pada pemodelan dengan metode *Geographically Weighted Negative Binomial Regression* (GWNBR).

Dalam pemodelan GWNBR dibutuhkan matriks pembobot, pada penelitian ini, matriks pembobot tersebut dihitung dengan menggunakan fungsi kernel *Adaptive Bisquare.* Langkah awalnya adalah menghitung jarak antar lokasi dengan rumus jarak euclidean berdasarkan koordinat lintang dan bujur. Setelah memperoleh jarak antar lokasi maka dilanjutkan dengan menentukan *bandwidth*, nilai *bandwidth* optimum diperoleh sebesar 0,9999. Nilai *bandwidth* ini digunakan untuk perhitungan matriks pembobot. Perhitungan matriks pembobot dilakukan dengan menggunakan fungsi kernel *adaptive bisquare* dan membentuk matriks pembobot.

Estimasi parameter model GWNBR pada taraf signifikansi 5%, diperoleh hasil bahwa seluruh variabel prediktor berpengaruh terhadap jumlah kasus Pneumonia di tiap kabupaten/kota Provinsi Sulawesi Tengah Tahun 2021 kecuali variabel BBLR $(\mathbf{X_4})$.

Pemodelan banyaknya kasus Pneumonia menggunakan metode GWNBR diharapkan memberikan hasil yang lebih baik dibandingkan dengan menggunakan metode regresi Binomial Negatif dan regresi Poisson. Dari ketiga model yang dibandingkan, model GWNBR memiliki nilai AIC yang lebih kecil (171,21) dan nilai *McFadden's R-Squared* yang lebih besar (92,39%) dibandingkan dengan regresi Binomial Negatif dan regresi Poisson. Hal tersebut menunjukkan bahwa model GWNBR lebih baik digunakan dalam memodelkan kasus Pneumonia di Provinsi Sulawesi Tengah dibandingkan model regresi Poisson dan regresi Binomial Negatif.

5. Kesimpulan

Berdasarkan hasil analisis dan pembahasan didapatkan kesimpulan sebagai berikut:

1. Provinsi Sulawesi Tengah merupakan salah satu provinsi dengan tingkat penemuan kasus Pneumonia tertinggi di pulau Sulawesi dan menduduki posisi ke-8 sebagai provinsi dengan kasus Pneumonia tertinggi di Indonesia, dengan total kasus sebesar 6.273 balita dengan rata-rata sebanyak 355 kasus.
2. Salah satu model yang didapatkan dari 13 model yang terbentuk dengan metode GWNBR adalah Kota Palu, dengan model sebagai berikut

$$\boldsymbol{\hat{\mu}_{Palu} = \exp\left(-15,8069 - 0,0053X_1 + 0,0011X_2 + 0,0002X_3 - 0,0017X_5\right)}$$

3. Berdasarkan pemodelan GWNBR dengan fungsi pembobot kernel *Adaptive Bisquare* pada taraf signifikansi 5%, diperoleh hasil bahwa variabel prediktor yang berpengaruh signifikan terhadap jumlah kasus Pneumonia di tiap kabupaten/kota Provinsi Sulawesi Tengah Tahun 2021 yaitu variabel Persentase Pemberian ASI Eksklusif ($\boldsymbol{X_1}$), Persentase Imunisasi Dasar Lengkap ($\boldsymbol{X_2}$), Persentase Balita mendapat Vitamin A ($\boldsymbol{X_3}$), dan Persentase Cakupan Pelayanan Anak Balita ($\boldsymbol{X_5}$). Sedangkan variabel yang tidak berpengaruh signifikan yaitu variabel BBLR ($\boldsymbol{X_4}$).

References

Alwi, W., Sauddin, A., & Islamiyah, N. I. (2022). Pemodelan Generalized Poisson Regression Pada Faktor-Faktor Yang Mempengaruhi Kasus Pneumonia Pada Balita Di Provinsi Sulawesi Selatan 2018. *Jurnal Matematika dan Statistika serta Aplikasinya*, *10*(1), 9-14.

Ambarwati, P. C., Indahwati, I., & Aidi, M. N. (2020). Kajian simulasi overdispersi pada regresi poisson dan binomial negatif terboboti geografis untuk data balita gizi buruk. *Indonesian Journal of Statistics and Its Applications*, *4*(3), 484-497.

Anselin, L., & Bera, A. K. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In A. Ullah (Ed.), *Handbook of Applied Economic Statistics* (pp. 237-290). CRC Press.

Cahyandari, R. (2014). Pengujian Overdispersi pada Model Regresi Poisson. *Statistika*, 14(2), 69–76.

Cameron, A. C., & Trivedi, P. K. (1998). *Regression Analysis of Count Data* (P. Hammond & A. Holly (eds.); Issue 30). Cambridge University Press.

Damayanti CR, M., & Yanti, T. S. (2022). Regresi Poisson Invers Gaussian (PIG) untuk Pemodelan Jumlah Kasus Pneumonia pada Balita di Provinsi Jawa Tengah Tahun 2019. *Jurnal Riset Statistika*, 1(2), 143–151.

Darsyah, M. (2021). Pemodelan *Geographically Weighted Negative Binomial Regression* (GWNBR) Pada Kasus Malaria di Indonesia. *Jurnal Litbang Edusaintech*, 2(2), 1–15.

Dean, A., & Hocking, R. R. (1997). *Methods and Applications of Linear Models*. 2nd ed., Wiley & Sons, Canada.

Dinas Kesehatan Provinsi Sulawesi Tengah. (2021). Profil Kesehatan Provinsi Sulawesi Tengah. Profil Kesehatan Provinsi Sulawesi Tengah, 1–377.

Dinas Kesehatan Provinsi Sulawesi Tengah. (2022). *Profil Kesehatan Provinsi Sulawesi Tengah*. Diakses pada 23 September 2023 dari https://dinkes.sultengprov.go.id/wp-content/uploads/2023/06/Profil-KEsehatan-2022.pdf

Famoye, F., Wulu, J., & Singh, K. (2004). On The Generalize Poisson Regression Model with an Application to Accident Data. *Journal of Data Science 2*, 287-295.

Faridi, A., Budiasih, Suswandari, Ahmad, I., Furqan, M., & Susanti, E. (2017). *Evaluasi Metode Penghitungan Cakupan Imunisasi Dasar Lengkap di Indonesia*. 1–11.

Greene, W. (2008). Functional forms for the negative binomial model for *count* data. *Economics Letters*, *99*(3), 585–590.

Gierke, R., Wodi, A.P. and Kobayashi, M. (2021) 'Pinkbook: Pneumococcal Disease', *Pinkbook*, pp. 255–274.

Ismail, I. (2019). Pemodelan Jumlah Kematian Bayi Di Provinsi Jawa Barat Tahun 2017 Dengan Pendekatan *Geographically Weighted Negative Binomial Regression (GWNBR)* (Doctoral dissertation, Muhammadiyah University, Semarang).

Josefa, R., Sovia, R., & Mandala, E. P. (2019). Sistem Pakar Diagnosa Penyakit Pneumonia Pada Anak Menggunakan Metode Case Based Reasoning. *Sainteks*, 6, 868–872.

Kemenkes RI. (2020). *Pneumonia Pada Anak bisa Dicegah dan Diobati*. Di akses pada 22 Juli 2023 dari https://www.kemkes.go.id/id/rilis kesehatan/pneumonia-anak-dicegah-dan-diobati

Kemenkes RI. (2022). Profil Kesehatan Indonesia 2021. In *Pusdatin.Kemenkes.Go.Id.*

Kurniasari, M., & Ariastita, P. G. (2014). Faktor - Faktor yang Mempengaruhi Alih Fungsi Lahan Pertanian Sebagai Upaya Prediksi Perkembangan Lahan Pertaniandi Kabupaten Lamongan. *Jurnal Teknik Pomits*, *3*(2), 27–40.

Kurniawan, I. (2017). Model Regresi Poisson Terbaik Menggunakan Zero-Inflated Poisson (ZIP) dan Zero-Inflated Negative Binomial (ZINB*). Unnes Journal of Mathematics,* 5(1), 1-10.

Mar'ah, Z., Ahmar, A. S., & Rais, Z. (2023). Pemodelan Regresi Data Panel pada IPM di Sulawesi Selatan. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *5*(1), 23–27.

Pratama, W., & Wulandari, S. P. (2015). Pemetaan dan Pemodelan Jumlah Kasus Penyakit Tuberculosis (TBC) di Provinsi Jawa Barat dengan Pendekatan *Geographically Weighted Negative Binomial Regression*. *Jurnal Sains Dan Seni ITS,* 4(1), 37–42..

Purba, D., & Purba, M. (2022). Aplikasi Analisis Korelasi dan Regresi menggunakan Pearson Product Moment dan Simple Linear Regression. *Citra Sains Teknologi*, 1(2), 97–103.

Rahmadeni Feni Fatkhuli, R. J. (2019). Pemodelan Generalized Poisson Regression (GPR) Pada Kasus Kematian Neonatal Di Provinsi Riau. *Jurnal Sains Matematika Dan Statistika : Jurnal Hasil Penelitian Matematika, Statistika, Dan Aplikasinya*, 5(2), 43–50.

Rini, D. S. (2018). *Geographically Weighted Negative Binomial Regression* untuk Jumlah Kasus Demam Berdarah Dengue Kabupaten/Kota Provinsi Bengkulu. *PRISMA, Prosiding Seminar Nasional Matematika*, *1*, 736–744.

Safrida, N., D. Ispriyanti, & T. Widiharih. 2013. Aplikasi Model Regresi Poisson Tergeneralisasi Pada Kasus Angka Kematian Bayi di Jawa Tengah Tahun 2007. *Jurnal Gaussian*, 2(2): 361-368.

Saputri, V. A., & Purhadi, P. (2022). Pemodelan Faktor – Faktor yang Mempengaruhi Kasus Pneumonia pada Balita di Provinsi Jawa Barat dengan Metode Geographically Weighted Generalized Poisson Regression. *Inferensi*, 5(2), 91. https://doi.org/10.12962/j27213862.v5i2.12619

Savitri, N., Miranda, I., Sitorus, A., Luh, N., Andini, E., Husna, N. L., & Balita, K. (2022). Determinan jumlah kematian balita akibat pneumonia di Indonesia tahun 2019 dengan pendekatan generalized poisson regression. *Median*, 5(1), 40–51.

WHO. (2022). *Pneumonia in children*. World Health Organization: WHO. Diakses pada 23 Juli 2023 dari https://www.who.int/news-room/fact-sheets/detail/pneumonia.

Widyaningsih, Y., Arum, G. P., & Prawira, K. (2021). Aplikasi K-Fold Cross Validation Dalam Penentuan Model Regresi Binomial Negatif Terbaik. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *15*(2), 315–322.

Wirawan, I. M. (2016). Pendekatan komputasi numerik metode regresi pada penelitian yang mengamati suatu kecenderungan/trends terhadap Peningkatan Prestasi/Hasil Belajar. *Tekno*, 25(1), 1–14.

Yatnaningtyas, R., Latra, I. N., & Andari, S. (2016). Identifikasi Faktor-Faktor yang Mempengaruhi Pneumonia pada Balita di Surabaya Menggunakan *Geographically Weighted Negative Binomial Regression*. *Jurnal Sains Dan Seni ITS*, 5(2), D283–D288.

Xing, Y. *et al.* (2020). Vitamin A deficiency is associated with severe Mycoplasma pneumoniae pneumonia in children. *Annals of Translational Medicine*, 8(4), pp. 120–120