PEMODELAN *GEOGRAPHICALLY WEIGHTED ZERO INFLATED POISSON REGRESSION* PADA KASUS KEMATIAN DEMAM BERDARAH *DENGUE* DI SULAWESI SELATAN

SKRIPSI

ANDI UMMI MELIN AICHA

H051181020

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

PEMODELAN *GEOGRAPHICALLY WEIGHTED ZERO INFLATED POISSON REGRESSION* PADA KASUS KEMATIAN DEMAM BERDARAH *DENGUE* DI SULAWESI SELATAN

SKRIPSI

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Sains pada Program Stutdi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

ANDI UMMI MELIN AICHA

H051181020

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

Pemodelan *Geographically Weighted Zero Inflated Poisson Regression* pada Kasus Kematian Demam Berdarah *Dengue* di Sulawesi Selatan

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 21 Oktober 2022

Andi Ummi Melin Aicha

NIM H051181020

PEMODELAN *GEOGRAPHICALLY WEIGHTED ZERO INFLATED POISSON REGRESSION* PADA KASUS KEMATIAN DEMAM BERDARAH *DENGUE* DI SULAWESI SELATAN

Disetujui Oleh:

Pembimbing Utama	Pembimbing Pertama
Dr. Anna Islamiyati, S.Si., M.Si.	**Andi Kresna Jaya, S.Si., M.Si.**
NIP. 197708082005012002	**NIP. 197312282000031001**

Ketua Program Studi

Dr. Nurtiti Sunusi, S.Si., M.Si.

NIP. 19720117 199703 2 002

Pada 21 Oktober 2022

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama	:	Andi Ummi Melin Aicha
NIM	:	H051181020
Program Studi	:	Statistika
Judul Skripsi	:	Pemodelan *Geographically Weighted Zero Inflated Poisson Regression* pada Kasus Kematian Demam Berdarah *Dengue* di Sulawesi Selatan

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Dr. Anna Islamiyati, S.Si., M.Si. (.....)
2. Sekretaris : Andi Kresna Jaya, S.Si., M.Si. (.....)
3. Anggota : Dr. Nurtiti Sunusi, S.Si., M.Si. (.....)
4. Anggota : Siswanto, S.Si., M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 21 Oktober 2022

KATA PENGANTAR

Alhamdulillahi Rabbil'alamin, puji syukur atas berkat Allah SWT., atas limpahan rahmat, karunia, serta kekuatan kepada penulis sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Salam dan shalawat kepada Nabi Muhammad SAW., Nabi yang membawa cahaya kebenaran, membebaskan ummat manusia dari kebodohan dan kegelapan.

Skripsi dengan judul *"**Pemodelan Geographically Weighted Zero Inflated Poisson Regression Pada Kasus Kematian Demam Berdarah Dengue di Sulawesi Selatan**"* ini penulis hadirkan sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis menyadari bahwa skripsi ini tidak akan terwujud tanpa orang-orang yang memberikan bantuan, dukungan, bimbingan kepada penulis. Oleh karena itu dalam kesempatan ini penulis menghaturkan terima kasih dan rasa hormat yang setinggi-tingginya kepada kedua orang tua penulis Almarhum Ayahanda **Drs. Andi Muhammad Akis** dan Ibunda **Hj. Cahaya, S.Pd.** yang dengan penuh kasih sayang merawat, membesarkan penulis, tidak henti mengirimkan doa dan memberikan restu kepada penulis, dan kepada Kakak-kakakku **Andi Shari Aicha** dan **Andi Abdillah Aicha** serta keluarga besar penulis yang senantiasa memberikan doa, semangat, dukungan dan kasih sayang.

Selanjutnya ucapan terima kasih dan penghargaan yang sedalam-dalamnya, penulis sampaikan kepada:

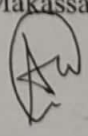
1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurtiti Sunusi, S.Si., M.Si.**, selaku Ketua Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.
4. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.** selaku Pembimbing Utama dan **Bapak Andi Kresna Jaya, S.Si., M.Si.** selaku Pembimbing Pertama dan Penasehat Akademik penulis yang telah rela dan sabar meluangkan waktu untuk

memberikan bantuan, arahan, bimbingan, tanggapan dan masukan yang berguna dalam penyelesaian skripsi ini.

5. **Ibu Dr. Nurtiti Sunusi, S.Si., M.Si.** dan **Bapak Siswanto, S.Si., M.Si.** selaku Tim Penguji yang telah rela dan sabar meluangkan waktu untuk memberikan kritikan yang membangun dalam penyempurnaan penyusunan skripsi ini.
6. Bapak dan Ibu Dosen Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin yang telah membagikan ilmunya kepada penulis dan segenap pegawai akademik yang selama ini selalu siap melajani segala urusan akademik penulis.
7. **Nurul Syafika, Isra' Riska Utami**, **Nurul Ikhsani**, **Riskayani**, dan **Rahma Alifah** selaku sahabat-sahabat penulis, terima kasih atas waktunya selama ini dengan saling mendengarkan keluh dan kesah masing-masing, selalu memberikan dukungan, semangat, bantuan, dan nasehat, serta selalu ada dalam suka dan duka.
8. **Viktor Liman, Musdalifah, Andi Sri Yulianti, Nurhidayah L, Sri Indriani Amil, Fiska Evryan, Anugrah, Ainun, dan Siti Aisyah** selaku teman-teman penulis yang telah memberikan dukungan, masukan, bantuan kepada penulis selama bersama-sama menjalani pendidikan di Prodi Statistika.
9. Teman-teman Statistika 2018 terima kasih atas ilmu, perjuangan, kerjasama, bantuan, dan dukungan serta kenangan indah hingga saat ini.
10. Seluruh pihak yang membantu penyelesaian tugas akhir ini, semoga menjadi pahala kebaikan bagi mereka pada hari kemudian kelak.

Penulis menyadari bahwa dalam penulisan skripsi ini masih banyak kesalahan dan kekurangan, sehingga penulis mengharapkan kritik dan saran yang membangun agar skripsi ini lebih baik dan bermanfaat.

Makassar, 21 Oktober 2022



Penulis

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama	: Andi Ummi Melin Aicha
NIM	: H051181020
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

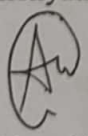
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-ekslusif** (***Non-exclusive Royalty- Free Right***) atas tugas akhir saya yang berjudul:

"Pemodelan *Geographically Weighted Zero Inflated Poisson Regression* pada Kasus Kematian Demam Berdarah *Dengue* di Sulawesi Selatan"

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.
Dibuat di Makassar pada tanggal, 21 Oktober 2022

Yang menyatakan



(Andi Ummi Melin Aicha)

ABSTRAK

Geographically Weighted Regression (GWR) merupakan salah satu metode yang dapat digunakan untuk menganalisis data spasial. Contoh bentuk data spasial diperoleh pada data penyebaran atau tingkat kematian yang dikarenakan sebuah penyakit di lokasi tertentu. Terdapat data spasial yang memuat banyak nilai nol pada data yang menyebabkan heterogenitas spasial dan overdispersi pada data, sehingga digunakan model *Geographically Weighted Zero Inflated Poisson Regression* (GWZIPR). Data yang dianalisis dalam penelitian ini adalah data jumlah kematian Demam Berdarah *Dengue* (DBD) di Sulawesi Selatan Tahun 2019 yang mengandung 66,67% nilai nol. Untuk memperoleh estimasi parameter model GWZIPR dilakukan dengan menggunakan metode estimasi *Maximum Likelihood* dan *Expectation Maximization*. Model GWZIPR menghasilkan variabel prediktor signifikan yang beragam antar kabupaten/kota di Sulawesi Selatan**.** Model GWZIPR terdiri dari model log dan model logit karena variabel respon (y_i) muncul dari dua keadaan yaitu Poisson *state* dan *zero state*. Pada model log Persentase jumlah sarana air minum yang memenuhi syarat kesehatan dan Kepadatan penduduk merupakan variabel yang berpengaruh terhadap jumlah kematian DBD di Sulawesi Selatan Tahun 2019 di sebagian besar kabupaten/kota di Sulawesi Selatan. Kondisi yang sama terjadi pada model logit, namun Persentase jumlah sarana air minum yang memenuhi syarat kesehatan berpengaruh pada semua kabupaten/kota di Sulawesi Selatan.

Kata Kunci: *Geographically Weighted Zero Inflated Poisson Regression*, *Maximum Likelihood*, *Expectation Maximization*, Demam Berdarah *Dengue*.

ABSTRACK

Geographically Weighted Regression (GWR) is a method can used to analyze spatial data. An example of the form of spatial data is obtained from data on the spread or rate of death due to a disease in a certain location. There is spatial data that contains many zero values in the data which causes spatial heterogeneity and overdispersion in the data, so the Geographically Weighted Zero Inflated Poisson Regression (GWZIPR) model's is used. The data analyzed in this study is data on the number of deaths from Dengue Hemorrhagic Fever (DHF) in South Sulawesi in 2019 which contains 66.67% of zero values. To obtain the estimated parameters of the GWZIPR model's, the Maximum Likelihood Estimation and Expectation Maximization methods were used. The GWZIPR model produces significant predictor variables that vary between districts/cities in South Sulawesi. The GWZIPR model consists of a log model and a logit model because the response variable (y_i) arises from two states, namely Poisson state and zero state. In the log model, the percentage of drinking water facilities satisfying health requirements and population density are variables that affect the number of deaths due to dengue fever in South Sulawesi in 2019 in most districts/cities in South Sulawesi. The same condition occurs in the logit model, but the percentage of drinking water facilities that meet health requirements affects all districts/cities in South Sulawesi.

Keywords: *Geographically Weighted Zero Inflated Poisson Regression, Maximum Likelihood Estimation, Expectation Maximization, Dengue Hemorrhagic Fever.*

DAFTAR ISI

HALAMAN SAMPUL i
HALAMAN JUDUL……..... ii
LEMBAR PERNYATAAN KEOTENTIKAN..... iii
HALAMAN PERSETUJUAN PEMBIMBING iv
HALAMAN PENGESAHAN..... v
KATA PENGANTAR vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI..... viii
ABSTRAK……..... ix
ABSTRACK x
DAFTAR ISI..... xi
DAFTAR TABEL..... xiii
DAFTAR LAMPIRAN..... xiv
BAB I PENDAHULUAN..... 1
1.1. Latar Belakang 1
1.2. Rumusan Masalah..... 3
1.3. Batasan Masalah..... 3
1.4. Tujuan Penelitian 3
1.5. Manfaat Penelitian 3
BAB II TINJAUAN PUSTAKA..... 4
2.1. Regresi Poisson..... 4
2.2. *Maximum Likelihood Estimation* (MLE) 4
2.3. Asumsi Multikolinieritas..... 5
2.4. Asumsi Heterogenitas Spasial..... 5
2.5. *Bandwidth* Optimum..... 6
2.6. Pembobotan Model GWR..... 6
2.7. Model *Geographically Weighted Poisson Regression* (GWPR) 7
2.8. Overdispersi 8
2.9. Regresi *Zero Inflated Poisson* (ZIP) 8
2.10. *Geographically Weighted Zero Inflated Poisson Regression* 11
(GWZIPR)..... 11
2.10.1 Pengujian Serentak Parameter Model GWZIPR 12

2.10.2 Pengujian Parsial Parameter Model GWZIPR 13
2.11. Demam Berdarah *Dengue* (DBD) 13
BAB III METODE PENELITIAN 15
3.1. Sumber Data dan Variabel Penelitian 15
3.2. Metode Analisis Penelitian 16
BAB IV HASIL DAN PEMBAHASAN 18
4.1. Model *Geographically Weighted Zero Inflated Poisson Regression* (GWZIPR) 18
4.2. Aplikasi pada data jumlah kasus kematian akibat penyakit Demam Berdarah *Dengue* di Sulawesi Selatan Tahun 2019 28
4.2.1 Deskriptif 28
4.2.2 Uji Multikolinieritas 29
4.2.3 Uji Heterogenitas Spasial 30
4.2.4 Overdispersi 30
4.2.5 *Bandwidth* Optimum 30
4.2.6 Pembobot Adaptive Bisquare Kernel 32
4.2.7 Uji Serentak Parameter Model GWZIPR 32
4.2.8 Uji Parsial Parameter Model GWZIPR 33
4.2.9 Pemodelan GWZIPR dengan Pembobot *Adaptive Bisquare Kernel* . 40
BAB V PENUTUP 42
5.1. Kesimpulan 42
5.2. Saran 43
DAFTAR PUSTAKA 44
LAMPIRAN 46

DAFTAR TABEL

Tabel 4.1 Statistika Deskriptif Variabel 28
Tabel 4.2 Nilai VIF Variabel Prediktor 29
Tabel 4.3 *Bandwidth* Optimum untuk setiap kabupaten/kota di Sulawesi Selatan 31
Tabel 4.4 Variabel yang Signifikan Di Sulawesi Selatan pada Model GWZIPR dengan Pembobot *Adaptive Bisquare Kernel* 34
Tabel 4.5 Hasil Pengelompokan Kabupaten/Kota Berdasarkan variabel Signifikan pada Model Log 35
Tabel 4.6 Hasil Pengelompokan Kabupaten/Kota Berdasarkan variabel Signifikan pada Model Logit 36
Tabel 4.7 Pengujian Parameter Model GWZIPR di Kabupaten Gowa dengan Pembobot *Adaptive Bisquare Kernel* 38
Tabel 4.8 Pengujian Parameter Model GWZIPR Di Kabupaten Maros dengan Pembobot *Adaptive Bisquare Kernel* 39
Tabel 4.9 Pengujian Parameter Model GWZIPR Di Kabupaten Enrekang dengan Pembobot *Adaptive Bisquare Kernel* 40

DAFTAR LAMPIRAN

Lampiran 1 Data dan Koordinat Kabupaten/Kota di Sulawesi Selatan 47
Lampiran 2 *Output* jarak *Euclidean* antar Kabupaten/kota..... 48
Lampiran 3 *Output* Pembobot *Adaptive Bisquare Kernel* 51
Lampiran 4 *Output* Nilai Hasil Parameter Awal $\boldsymbol{\beta}$ dan $\boldsymbol{\gamma}$ 54
Lampiran 5 *Output* Penduga Parameter Model GWZIPR dengan Pembobot *Adaptive Bisquare Kernel* 55
Lampiran 6 *Output* Statistik Uji *W* Penduga Parameter Model GWZIPR dengan Pembobot *Adaptive Bisquare Kernel* 57
Lampiran 7 Peta Hasil Pengelompokan Kabupaten/Kota di Sulawesi Selatan Berdasarkan Variabel Signifikan pada Model Log 59
Lampiran 8 Peta Hasil Pengelompokan Kabupaten/Kota di Sulawesi Selatan Berdasarkan Variabel Signifikan pada Model Logit..... 60
Lampiran 9 Pengujian Parameter Model GWZIPR di Sulawesi Selatan dengan Pembobot *Adaptive Bisquare Kernel* 61
Lampiran 10 *Syntax* R Studio 73

BAB I
PENDAHULUAN

1.1. Latar Belakang

Model regresi linier secara umum bersifat global dengan estimasi parameternya bernilai sama untuk semua lokasi pengamatan. Menurut Fotheringham dkk. (2002) jika keragaman antar wilayah lokal tidak ada atau hanya sedikit maka model persamaan global akan memberikan informasi yang akurat untuk wilayah lokal. Pada kenyataannya, data antar pengamatan sangat sulit dianalisis menggunakan regresi linier karena antar lokasi memiliki kondisi berbeda yang diduga akan memberikan efek keragaman spasial. Seperti yang dikemukakan Anselin dan Getis (1992) yang mengatakan pengaruh spasial antar lokasi dapat disebabkan oleh keragaman spasial, yaitu satu peubah penjelas yang sama memberikan respon yang berbeda antara satu lokasi dengan lokasi lainnya. Jika uji keragaman spasial diabaikan dalam model regresi maka hasil yang diperoleh kurang sesuai (Pratiwi dkk., 2019).

Data yang berorientasi pada geografis atau yang memiliki gambaran suatu wilayah disebut dengan data spasial. Geografis memiliki arti bahwa segala sesuatu saling berpengaruh satu sama lain tetapi jarak yang lebih dekat akan memberikan pengaruh yang lebih besar. Data spasial biasanya digunakan dalam menangani permasalahan bencana, kemiskinan, sebaran penyakit (epidemik), kematian bayi yang dilahirkan, dan lain sebagainya. Regresi spasial merupakan metode yang dapat digunakan untuk menganalisis hubungan antar satu variabel atau lebih yang melibatkan efek spasial pada beberapa wilayah atau lokasi yang menjadi pusat pengamatan. Pengembangan metode untuk menganalisis data spasial yaitu *Geographically Weighted Regression* (GWR) yang dikembangkan oleh Fotheringham dkk. (2002). Hasil dari model GWR yaitu model regresi yang nilai-nilai parameternya berlaku hanya pada setiap lokasi pengamatan dan berbeda dengan lokasi yang lainnya(Ismah dkk., 2020).

Data berupa penyebaran penyakit atau tingkat kematian yang dikarenakan penyakit tertentu sering diperoleh banyak nilai nol pada data (*excess zeros)*. Banyaknya nilai nol pada hasil observasi dapat menjadi salah satu penyebab terjadinya overdispersi. Overdispersi merupakan data dengan nilai varians dari

variabel respon lebih besar dari nilai *mean* variabel respon. Dalam regresi Poisson, semakin banyak hasil observasi yang bernilai nol akan melampaui nilai prediksi. Salah satu metode yang dapat digunakan untuk mengatasi masalah tersebut yaitu model *Zero Inflated Poisson Regression* (ZIPR). *Maximum Likelihood Estimation* (MLE) merupakan metode yang dapat digunakan untuk mengestimasi beberapa parameter pada model regresi ZIPR (Lambert, 1992).

Lambert (1992) menyatakan ZIPR merupakan model campuran untuk data diskrit dengan banyak peristiwa yang bernilai nol. Akan tetapi, terdapat kasus yang memiliki banyak observasi bernilai nol yang mengandung heterogenitas seperti pada data spasial. Ismah dkk. (2020) menjelaskan bahwa model ZIPR kurang tepat digunakan pada kondisi data yang melibatkan efek spasial sehingga dapat digunakan pengembangan dari model GWR dan ZIP yaitu model *Geographically Weighted Zero Inflated Poisson Regression* (GWZIPR) untuk mengatasi kasus yang terdapat heterogenitas spasial dan kondisi data *excess zeros*. Penelitian terkait dengan GWZIPR telah dilakukan sebelumnya oleh Nusantara dan Purhadi (2015) dengan hasil penelitian bahwa terdapat perbedaan pada pengujian kesamaan model GWZIPR dan ZIP, pengujian serentak parameter model GWZIPR memberikan kesimpulan bahwa model GWZIPR layak digunakan pada kasus Tetanus Neonatorum.

Pada pendugaan parameter model GWZIPR digunakan matriks pembobot yang besarnya bergantung dengan kedekatan antara wilayah pengamatan ke-i dan ke-j. Fungsi pembobot tersebut bertujuan agar setiap wilayah pengamatan memiliki pendugaan yang berbeda-beda. Amaliana dkk. (2018) membandingkan fungsi pembobot *Fixed Bisquare Kernel* dengan *Fixed Gaussian Kernel* pada model GWZIPR pada data filariasis, hasil yang diperoleh *Fixed Gaussian Kernel* lebih baik daripada *Fixed Bisquare Kernel.* Faricha (2016) membandingkan fungsi pembobot *Adaptive Gaussian Kernel* dengan *Adaptive Bisquare Kernel* pada model GWZIPR, hasil yang diperoleh dari kedua pembobot sama baiknya dalam memodelkan jumlah kematian ibu di usia dibawah 20 tahun di Jawa Timur tahun 2013. Berdasarkan uraian tersebut, penulis mengusulkan kajian penelitian tentang pemodelan GWZIPR dengan pembobot *Adaptive Bisquare Kernel* pada kasus kematian demam berdarah *dengue* (DBD) di Sulawesi Selatan tahun 2019.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana estimasi parameter model GWZIPR dengan pembobot *Adaptive Bisquare Kernel* yang bersesuaian dengan data?
2. Bagaimana model kasus kematian akibat penyakit DBD di Sulawesi Selatan tahun 2019 berdasarkan metode GWZIPR?

1.3. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Jarak yang digunakan dalam penentuan pembobot yaitu jarak *Euclidean.*
2. Pembobot yang digunakan dalam pendugaan parameter yaitu *Adaptive Bisquare Kernel.*
3. Pendugaan parameter model menggunakan metode MLE.
4. Data penelitian yang digunakan yaitu kasus kematian akibat penyakit DBD di Sulawesi Selatan tahun 2019.

1.4. Tujuan Penelitian

Berdasarkan rumuan masalah di atas, maka tujuan penelitian adalah sebagai berikut:

1. Untuk memperoleh estimator parameter model GWZIPR dengan pembobot *Adaptive Bisquare Kernel* yang bersesuaian dengan data.
2. Untuk mendapatkan model kasus kematian akibat penyakit DBD di Sulawesi Selatan tahun 2019 melalui metode GWZIPR.

1.5. Manfaat Penelitian

Peneltian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Sebagai bahan informasi, masukan, dan pertimbangan masyarakat ataupun pemerintah dalam mengambil keputusan/menetapkan kebijakan yang dapat menekan angka kasus kematian akibat penyakit DBD di Sulawesi Selatan.
2. Dapat digunakan sebagai bahan refrensi bagi peneliti lain yang akan mengadakan penelitian dalam bidang yang sama yaitu tentang kematian akibat penyakit DBD maupun tentang metode GWZIPR.

BAB II
TINJAUAN PUSTAKA

2.1. Regresi Poisson

Regresi Poisson merupakan suatu bentuk analisis regresi yang dapat digunakan untuk memodelkan banyaknya kejadian yang terjadi dalam suatu periode waktu dan/atau wilayah tertentu, dengan kata lain data yang berbentuk jumlah yang dilambangkan dengan *Y*. Adeliana dan Harini (2017) menjelaskan bahwa dalam Regresi Poisson, variabel random *Y* berdistribusi Poisson apabila fungsi peluangnya diberikan sebagai berikut:

$$f_Y(y) = f_Y(y;\mu)\begin{cases}\frac{\mu^y e^{-\mu}}{y!} & y = 0,1,2\ldots \\ 0 & lainnya\end{cases}$$

dengan parameter $\mu > 0$.

Myers (1990) menjelaskan model regresi Poisson adalah $y_i = \mu_i + \varepsilon_i$. Model regresi Poisson merupakan salah satu model *Generalized Linier Model* dengan fungsi penghubung *log* yaitu $\ln(\mu_i) = \sum_{k=0}^{p} x_{ik}\beta_k$. Sehingga model regresi Poisson adalah

$$\ln(\mu_i) = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ip}\beta_p$$

$$\mu_i = \exp\left(\boldsymbol{x_i^T\beta}\right)$$

diketahui:

i : 1,2,...,*n* ; *n* = banyaknya lokasi pengamatan

k : 1,2,...,*p* ; *p* = banyaknya variabel independen

y_i : variabel dependen pada lokasi pengamatan ke-*i*

μ_i : rata-rata variabel dependen pada lokasi pengamatan ke-*i* yang merupakan ekspektasi dari *Y* dengan *Y*~Poisson (μ)

$\boldsymbol{x_i}$: vektor variabel independen pada lokasi pengamatan ke-*i*

$\boldsymbol{\beta}$: vektor parameter regresi Poisson berukuran (*p* + 1) x 1

2.2. *Maximum Likelihood Estimation* (MLE)

Kusuma dkk. (2013) menjelaskan bahwa fungsi densitas bersama $f(x_1 \ldots x_n;\theta)$ dari variabel random $x_1, x_2, \ldots, x_n$ dinamakan fungsi *likelihood*.

Misal terdapat n pengamatan $x_1, x_2, \dots, x_n$ yang masing-masing mempunyai fungsi kepadatan peluang $f(x_i; \theta)$, fungsi *likelihood* suatu fungsi dari θ adalah:

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) \dots f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

Selanjutnya menentukan nilai maksimum dari logaritma natural fungsi *likelihood* atau disebut sebagai fungsi *log-likelihood*. Fungsi *log-likelihood* merupakan fungsi kepadatan bersama yang diubah menjadi bentuk logaritma. Fungsi *log-likelihood* dapat ditulis dalam bentuk:

$$l(\theta) = \ln L(\theta)$$

2.3. Asumsi Multikolinieritas

Asumsi Multikolinieritas dapat dideteksi pada suatu data dengan mengetahui nilai *Varian Inflation Factor* (VIF). Apabila salah satu nilai VIF melebihi angka 10 maka koefisien regresi merupakan penduga yang tidak baik dikarenakan terdapat multikolinieritas pada data. Persamaan untuk mencari nilai VIF dituliskan sebagai berikut (Praditia dkk., 2020):

$$VIF_k = \frac{1}{1 - R_k^2}$$

2.4. Asumsi Heterogenitas Spasial

Heterogenitas spasial terjadi karena adanya perbedaan karakteristik dari pengamatan di lokasi yang satu dengan pengamatan di lokasi yang lain. Heterogenitas spasial dapat diidentifikasi menggunakan pengujian *Breusch Pagan* dengan hipotesis sebagai berikut:

H_0 : $\sigma^2(u_1, v_1) = \cdots = \sigma^2(u_n, v_n) = \sigma^2$

H_1 : paling sedikit terdapat satu $\sigma^2(u_i, v_i) \neq \sigma^2$

Statistik uji *BP* adalah:

$$BP = \frac{1}{2}\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{Z}(\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{Z})^{-\boldsymbol{1}}\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{f} \sim \chi^2_{(p)} \qquad (2.1)$$

Dengan vektor $\boldsymbol{f}$ berukuran n x 1 adalah

$$\boldsymbol{f} = \left[\left(\frac{e_1^2}{\sigma^2}\right) - 1, \dots, \left(\frac{e_n^2}{\sigma^2}\right) - 1\right]^T$$

diketahui:

$\boldsymbol{Z}$: matriks berdimensi n x (p + 1) yang berisi variabel independen yang telah distandarisasikan untuk setiap pengamatan

σ^2 : ragam residual (e_i) dari model regresi Poisson

e_i : residual untuk pengamatan ke-i dari model regresi Poisson

Kriteria penolakan yaitu jika nilai $BP > \chi^2_{(\alpha;p)}$ atau $p - value < \alpha$ maka tolak H_0, artinya terdapat heterogenitas spasial pada data (Diastina dkk., 2019).

2.5. *Bandwidth* Optimum

Menurut Fadlilah dkk. (2019) *Bandwidth* dapat dianalogikan sebagai radius dari suatu lingkaran, sehingga sebuah titik yang terdapat dalam radius lingkaran tersebut masih dianggap memiliki pengaruh. Metode *Cross Validation* (CV) dapat digunakan untuk memperoleh nilai *bandwidth* yang optimum, penentuan nilai *bandwidth* optimum diperoleh dari nilai CV yang minimum. Persamaan CV sebagai berikut:

$$CV = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(h)]^2 \qquad (2.2)$$

diketahui:

$\hat{y}_{\neq i}$: nilai penduga lokasi pengamatan ke-i yang diperoleh tanpa melibatkan lokasi pengamatan ke-i

2.6. Pembobotan Model GWR

Pembobot digunakan untuk memberikan penekanan yang berbeda untuk observasi yang berbeda dalam menghasilkan penduga parameter. Menurut Leung, Mei dan Zhang (2000) penduga parameter di suatu lokasi ke-i akan lebih dipengaruhi oleh titik-titik yang dekat lokasi daripada titik-titik yang lebih jauh. Lokasi yang berdekatan menunjukkan hubungan kemiripan dan sebaliknya lokasi yang berjauhan akan menunjukkan keragaman spasial. Keragaman spasial antara lokasi pengamatan satu dengan yang lainnya ditunjukkan dengan adanya matriks pembobot $\boldsymbol{W}(u_i, v_i)$ berukuran n x n.

$$\boldsymbol{W}(u_i, v_i) = diag\left[w_{1j}(u_i, v_i) + w_{2j}(u_i, v_i) + \cdots + w_{nj}(u_i, v_i)\right]$$

Sebelum matriks pembobot $\boldsymbol{W}(u_i, v_i)$ ditentukan, terlebih dahulu menentukan fungsi pembobot yang akan digunakan dan juga ditentukan d_{ij} yang merupakan jarak lokasi ke-i dengan lokasi ke-j menggunakan jarak *Euclidean* dengan persamaan berikut:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \tag{2.3}$$

Penentuan pembobot $w_{ij}(u_i, v_i)$ menggunakan fungsi kernel yang memberikan pembobot sesuai *bandwidth* optimum pada masing-masing lokasi pengamatan. Model *adaptive* kernel dapat menyesuaikan dengan kondisi titik pengamatan sehingga nilai *bandwidth* akan berbeda-beda untuk setiap lokasi pengamatan.

Adapun jenis fungsi pembobot *Adaptive Kernel* yang digunakan yaitu *Adaptive Bisquare Kernel* dengan persamaan sebagai berikut:

$$w_{ij}(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h_i}\right)^2\right)^2 & d_{ij} \leq h_i \\ 0 & d_{ij} > h_i \end{cases} \tag{2.4}$$

Diketahui h_i adalah nilai *bandwidth* pada lokasi pengamatan ke-*i* (Faricha, 2016).

2.7. Model *Geographically Weighted Poisson Regression* (GWPR)

Qomariyah dkk. (2018) menjelaskan bahwa GWPR merupakan pengembangan metode statistika dari regresi Poisson, namun yang membedakan adalah metode ini memperhatikan pembobot berupa letak lintang dan bujur dari titik-titik pengamatan yang diamati. Model GWPR menghasilkan penaksir parameter model yang bersifat lokal untuk setiap titik pengamatan (Fotheringham dkk., 2002). Berikut ini model GWPR dengan u_i sebagai koordinat lintang dan v_i sebagai koordinat bujur yang digunakan untuk menentukan pembobot penaksiran parameter nantinya.

$$y_i \sim \text{Poisson}(\mu_i)$$

Model regresi GWPR adalah :

$$\hat{\mu}_i = \exp\left(\beta_0(u_i, v_i) + \sum_{k=1}^{p} x_{ik}\beta_k(u_i, v_i)\right)$$

$$\hat{\mu}_i = \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}(\boldsymbol{u}_i, \boldsymbol{v}_i)\right)$$

diketahui:

$\beta_0(u_i, v_i)$: koefisien intersep model GWPR

x_{ik} : nilai variabel independen ke-*k* pada setiap lokasi pengamatan ke-*i*

$\hat{\mu}_i$: penduga rata-rata variabel dependen pada lokasi pengamatan ke-*i*

$$\boldsymbol{\beta}(\boldsymbol{u_i}, \boldsymbol{v_i}) = \left(\beta_0(u_i, v_i), \beta_1(u_i, v_i), \dots, \beta_p(u_i, v_i)\right)^T$$

$$\boldsymbol{x_i^T} = (1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ip})$$

2.8. Overdispersi

Fatima dkk. (2016) menjelaskan bahwa regresi Poisson dikatakan mengandung overdispersi apabila nilai variansi lebih besar daripada *mean*. Jika pada data diskrit terjadi overdispersi namun tetap digunakan regresi poisson maka dugaan dari parameter koefisien regresinya tetap konsisten tetapi tidak efisien. Hal ini berdampak pada nilai *standar error* yang menjadi *under estimate,* sehingga kesimpulannya menjadi tidak valid. Overdispersi dapat dituliskan Var (*Y*) > *E* (*Y*). Cara yang dapat digunakan untuk mendeteksi overdispersi yaitu dengan melihat nilai statistik *deviance* dan *Pearson chi-square* yang dibagi dengan derajat bebasnya. Jika kedua nilai lebih dari 1 maka terjadi overdispersi pada data.

2.9. Regresi *Zero Inflated Poisson* (ZIP)

Jansakul dkk. (2001) *dalam* Adeliana dan Harini (2017) menyebutkan salah satu penyebab terjadinya overdispersi adalah lebih banyak observasi bernilai nol daripada yang ditaksir untuk model regresi Poisson. Salah satu metode analisis yang cocok digunakan untuk masalah tersebut adalah model analisis regresi ZIP.

Jika y_i adalah variabel random independen yang mempunyai distribusi ZIP nilai nol pada observasi diduga muncul dalam dua cara yang sasuai untuk keadaan (*state*) yang tepisah. Keadaan pertama disebut *zero state* terjadi dengan probabilitas ω_i dan menghasilkan hanya observasi bernilai nol, sementara keadaan kedua disebut Poisson *state* terjadi dengan probabilitas (1 - ω_i) dan berdistribusi Poisson dengan *mean* μ_i. ω_i dapat dikatakan sebagai peluang tidak terjadinya suatu kasus atau peluang variabel y bernilai 0.

Untuk setiap pengamatan Y_i yang saling bebas *i = 1,2,..,n* dan

$$Y_i \sim \begin{cases} 0 & \text{dengan peluang } \omega_i \\ \text{Possion } (\mu_i) & \text{dengan peluang } (1-\omega_i) \end{cases}$$

maka:

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu_i} & y_i = 0 \\ (1-\omega_i)\frac{e^{-\mu_i}\mu_i{}^{y_i}}{y_i!} & y_i = 1,2,.. \end{cases} \qquad (2.5)$$

Lambert (1992) menyarankan model gabungan untuk μ dan ω yaitu:

$$\ln(\mu_i) = X_i^T\beta \text{ dan logit } (\omega_i) = \ln\left(\frac{\omega_i}{1-\omega_i}\right) = X_i^T\gamma \qquad (2.6)$$

Diketahui X_i^T adalah matriks variabel prediktor, β dan γ adalah vektor parameter yang akan ditaksir, dan ω_i adalah probabilitas observasi bernilai nol.

Taufan dkk. (2012) *dalam* Kurniawan (2017) menyatakan estimasi parameter regresi ZIP menggunakan metode MLE. Metode ini biasanya digunakan untuk menaksir parameter suatu model yang diketahui fungsi densitasnya. Berdasarkan Persamaan (2.6) didapat:

$$\mu_i = \exp(x_i^T\beta), \omega_i = \frac{\exp(x_i^T\gamma)}{1+\exp(x_i^T\gamma)} \text{dan } (1-\omega_i) = \frac{1}{1+\exp(x_i^T\gamma)} \qquad (2.7)$$

Selanjutnya persamaan (2.7) disubtitusikan ke persamaan (2.5) didapat:

$$P(\mathrm{Y}_i = \mathrm{y}_i) = \begin{cases} \frac{\exp(x_i^T\gamma)+(\exp(-\exp(x_i^T\beta)))}{1+\exp(x_i^T\gamma)} & \mathrm{y}_i = 0 \\ \frac{\frac{1}{1+\exp(x_i^T\gamma)}(\exp(-\exp(x_i^T\beta)))(\exp(x_i^T\beta))^{\mathrm{y}_i}}{\mathrm{y}_i!} & \mathrm{y}_i > 0 \end{cases}$$

Sehingga fungsi *likelihood* diperoleh dengan mengalikan fungsi probabilitas dari Y_i yaitu

$$L(\beta,\gamma|\mathrm{y}_i) = \begin{cases} \prod_{i=1}^n \frac{\exp(x_i^T\gamma)+(\exp(-\exp(x_i^T\beta)))}{1+\exp(x_i^T\gamma)} & \mathrm{y}_i = 0 \\ \frac{\prod_{i=1}^n \frac{1}{1+\exp(x_i^T\gamma)}(\exp(-\exp(x_i^T\beta)))(\exp(x_i^T\beta))^{\mathrm{y}_i}}{\prod_{i=1}^n \mathrm{y}_i!} & \mathrm{y}_i > 0 \end{cases}$$

Dibentuk fungsi *ln-likelihood* sebagai berikut:

$$\ln L(\beta,\gamma|\mathrm{y}_i) = \begin{cases} \sum_{i=1}^n \ln\left(\exp(x_i^T\gamma) + (\exp(-\exp(x_i^T\beta)))\right) - \sum_{i=1}^n \ln\left(1+\exp(x_i^T\gamma)\right) & \mathrm{y}_i = 0 \\ \sum_{i=1}^n \left(\mathrm{y}_i x_i^T\beta - \exp(x_i^T\beta)\right) - \sum_{i=1}^n \ln\left(1+\exp(x_i^T\gamma)\right) - \sum_{i=1}^n \ln(\mathrm{y}_i!) & \mathrm{y}_i > 0 \end{cases} \qquad (2.8)$$

Penjumlahan fungsi *ln-likelihood* pada Persamaan (2.8) akan menyulitkan perhitungan karena tidak diketahui nilai nol mana yang berasal *zero state* dan mana yang berasal dari Poisson *state* sehingga fungsi *ln-likelihood* ini tidak dapat diselesaikan dengan numerik biasa. Oleh karena itu, Persamaan (2.8) disebut juga *incomplete data likelihood.*

Hall dan Shen (2009) menjelaskan bahwa untuk memaksimalkan fungsi ln-*likelihood* pada Persamaan (2.8) digunakan algoritma EM (*Expectation Maximization*) yang merupakan salah satu metode optimasi yang banyak digunakan sebagai alternatif dalam memaksimumkan fungsi *likelihood* yang

mengandung data hilang. Misalkan setiap variabel Y_i berkaitan dengan variabel indikator Z_i yaitu:

$$Z_i = \begin{cases} 1, \text{jika } y_i \text{berasal dari } zero\ state \\ 0, \text{jika } y_i \text{berasal dari Poisson } state \end{cases} \quad (2.9)$$

Langkah-langkah estimasi parameter dengan algoritma EM sebagai berikut:

1. Menentukan distribusi dari variabel Z_i:

Berdasarkan Persamaan (2.5) dan (2.9), maka:

$$P(Z_i = 1) = \omega_i$$

$$P(Z_i = 0) = P(Y_i \sim \text{Poisson } (\mu)) = 1 - \omega_i$$

Sehingga $Z_i \sim$ Bernoulli $(\omega_i), E(Z_i) = \omega_i, Var(Z_i) = \omega_i(1 - \omega_i)$

2. Membentuk distribusi gabungan antara y_i dan z_i yaitu

$$f(y_i, z_i|\omega_i, \mu_i) = f(z_i)f(y_i|z_i) = (1 - \omega_i)^{(1-z_i)}(\omega_i)^{(z_i)} \left(\frac{\exp(-(\mu_i)\mu_i{}^{(y_i)}}{y_i!}\right)^{(1-z_i)} \quad (2.10)$$

Kemudian dengan mensubstitusikan Persamaan (2.7) ke Persamaan (2.10) hingga diperoleh persamaan *ln-likelihood:*

$$\ln L(\beta, \gamma|y, z) = \sum_{i=1}^{n}(z_i x_i^T \gamma - \ln(1 + \exp(x_i^T \gamma))) + \sum_{i=1}^{n}(1 - z_i)(y_i x_i^T \beta - \exp(x_i^T \beta)) - \sum_{i=1}^{n}(1 - z_i)\ln(y_i!) \quad (2.11)$$

Persamaan (2.11) biasa disebut *complete data likelihood.* Persamaan ini yang akan dimaksimumkan menggunakan algoritma EM, parameter β dan γ dapat diestimasi secara terpisah dengan menuliskan Persamaan (2.11) menjadi:

$$\ln L(\beta, \gamma, y, z) = \ln L(\beta, y, z) + \ln L(\gamma, y, z) - \sum_{i=1}^{n}(1 - z_i)\ln(y_i!)$$

Dengan

$$\ln L(\gamma, y, z) = \sum_{i=1}^{n}(z_i x_i^T \gamma - \ln(1 + \exp(x_i^T \gamma))) \quad (2.12)$$

$$\ln L(\beta, y, z) = \sum_{i=1}^{n}(1 - z_i)(y_i x_i^T \beta - \exp(x_i^T \beta)) \quad (2.13)$$

Sedangkan $\sum_{i=1}^{n}(1 - z_i)\ln(y_i!)$ dapat diabaikan karena tidak mengandung vektor parameter β dan γ.

3. Tahap Ekspektasi

Mengganti variabel z_i dengan $z_i{}^{(m)}$ yang merupakan ekspektasi dari z_i

$$z_i^{(m)} = E\left(z_i \middle| y_i, \gamma^{(m)}, \beta^{(m)}\right) = \begin{cases} P\left(z_i = 1 \middle| y_i = 0, \gamma^{(m)}, \beta^{(m)}\right) & jika\ y_i = 0 \\ 0 & jika\ y_i > 0 \end{cases}$$

$$P\left(z_i = 1 \middle| y_i = 0, \gamma^{(m)}, \beta^{(m)}\right) = \frac{1}{1 + \exp(-x_i^T \gamma^{(m)} - \exp(x_i^T \beta^{(m)}))}$$

Sehingga Persamaan (2.12) dan (2.13) menjadi

$$\ln L\left(\gamma^{(m)}, y, z^{(m)}\right) = \sum_{i=1}^{n}\left(z_i^{(m)} x_i^T \gamma^{(m)} - \ln(1 + \exp(x_i^T \gamma^{(m)}))\right) \quad (2.14)$$

$$\ln L\left(\beta^{(m)}, y, z^{(m)}\right) = \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\left(y_i x_i^T \beta^{(m)} - \exp(x_i^T \beta^{(m)})\right) \quad (2.15)$$

4. Tahap Maksimalisasi

Memaksimalkan β dan γ pada Persamaan (2.14) dan (2.15) dengan menghitung $\beta^{(m)}$ dan $\gamma^{(m)}$ dengan metode Newton-Raphson. Misalkan $\beta^{(m-1)}$ dan $\gamma^{(m-1)}$ adalah aproksimasi metode maksimum *likelihood* untuk mengestimasi $\hat{\beta}$ dan $\hat{\gamma}$. Dengan menggunakan metode Newton-Raphson maka:

$$\hat{\beta}^{(m)} = \hat{\beta}^{(m-1)} - \left(H^{(m-1)}\hat{\beta}^{(m-1)}\right)^{-1} U^{(m-1)}\hat{\beta}^{(m-1)}$$

dan

$$\hat{\gamma}^{(m)} = \hat{\gamma}^{(m-1)} - \left(H^{(m-1)}\hat{\gamma}^{(m-1)}\right)^{-1} U^{(m-1)}\hat{\gamma}^{(m-1)}$$

Untuk m = 1,2,...,n dengan $\beta^{(0)}$ dan $\gamma^{(0)}$ adalah nilai awal, H adalah turunan kedua dari $\ln L\left(\beta^{(m)}, y, z^{(m)}\right)$ dan $\ln L\left(\gamma^{(m)}, y, z^{(m)}\right)$, U adalah turunan pertama dari $\ln L\left(\beta^{(m)}, y, z^{(m)}\right)$ dan $\ln L\left(\gamma^{(m)}, y, z^{(m)}\right)$.

5. Pengulangan Tahap ke-3 dan ke-4

Ganti $\beta^{(m-1)}$ dan $\gamma^{(m-1)}$ dengan $\hat{\beta}^{(m)}$ dan $\hat{\gamma}^{(m)}$ pada iterasi selanjutnya, tahap ke-3 dan ke-4 ini dilakukan berulang-ulang sampai diperoleh penaksir parameter yang konvergen $\left(\left|\hat{\beta}^{(m)} - \hat{\beta}^{(m+1)}\right| \leq \varepsilon \text{ dan } \left|\hat{\gamma}^{(m)} - \hat{\gamma}^{(m+1)}\right| \leq \varepsilon\right)$, dengan ε merupakan nilai yang sangat kecil misalnya 10^{-5}.

2.10. *Geographically Weighted Zero Inflated Poisson Regression* (GWZIPR)

GWZIPR melibatkan efek spasial dalam menyelesaikan data cacahan dan memiliki nilai nol yang banyak (*axcess zeros*). GWZIPR merupakan pengembangan model ZIP maka fungsi peluangnya seperti pada Persamaan (2.5) diketahui $\mu_i = \exp(x_i^T \beta(u_i, v_i))$ dan $\omega_i = \frac{\exp(x_i^T \gamma(u_i, v_i))}{1+\exp(x_i^T \gamma(u_i, v_i))}$, dengan melibatkan

efek spasial pada masing-masing pengamatan dari wilayah yang berbeda-beda (u_i, v_i). $\beta(u_i, v_i)$ dan $\gamma(u_i, v_i)$ adalah parameter regresi pada wilayah (u_i, v_i). Estimasi parameter model GWZIPR dapat diperoleh dengan menggunakan metode MLE dan metode iteratif algoritma EM. Fungsi *likelihood* dan fungsi total *ln-likelihood* model GWZIPR yaitu:

$$L\big(\beta(u_i, v_i), \gamma(u_i, v_i)\big) =$$

$$\begin{cases} \prod_{i=1}^{n} \frac{\exp\left(x_i^T \gamma(u_i,v_i)\right)+(\exp(-\exp(x_i^T \beta(u_i,v_i))))}{1+\exp\,(x_i^T \gamma(u_i,v_i))} & \mathrm{y}_i = 0 \\ \frac{\prod_{i=1}^{n} \frac{1}{1+\exp\,(x_i^T \gamma(u_i,v_i))} (\exp(-\exp(x_i^T \beta(u_i,v_i))))(\exp(x_i^T \beta(u_i,v_i)))^{\mathrm{y}_i}}{\prod_{i=1}^{n} \mathrm{y}_i!} & \mathrm{y}_i = 1,2,.. \end{cases}$$

$$\ln L_T\big(\beta(u_i, v_i), \gamma(u_i, v_i)\big) = \sum_{i=1}^{n} \ln\Big(\exp\left(x_i^T \gamma(u_i, v_i)\right) +$$

$$\Big(\exp - \Big(\exp\left(x_i^T \beta(u_i, v_i)\right)\Big)\Big)\Big) w_{ij}(u_i, v_i) -$$

$$\sum_{i=1}^{n} \ln\left(1 + \exp\left(x_i^T \gamma(u_i, v_i)\right)\right) w_{ij}(u_i, v_i) + \sum_{i=1}^{n} (\mathrm{y}_i x_i^T \beta(u_i, v_i) -$$

$$\exp(x_i^T \beta(u_i, v_i))) w_{ij}(u_i, v_i) - \sum_{i=1}^{n} \ln(\mathrm{y}_i!)\, w_{ij}(u_i, v_i)$$

$w_{ij}(u_i, v_i)$ adalah pembobot untuk setiap lokasi (u_i, v_i) (Ismah dkk., 2020).

2.10.1 Pengujian Serentak Parameter Model GWZIPR

Uji hipotesis yang pertama dilakukan adalah pengujian model secara serentak untuk menguji signifikansi pada seluruh parameter $\boldsymbol{\beta}(u_i, v_i)$ dan $\boldsymbol{\gamma}(u_i, v_i)$. Bentuk hipotesisnya:

$H_0: \beta_1(u_i, v_i) = \cdots = \beta_p(u_i, v_i) = \gamma_1(u_i, v_i) = \cdots = \gamma_p(u_i, v_i) = 0$

H_1: paling tidak ada satu $\beta_k(u_i, v_i) \neq 0$ atau $\gamma_k(u_i, v_i) \neq 0$

$i = 1,2,..,n; k = 1,2,\ldots,p$

Statistik uji yang digunakan yaitu uji rasio *likelihood* (devians). Di mana nilai devians dibentuk dari parameter model penuh $\left(\Omega = \left(\boldsymbol{\beta}(u_i, v_i), \boldsymbol{\gamma}(u_i, v_i)\right)\right)$ dan model intersep $\left(\omega = \left(\beta_0(u_i, v_i), \gamma_0(u_i, v_i)\right)\right)$. Devians terbentuk dari dua ln-*likelihood* dari $L(\widehat{\Omega})$ dan $L(\widehat{\omega})$ yakni:

$$G = -2ln\left[L\left(\frac{\widehat{\omega}}{\widehat{\Omega}}\right)\right] = 2\left[\ln L(\widehat{\Omega}) - \ln L(\widehat{\omega})\right] \sim \chi_v^2$$

G berdistribusi *Chi-Square* dengan derajat bebas p. Tolak H_0 jika nilai $G > \chi_{v,\alpha}^2$.

2.10.2 Pengujian Parsial Parameter Model GWZIPR

Uji hipotesis yang kedua adalah pengujian parameter model secara parsial. Pengujian ini untuk mengetahui signifikansi masing-masing parameter $\boldsymbol{\beta}(u_i, v_i)$ dan $\boldsymbol{\gamma}(u_i, v_i)$. Bentuk hipotesis pengujian secara parsial pada kedua parameter model adalah sebagai berikut:

$H_0: \left(\beta_k(u_i, v_i)\right) = 0 \quad H_1: \beta_k(u_i, v_i) \neq 0; k = 1,2, \dots, p$

$H_0: \left(\gamma_k(u_i, v_i)\right) = 0 \quad H_1: \gamma_k(u_i, v_i) \neq 0; k = 1,2, \dots, p$

Statistik uji yang digunakan yaitu statistik uji Wald (W):

$W = \frac{\widehat{\boldsymbol{\beta}}(u_i, v_i)}{se\left(\widehat{\boldsymbol{\beta}}(u_i, v_i)\right)}$ dan $W = \frac{\widehat{\gamma}(u_i, v_i)}{se\left(\widehat{\gamma}(u_i, v_i)\right)}$

Kriteria keputusan tolak H_0 jika statistik uji $W > Z_{a/2}$ (Faricha, 2016).

2.11. Demam Berdarah *Dengue* (DBD)

Jumlah penderita dan luas daerah penyebaran penyakit DBD semakin bertambah dari tahun ke tahun seiring meningkatnya mobilitas dan kepadatan penduduk di Indonesia. Penyakit ini pertama kali ditemukan di Indonesia pada tahun 1968 di Surabaya, terdapat 58 orang yang terinfeksi dan 24 orang diantaranya meninggal dunia. Dan sejak saat itu DBD menyebar luas seluruh Indonesia. Penyakit DBD disebabkan oleh virus dari genus *Flavivirus* yang ditularkan ke manusia melalui gigitan nyamuk *Aedes* yang terinfeksi virus *dengue*. Penyakit DBD biasanya menyebar dan memburuk dalam waktu yang relatif singkat yaitu lima sampai tujuh hari jika tidak segera ditangani dengan cepat dan tepat (Asrirawan dan Khaerati, 2017).

Pada penelitian Fatati dkk. (2017) yang menganalisis data kasus DBD di provinsi Jawa Tengah dengan analisis regresi spasial, menghasilkan bahwa faktor yang berpengaruh terhadap jumlah kasus DBD adalah jumlah puskesmas per 1000 penduduk, jumlah polides per 1000 penduduk, kepadatan penduduk tiap km^2, persentase penduduk terhadap akses air minum berkelanjutan layak, persentase kualitas air bersih yang bebas bakteri, jamur dan bahan kimia, dan jumlah sarana mata air terlindungi. Selain dari enam variabel penjelas yang signifikan, muncul juga variabel baru yang signifikan yaitu kasus DBD di kabupaten/kota di sekitarnya. Adapun dalam penelitian Khairatunnisa dan Budiman (2018) menganalisis faktor perilaku yang berhubungan dengan kejadian DBD di

Kecamatan Binjai Timur Kota Binjai tahun 2017, hasil penelitian menunjukkan terdapat hubungan antara kebiasaan menutup penampungan air, kebiasaan memakai kelambu, kebiasaan menguras penampungan air, kebiasaan mengubur barang bekas, dan kebiasaan menggantung pakaian dengan kejadian DBD.