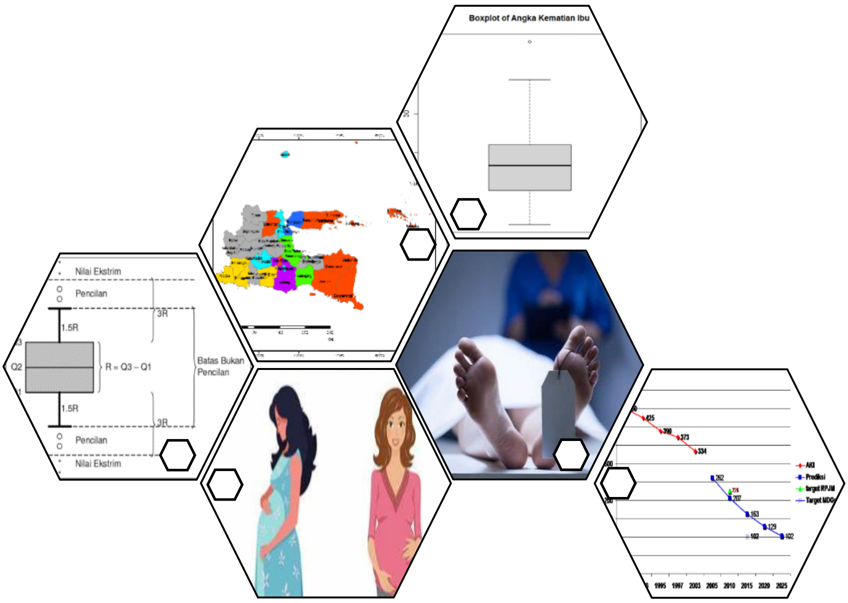
PEMODELAN *GEOGRAPHICALLY WEIGHTED POISSON REGRESSION SEMIPARAMETRIC* PADA DATA YANG MENGANDUNG PENCILAN

GEOGRAPHICALLY WEIGHTED POISSON REGRESSION SEMIPARAMETRIC MODELING ON DATA CONTAINING OUTLIERS



FITRIAYU

H062221006

PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

PEMODELAN *GEOGRAPHICALLY WEIGHTED POISSON REGRESSION SEMIPARAMETRIC* PADA DATA YANG MENGANDUNG PENCILAN

FITRIAYU

H062221006

MAGISTER STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
2024

PEMODELAN *GEOGRAPHICALLY WEIGHTED POISSON REGRESSION SEMIPARAMETRIC* PADA DATA YANG MENGANDUNG PENCILAN

Tesis
sebagai salah satu syarat untuk mencapai gelar magister

Program Studi Magister Statistika

Disusun dan diajukan oleh

FITRIAYU
H062221006

Kepada

PROGRAM STUDI MAGISTER STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2024

TESIS

PEMODELAN *GEOGRAPHICALLY WEIGHTED POISSON REGRESSION SEMIPARAMETRIC* PADA DATA YANG MENGANDUNG PENCILAN

FITRIAYU
H062221006

telah dipertahankan di depan Panitia Ujian Magister pada 16 Agustus 2024
dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Magister Statistika
Departemen Statisika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan

Pembimbing Utama	Pembimbing Pendamping
Dr. Anna Islamiyati, S.Si., M.Si.	**Dr. Erna Tri Herdiani, S.Si., M.Si.**
NIP. 19770808 200501 2 002	NIP. 19750429 200003 2 001
Ketua Program Studi Magister Statistika	Dekan Fakultas Matematika dan Magister Ilmu Pengetahuan Alam Universitas Hasanuddin
Dr. Erna Tri Herdiani, S.Si., M.Si.	**Dr. Eng. Amiruddin, M.Si.**
NIP. 19750429 200003 2 001	NIP. 19720515 199702 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul "Pemodelan *Geographically Weighted Poisson Regression Semiparametric* Pada Data Yang Mengandung Pencilan" adalah benar karya saya dengan arahan dari tim pembimbing (Dr. Anna Islamiyati, S.Si., M.Si. sebagai Pembimbing Utama dan Dr. Erna Tri Herdiani, S.Si., M.Si. sebagai Pembimbing Pendamping). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini akan dipublikasikan di SCIK *Publishing Corporation* sebagai artikel dengan judul "*Regional Classification Based on Maternal Mortality Rate Using a Robust Semiparametric Geographically Weighted Poisson Regression Model*".

Dengan ini saya melimpahkan hak cipta dari karya saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 16 Agustus 2024

METERAI TEMPEL
10000
CA6ALX326563249

FITRIAYU
NIM. H062221006

UCAPAN TERIMA KASIH

Segala puji hanya milik Allah *Subhanallahu Wa Ta'ala* atas limpahan rahmat dan hidayah-Nya kepada penulis. Shalawat dan salam tercurahkan kepada Rasulullah *Shallallahu 'Alaihi Wa sallam*, keluarganya, *tabi'in*, *tabi'ut tabi'in*, serta orang-orang sholeh yang haq hingga kadar Allah berlaku atas diri mereka. *Alhamdulillaahirobbil'aalamiin*, berkat rahmat dan kemudahan dari Allah *Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan tesis berjudul " Pemodelan *Geographically Weighted Poisson Regression Semiparametric* Pada Data Yang Mengandung Pencilan" sebagai salah satu syarat memperoleh gelar magister pada Program Studi Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis menyadari bahwa masih banyak kekurangan dalam tesis ini, untuk itu dengan segala kerendahan hati, penulis memohon maaf.

Makassar, 16 Agustus 2024
Penulis,

Fitriayu

ABSTRAK

FITRIAYU. **Pemodelan *Geographically Weighted Poisson Regression Semiparamterik* Pada Data Yang Mengandung Pencilan** (dibimbing oleh Dr. Anna Islamiyati, S.Si., M.Si. dan Dr. Erna Tri Herdiani, S.Si., M.Si.).

Latar Belakang. Model *Geographically Weighted Poisson Regression Semiparametric* (GWPRS) merupakan metode yang dikembangkan dari model GWPR yang mengkombinasi antara parameter yang bersifat lokal dan parameter yang bersifat global terhadap lokasi. Dalam menganalisis data dengan menggunakan model GWPRS, terkadang ditemukan adanya pencilan. Pencilan ini dapat diidentifikasi secara jelas karena berbeda dengan titik sampel yang lainnya. Adanya pencilan dapat berdampak terhadap hasil estimasi parameter model yang menyebabkan estimasi parameter menjadi bias. **Tujuan**. Penelitian ini bertujuan untuk mendapatkan estimasi parameter model GWPRS yang mengandung pencilan kemudian model tersebut akan diaplikasikan pada data angka kematian ibu di Provinsi Jawa Timur dengan menggunakan matriks pembobot *Tukey Bisquare*. **Metode**. Adapun metode yang digunakan pada penelitian ini adalah metode *GM-estiamator* dengan iterasi IRLS (*Iteratively Reweighted Least Square*). **Hasil**. Hasil penelitian ini menunjukkan bahwa model GWPRS yang mengandung pencilan dapat mendekripsikan angka kematian ibu beserta faktor-faktor yang mempengaruhinya dengan baik serta dapat digunakan sebagai bahan untuk menurunkan angka kematian ibu. **Kesimpulan**. Dari hasil penelitian dapat disimpulkan bahwa faktor-faktor yang berpengaruh signifikan terhadap angka kematian ibu yaitu persentase persalinan komplikasi ditangani tenaga kesehatan (x_1), pelayanan kesehatan ibu hamil (K1) dan (K2) (x_2) dan wanita menikah dibawah 17 tahun (x_4). Adapun variabel yang berpengaruh signifikan secara global terhadap angka kematin ibu yaitu ibu hamil memperoleh tablet Fe1 dan Fe3 (x_3) dan ibu nifas memperoleh vitamin A (x_5).

Kata Kunci: GWPRS; pencilan; *GM-estimator*; *Tukey Bisquare*; angka kematian ibu

ABSTRACT

FITRIAYU. **Geographically Weighted Poisson Regression Semiparametric Modeling On Data Containing Outliers** (supervised by Dr. Anna Islamiyati, S.Si., M.Si. dan Dr. Erna Tri Herdiani, S.Si., M.Si.).

Background. The Semiparametric Geographically Weighted Poisson Regression (GWPRS) model is an extension of the GWPR model that combines local and global parameters in relation to location. When analyzing data using the GWPRS model, outliers are sometimes encountered. These outliers can be distinctly identified because they differ from other sample points. The presence of outliers can impact the estimation results of the model's parameters, leading to biased parameter estimates. **Objective**. This study aims to estimate the parameters of the GWPRS model in the presence of outliers and apply the model to maternal mortality data in East Java Province using the Tukey Bisquare weighting matrix. **Method**. The method used in this study is the GM-estimator method with IRLS (Iteratively Reweighted Least Squares) iterations. **Results.** The results of this study indicate that the GWPRS model, which accounts for outliers, effectively describes maternal mortality rates and the factors influencing them. This model can be used as a basis for reducing maternal mortality rates. **Conclusion**. From the study results, it can be concluded that the factors significantly influencing maternal mortality are the percentage of complicated deliveries handled by healthcare professionals (x_1), antenatal care services (K1 and K2) (x_2), and women married under the age of 17 (x_4). The variables that have a significant global impact on maternal mortality are pregnant women receiving Fe1 and Fe3 tablets (x_3) and postpartum women receiving vitamin A (x_5).

Keywords: GWPRS; outliers; GM-estimator; Tukey Bisquare; maternal mortality rate

DAFTAR ISI

Halaman

DAFTAR TABEL

Nomor urut Halaman

DAFTAR GAMBAR

Nomor urut Halaman

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan suatu metode statistik yang dapat digunakan dalam mengevaluasi hubungan antara variabel respon dan variabel prediktor. Metode ini bertujuan untuk mengidentifikasi faktor-faktor yang diperkirakan memiliki pengaruh terhadap variabel yang sedang diteliti serta hasilnya disajikan dalam bentuk model regresi. Pada umumnya, analisis regresi diterapkan pada data observasi di mana variabel respon bersifat kontinu (Nohe, 2013). Pemodelan regresi dapat diterapkan di berbagai disiplin ilmu, termasuk ekonomi, ilmu sosial, dan kesehatan. Di lapangan, data yang digunakan pada penelitian sering kali berupa data spasial (Fadlilah, 2018). Data spasial adalah data yang berkaitan dengan lokasi tertentu di permukaan bumi. Informasi dalam data ini mencakup koordinat, aspek geografi, dan geometri dari berbagai objek (titik, garis, atau poligon), serta atribut tambahan yang berhubungan dengan objek-objek tersebut. Keputusan yang berkaitan dengan aspek spasial diambil berdasarkan analisis data yang dilakukan di lapangan (Leung, 2000). Seiring perkembangan zaman, data spasial telah menjadi alat yang sangat penting dalam perencanaan pengelolaan dan pembangunan sumber daya, dengan perhatian yang besar terhadap kondisi pada setiap wilayah. Dalam data spasial terdapat faktor spasial yang berarti lokasi geografisnya, karena terdapat heterogenitas spasial (perbedaan antara lokasi) di dalamnya (Anselin, 1988). Heterogenitas ini menyebabkan koefisien regresi yang bervariasi secara spasial, sehingga variabel prediktor yang sama dapat menghasilkan respons yang berbeda pada setiap lokasi pengamatan (Aguero, 2006). Maka sebab itu, untuk mengolah data spasial menjadi informasi yang berguna, diperlukan model regresi yang dapat mempertimbangkan pengaruh heterogenitas spasial yang terdapat pada data.

Regresi Berbobot Geografis (*Geographically Weighted Regression*/GWR) adalah sebuah model statistik yang dirancang untuk mengatasi heterogenitas spasial (Hailegebreal, 2023). GWR memberikan kemampuan kepada peneliti untuk mengevaluasi variasi spasial (geografis) yang mungkin terdapat dalam hubungan antara variabel dependen dan independen di seluruh wilayah studi (Clement dkk, 2009). Fotheringham dkk. (2002) menyatakan bahwa GWR adalah model regresi linier bersifat lokal dan dapat menghasilkan estimasi parameter pada setiap lokasi data pengamatan, oleh karena itu setiap lokasi dapat memiliki interpretasi yang berbeda-beda. Meskipun GWR didasarkan pada asumsi kerangka model regresi yang sederhana, dalam memodelkan data diskrit yang mengikuti distribusi Poisson model GWR kurang tepat untuk digunakan. Oleh karena itu, Nakaya dkk. (2005) merekomendasikan analisis statistik baru yang dikenal sebagai Regresi Poisson Berbobot Geografis (*Geographically Weighted Poisson Regression*/GWPR). Model

GWPR dirancang agar dapat memodelkan data dimana variabel responnya berupa data diskrit.

Dalam penelitian yang dilakukan oleh Erdkhadifa (2022), ditemukan bahwa beberapa koefisien regresi dalam model Regresi Poisson Berbobot Geografis (GWPR) tidak menunjukkan variasi spasial, di mana beberapa koefisien regresi pada model GWPR berbeda di beberapa wilayah. Untuk mengatasi masalah ini, model GWPR dikembangkan lagi menjadi model *Geographically Weighted Poisson Regression Semiparametrik* (GWPRS). Model GWPRS merupakan bentuk lokal dari regresi Poisson yang memperhatikan lokasi dan dapat secara bersamaan menangani masalah korelasi spasial. Menurut Nakaya (2005), untuk membangun model GWPRS, diperlukan kombinasi antara parameter yang bervariasi di setiap lokasi sebagai variabel yang bersifat lokal dengan parameter yang tetap pada setiap lokasi sebagai variabel yang bersifat global.

Pada penerapannya, model GWPRS sering kali menghadapi masalah pencilan. Pencilan adalah pengamatan yang letaknya jauh dari pusat data dan dapat memengaruhi koefisien regresi. Keberadaan pencilan ini dapat mengakibatkan estimasi parameter pada model menjadi tidak konsisten yang dapat dilihat dari nilai standar *error* yang tinggi saat metode kuadrat terkecil digunakan. Maka untuk menangani pencilan pada model regresi spasial dapat menggunakan metode regresi *robust*. Regresi *robust* merupakan alat analisis yang dirancang untuk menangani data yang mengandung pencilan. Menurut Chen (2002), terdapat beberapa metode dalam regresi *robust* untuk estimasi, salah satunya adalah metode *Genralized M-estimator* atau yang sering disingat menjadi *GM-estimator* yang dapat menganalisis data dengan asumsi bahwa sebagian besar pencilan berada pada variabel prediktor (Obikee dkk., 2014). Maka dari itu, pada penelitian ini data yang mengandung pencilan akan diestimasi menggunakan metode *GM-estimator*.

Model GWPRS yang menggabungkan parameter yang berbeda di setiap lokasi dengan parameter yang tetap akan diterapkan pada penelitian ini, karena dalam penerapannya tidak semua variabel berpengaruh spasial. Selain itu, penelitian ini mengacu pada studi sebelumnya, termasuk yang dilakukan oleh Aini (2013), yang menggunakan model *Geographically Weighted Poisson Regression Semiparametric* (GWPRS) pada data yang mengandung pencilan dan diestimasi dengan metode *GM-estimator*. Selain itu, penelitian ini juga merujuk pada karya Putra (2015), yang mengembangkan model *Geographically Weighted Poisson Regression Semiparametric* (GWPRS) yang menggabungkan parameter lokal dan global dengan asumsi distribusi poisson menggunakan pembobot *Gaussian Kernel*. Oleh karena itu, penelitian ini akan mengeksplorasi model GWPRS dengan estimasi parameter menggunakan metode *GM-estimator* dengan pembobot *Tukey Bisquare* pada data yang mengandung pencilan.

Berdasarkan uraian tersebut, penulis menyusun penelitian ini dengan judul “Pemodelan *Geographically Weighted Poisson Regression Semiparametric* Pada Data yang Mengandung Pencilan”. Model GWPRS yang mencakup pencilan akan diterapkan pada Angka Kematian Ibu (AKI). Angka Kematian Ibu adalah indikator yang penting dimana dapat mencerminkan kesejahteraan masyarakat di suatu negara. Secara umum, kematian ibu dipengaruhi oleh beragam faktor, termasuk kualitas lingkungan, status sosial, ekonomi, dan layanan kesehatan. Maka dari itu untuk menganalisis hubungan antara variabel respon berupa data diskrit dengan variabel lain yang mungkin diskrit, kontinu, kategorik, atau kombinasi dari semuanya dalam konteks kematian ibu, model regresi Poisson menjadi salah satu metode analisis regresi yang dapat digunakan.Namun, metode regresi Poisson kurang responsif ketika diterapkan pada data spasial yang mempertimbangkan kondisi geografis (Millah, 2015).

Metode yang tepat untuk mengatasi keheterogenan spasial pada data diskrit yang berdistribusi Poisson yaitu dapat digunakan metode model GWPRS yang merupakan model yang dikembangkan berdasarkan model GWR dan model GWPR. Model GWPRS ini mampu menggabungkan parameter yang berubah setiap lokasi atau peubah lokal dengan parameter tetap di setiap lokasi atau peubah global (Nakaya dkk., 2005). Pada model GWPRS, terdapat pembobot yang didasarkan pada posisi atau jarak antara satu lokasi dan lokasi pengamatan yang lainnya. Dengan demikian, model ini dapat digunakan untuk menganalisis angka kematian ibu di Provinsi Jawa Timur. Memahami faktor-faktor yang diduga berpengaruh terhadap kematian ibu sehingga dapat memberikan wawasan bagi pemerintah daerah dalam upaya menurunkan angka kematian ibu.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini disusun berdasarkan latar belakang yang telah dijelaskan yaitu sebagai berikut:

1. Bagaimana bentuk dari estimasi parameter model GWPRS pada data yang mengandung pencilan?
2. Bagaimana penerapan dari model GWPRS yang mengandung pencilan pada data angka kematian ibu di Provinsi Jawa Timur tahun 2021?

1.3 Tujuan Penelitian

Selain itu terdapat tujuan penelitian yang akan dicapai pada penelitian ini berdasarkan rumusan masalah adalah:

1. Memperoleh hasil estimasi parameter dari model GWPRS pada data yang mengandung pencilan.
2. Memperoleh hasil penerapan dari model GWPRS yang mengandung pencilan pada data angka kematian ibu di Provinsi Jawa Timur tahun 2021.

1.4 Manfaat Penelitian

Hasil dari penelitian, diharapkan dapat memberikan berbagai manfaat antara lain yaitu:

1. Menambah pengetahuan yang lebih luas lagi mengenai estimasi parameter model GWPRS pada data yang mngandung pencilan.
2. Dapat memberikan wawasan mengenai penerapan model GWPRS pada angka kematian ibu, yang diharapkan dapat berkontribusi dalam upaya mengurangi angka kematian ibu.

1.5 Batasan Penelitian

Selain itu perlu diadakan batasan masalah untuk mendekati hasil yang diaharapkan diantaranya yaitu:

1. Pencilan yang akan dipergunakan pada penelitian ini yaitu terdapat pada variabel prediktor.
2. Metode yang digunakan dalam proses estimaso parameter pada model GWPRS yang mengandung pencilan adalah metode *GM-estimator* dimana fungsi pembobotnya adalah *Tukey Bisquare.*
3. Data yang digunakan untuk menerapkan model GWPRS yaitu data Angka Kematian Ibu (AKI) di Jawa Timur tahun 2021.

1.6 Teori Pembahasan

Pada teori pembahasan akan dibahas mengenai teori apa saja yang akan digunakan pada penelitian ini. Berikut adalah teori-teori pembahasan pada penelitian ini diantaranya sebagai berikut:

1.6.1 Distribusi Poisson

Distribusi Poisson merupakan distribusi untuk variabel acak yang berupa diskrit, dimana berarti nilai data yang dihasilkan yaitu berupa data cacah atau bilangan bulat yang positif (Nohe, 2013). Contoh dari data cacah yaitu jumlah kejadian yang terjadi pada rentang waktu atau pada wilayah-wilayah tertentu. Distribusi Poisson juga sering diterapkan untuk memodelkan kejadian yang jarang terjadi dalam periode waktu tertentu (Kleinbaum, 1988). Menurut Walpole (1995), sebuah variabel acak dikatakan memiliki distribusi Poisson jika fungsi densitas (fungsi peluang) dinyatakan sebagai berikut:

$$P(Y = y) = \frac{e^{-\mu}\mu^{y}}{y!}, y = 0,1,2,... \tag{1}$$

dimana:

μ : rata-rata jumlah kejadian yang terjadi dalam suatu interval waktu atau di area tertentu serta memenuhi $\mu > 0$

y : banyaknya jumlah yang berhasil (kejadian dalam interval waktu dan are yang ditentukan).

Persamaan (1) juga dikenal sebagai fungsi peluang Poisson. Jika y adalah variabel acak yang mengikuti distribusi Poisson, maka nilai rata-rata dan variansinya adalah sama yaitu μ.

1.6.2 Pengujian Overdispersi

Pada model regresi Poisson, ada beberapa asumsi yang perlu dipenuhi, salah satunya yairu kesetaraan antara rata-rata atau mean dan variansinya atau variance yang dikenal sebagai ekuidispersi (Darnah, 2011). Tetapi, dalam menganalisis data statistik sering kali ditemui situasi di mana variansi lebih besar ataupun lebih kecil dari rata-rata. Kondisi ini dikenal dengan underdispersi (*underdispersion*) atau overdispersi (*overdispersion*). Adapun salah satu penyebab terjadinya overdispersi yaitu adanya jumlah nilai nol yang berlebihan pada variabel respon (Kusuma dkk., 2013). Overdispersi terjadi ketika varians lebih besar daripada mean dimana data juga dianggap overdispersi ketia nilai taksiran dispersi lebih besar dari 1, sementara data dikategorikan underdispersi ketika nilai taksiran dispersinya lebih kecil atau kurang dari 1 (Wang, 2004). Overdispersi dapat ditulis sebagai berikut:

$$\boldsymbol{Var}(\boldsymbol{Y}) > \boldsymbol{E}(\boldsymbol{Y})$$

Selain itu, untuk mendeteksi overdispersi dalam suatu data uji statistik yang dapat digunakan adalah uji overdispersi dimana bisa diakses melalui package AER yang terdapat dalam *software R* (Herindrawati dkk., 2017). Dengan hipotesis sebagai berikut:

H_0 : tidak terjadi Overdispersi

H_1 : terjadi Overdispersi

Keputusan pada uji overdispersi adalah ketika taksiran dispersi lebih besar dari 1 dan nilai *p-value* lebih kecil dari α, maka hipotesis H_0 ditolak, yang menunjukkan bahwa terjadi overdispersi. Sebaliknya, jika taksiran dispersi kurang dari atau lebih kecil dari 1 dan nilai *p-value* lebih besar dari α, maka hipotesis H_0 diterima, dimana berarti tidak terjadi overdispersi.

1.6.3 Pengujian Heterogenitas Spasial

Pada penelitian yang melibatkan informasi lokasi geografis, penting untuk mempertimbangkan adanya efek spasial dalam model. Efek spasial merujuk pada ketergantungan spasial yang muncul akibat korelasi antar wilayah (dependensi spasial) dan keragaman (heterogenitas) antar lokasi. Lokasi suatu pengamatan dapat menciptakan hubungan dengan pengamatan lain yang berdekatan, baik melalui hubungan antar pengamatan maupun jarak antara pengamatan tersebut. Perbedaan karakteristik di antara satu titik lokasi pengamatan dengan titik lokasi lainnya menyebabkan munculnya keragaman spasial. Pengujian heterogenitas (keragaman) spasial dilakukan untuk mengidentifikasi ciri khas di setiap lokasi pengamatan, yang dapat menyebabkan perbedaan parameter regresi secara spasial. Uji heterogenitas spasial dapat dilakukan dengan menggunakan statistik uji *Breusch-Pagan* (BP) (Anselin, 1988) dengan hipotesis sebagai berikut:

H_0 : $\sigma^2(u_1, v_1) = \sigma^2(u_2, v_2) = \cdots = \sigma^2(u_n, v_n)$

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

Pengujian heterogenitas spasial menggunakan statistik uji *Breusch-Pagan* (BP) yang dinyatakan sebagai berikut:

$$BP = \left(\frac{1}{2}\right)\boldsymbol{f}^{\boldsymbol{t}}\boldsymbol{Z}(\boldsymbol{Z}^{\boldsymbol{t}}\boldsymbol{Z})^{-\boldsymbol{1}}\boldsymbol{Z}^{\boldsymbol{t}}\boldsymbol{f} + \left(\frac{\boldsymbol{1}}{\boldsymbol{T}}\right)\left(\frac{\boldsymbol{e}'\boldsymbol{W}\boldsymbol{e}}{\sigma^2}\right)^{\boldsymbol{2}} \quad \sim \chi^2_{(\alpha,p)} \tag{2}$$

dengan vektor $\boldsymbol{f}$ adalah

$$\boldsymbol{f} = \left(\frac{{e_j}^2}{\sigma^2} - 1\right)$$

dimana e_j adalah sisaan yang terdapat pada lokasi pengamatan yang ke-j, $\boldsymbol{Z}$ adalah matriks yang berukuran $n \times (p+1)$ dimana berisikan vektor yang telah dinormal bakukan pada tiap lokasi pengamatan, σ^2 merupakan nilai ragam dari residual (e_i), T yaitu *trace* $[\boldsymbol{W}^{\boldsymbol{t}}\boldsymbol{W} + \boldsymbol{W}^{\boldsymbol{2}}]$ serta $\boldsymbol{W}$ adalah matriks pembobot antara lokasi pengamatan. Adapun kriteria untuk menolak H_0, yaitu ketika $BP > \chi^2_{(\alpha,p)}$ atau $p-value > \alpha$ maka dapat diambil kesimpulan yaitu terdapat heterogenitas spasial.

1.6.4 Fungsi Pembobot dan Matriks *Bandwidth*

Penaksir parameter pada model GWPR menggunakan pembobot spasial dimana memberikan pembobot yang berbeda-beda pada setiap pengamatan di setiap lokasi. Pembobot berfungsi untuk merepresentasikan data pengamatan di antara lokasi-lokasi yang ada. Data pengamatan yang berada lebih dekat maka akan memiliki pengaruh yang lebih signifikan dibandingkan dengan sata pengamatanyang berada lebih jauh, sehingga diberikan bobot yang lebih besar. Umumnya, nilai bobot ini bergantung pada jarak lokasi pengamatan (Chasco dkk, 2007).

Pembobot spasial ditentukan oleh data koordinat dari titik-titik pada lokasi pengamatan. Koordinat ini digunakan untuk menghitung jarak antara lokasi-lokasi pengamatan. Jarak antara lokasi (u_i, v_i) dengan lokasi (u_j, v_j) dapat disimbolkan dengan d_{ij} serta dihitung menggunakan jarak *Euclidean* yaitu sebagai berikut (Chaso dkk, 2007).

$$d_{ij} = \sqrt{(u_i, u_j)^2 + (v_i, v_j)^2} \tag{3}$$

dimana u_i adalah *latitude* atau garis lintang serta v_i merupakan garis *longitude* atau garis bujur.

Pembobot spasial dapat dihitung dengan menggunakan fungsi pembobot. Salah satu fungsi pembobot yang dapat digunakan yaitu fungsi *Gaussian Kernel*. Dimisalkan w_{ij} merupakan pembobot spasial yang diberikan oleh pengamatan pada lokasi ke-j untuk model GWPR pada lokasi ke-i maka pembobot spasial w_{ij} berdasarkan fungsi *Fixed Gaussian Kernel* dapat dinyatakan sebagai:

$$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right), j = 1,2, \ldots, n \tag{4}$$

dengan d_{ij} adalah jarak *Euclidean* dan b merupakan parameter panghalus (*bandwith*). Nilai *bandwidth* dalam persamaan (4) merupakan konstanta pada setiap pengamatan dan disebut *Fixed Bandwidth*. Sebagai alternatif, fungsi pembobot lain yang dapat digunakan yaitu fungsi pembobot *Gaussian* dengan *adaptive bandwidth*, di mana *bandwidth* berbeda pada setiap lokasi pengamatan. Nilai pembobot dari fungsi *Adaptive Gaussian Kernel* dapat dinyatakan sebagai:

$$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b_i}\right)^2\right), j = 1,2, \ldots, n \tag{5}$$

Nilai pembobot spasial yang diberikan akan sangat dipengaruhi oleh *bandwidth*, sehingga pemilihan nilai *bandwidth* menjadi hal yang penting. Jika nilai *bandwidth* terlalu besar maka pembobot pada setiap pengamatan akan mendekati nilai satu yang dapat menghasilkan model global. Di sisi lain, jika nilai *bandwidth* terlalu kecil maka pembobot akan mendekati nol sehingga estimasi parameter akan bergantung pada lokasi pengamatan yang berdekatan serta berpotensi menghasilkan model yang bias.

Pemilihan *bandwidth* optimum juga mempunyai peran penting dalam pembentukan matriks pembobot (Fotheringham et al., 2002). Ukuran *bandwidth* yang digunakan akan mempengaruhi akurasi model terkait dengan variansi dan bias dari estimasi yang didapatkan. Maka dari itu, diperlukan *bandwidth yang* optimum agar dapat mengatur besaran variansi dan bias tersebut (Nakaya dkk., 2005). Salah satu metode yang dapat digunakan untuk menentukan *bandwidth* yang optimum adalah dengan menggunakan pendekatan *Cross Validation* (CV). Ketika nilai CV yang dihasilkan mencapai nilai minimum maka akan menghasilkan nilai b*andwidth* yang optimum.

$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(b)]^2 \tag{6}$$

dimana y_i merupaka nilai untuk menaksir pengamatan di lokasi (u_i, v_i), b_i adalah nilai penaksir parameter y_i untuk setiap pengamatan yang ada di lokasi (u_i, v_i) dapat dihilangka dari proses penaksir pada lokasi ke-i dan n merupakan banyaknya sampel penelitian. Sedangkan b*andwidth* yang optimum yaitu $\{\hat{b}_1, \hat{b}_2, \dots, \hat{b}_n\}$ sedemikian sehinggan nilai CV adalah minimum.

1.6.5 *Geographically Weighted Poisson Regression* (GWPR)

Geographically Weighted Poisson Regression (GWPR) merupakan bentuk lokal dari regresi Poisson yang dapat diterapkan pada data spasial. Perbedaan yang terdapat antara regresi Poisson global dan GWPR terletak pada estimasi parameternya. Pada model regresi Poisson global estimasi parameternya adalah bernilai sama di setiap lokasi pengamatan, sehingga bersifat global. Sementara itu, model GWPR menghasilkan estimasi parameter yang bervariasi untuk masing-masing wilayah, sehingga model GWPR dikatakan bersifat lokal (Nakaya dkk., 2005). Pemodelan pada data spasial dengan regresi Poisson global menjadi kurang representatif karena adanya interdependensi antara data variabel respon dengan lokasi pengamatan yang dapat menyebabkan nilai estimasi parameter regresi berbeda pada satu lokasi pengamatan dengan lokasi yang lain. Salah satu pendekatan pemodelan spasial yang sesuai untuk data spasial yaitu model GWPR.

GWPR memiliki struktur model regresi yang serupa dengan regresi Poisson, tetapi dalam model GWPR terdapat aspek geografis yang berfungsi sebagai pembobot. Oleh karena itu, menurut Fotheringham dkk., (2002) model GWPR dinyatakan dalam bentuk berikut:

$y_i \sim$ Poisson (μ_i) dengan

$$\mu_i = \exp \sum_{j=1}^{T} \beta_j(u_i, v_i) x_{ij} + \varepsilon_i \tag{7}$$

dimana y_i adalah nilai observasi dari variabel respon ke-i, adapun x_{ji} adalah nilai observasi dari variabel prediktor ke-j di setiap pengamatan ke-i, $\beta_j(u_i, v_i)$ adalah nilai parameter dengan $j = 1,2, \dots, k$, (u_i, v_i) merupakan titik koordinat yaitu lintang dan bujur untuk lokasi ke-i serta ε_i merupakan nilai *error* pada regresi ke-i.

Salah satu metode untuk memperkirakan parameter dalam model GWPR adalah dengan menggunakan *Maximum Likelihood Estimation* (MLE). Proses estimasi model GWPR dilakukan di setiap lokasi pengamatan dengan mempertimbangkan pembobot spasial. Langkah pertama yang dilakukan dari metode MLE adalah membentuk fungsi *likelihood*. Dikarenakan variabel respon mengikuti distribusi Poisson ($y_i \sim Poisson(x_i, \beta)$), maka fungsi *likelihood* dapat dinyatakan sebagai berikut:

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

Setelah didapatkan bentuk *likelihood* maka selanjutnya yaitu dilakukan operasi algoritma natural dengan begitu diperoleh sebagai berikut:

$$\ln L(\beta) = \sum_{i=1}^{n} (-\mu(x_i, \beta)) + y_i \ln(x_i, \beta) - \ln y_i!$$

dan dapat dituliskan kedalam bentuk berikut:

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

Faktor letak geografis berperan sebagai faktor pembobot dalam model GWPR. Nilai dari faktor ini bervariasi pada setiap daerah, dimana mencerminkan sifat lokal dari model GWPR. Maka dari itu, ketika pembobot diterapkan pada bentuk log-likelihood pada model lokal GWPR, didapatkan:

$$\ln L^*\big(\beta(u_i, v_i)\big) = \sum_{j=1}^{n} \big(y_j {x_j}^T \beta(u_i, v_i) - \exp\big({x_i}^T \beta(u_i, v_i)\big) - \ln y_i!\big) w_{ij}(u_i, v_i) \tag{8}$$

Estimasi dari parameter $\beta(u_i, v_i)$ dapat diperoleh dengan mendifferensialkan persamaan (8) terhadap $\beta(u_i, v_i)$ sehingga didapatkan:

$$\frac{\partial \ln L^*(\beta(u_i, v_i))}{\partial \beta^T(u_i, v_i)} = \sum_{j=1}^{n} \left(y_j x_j - x_j \exp\left({x_j}^T \beta(u_i, v_i)\right)\right) w_{ij}(u_i, v_i)$$

Dengan memaksimumkan bentuk diferensial tersebut maka diperoleh nilai estimasi, sehingga didapatkan:

$$\frac{\partial \ln L^*(\beta(u_i, v_i))}{\partial \beta^T(u_i, v_i)} = \sum_{j=1}^{n} \left(y_j x_j - x_j \exp\left({x_j}^T \beta(u_i, v_i)\right)\right) w_{ij}(u_i, v_i) = 0 \tag{9}$$

dikarenakan fungsi yang terdapat pada persamaan (8) berbetuk implisit, maka penaksir $\hat{\beta}(u_i, v_i)$ dapat dihasilkan berdasarkan algoritma iterasi Newton-Raphson sebagai berikut:

$$\hat{\beta}_{m+1}(u_i, v_i) = \left(\sum_{i=1}^{n} x_i w_{ij}(u_i, v_i) \hat{\mu}_{i(m)} {x_j}^T\right)^{-1} \sum_{i=1}^{n} x_i w_{ij}(u_i, v_i) \hat{\mu}_{i(m)} \left\{\left(\frac{y_i - \hat{\mu}_{i(m)}}{\hat{\mu}_{i(m)}}\right) + {x_j}^T \hat{\beta}_m(u_i, v_i)\right\}$$

1.6.6 Pengujian Variabilitas Model GWPR

Pengujian variabilitas GWPR dilakukan guna mengidentifikasi sifat lokal serta sifat global pada suatu variabel melalui pengujian pengaruh lokasi secara parsial (Leung dkk., 2000). Uji variabilitas bertujuan untuk menentukan adanya perbedaan pengaruh yang signifikan antara variabel x_j di berbagai lokasi pengamatan. Proses pengujian tersebut dilakukan dengan hipotesis berikut:

H_0 : $\beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \cdots = \beta_k(u_i, v_i)$ untuk $i = 1,2, \dots, n$

H_1 : minimal terdapat satu $\beta_1(u_i, v_i) \neq \beta_k(u_i, v_i)$ untuk $i = 1,2, \dots, n$

Dengan statistik uji yaitu:

$$F = \frac{V_j^2 \Big/ tr\left(\frac{1}{n}\boldsymbol{B}_j^t\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\right)\boldsymbol{B}_j}{SSE(H_1) \Big/ b_1} \sim \boldsymbol{F}_{(\alpha,P)}$$

dengan:

$$V_j^2 = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left(\hat{\beta}_j(u_i, v_i) - \frac{1}{n}\sum\nolimits_{i=1}^{n}\hat{\beta}_j(u_i, v_i)\right)^2 = \frac{1}{n}\boldsymbol{B}_j^t\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\boldsymbol{B}_j$$

$$\boldsymbol{B}_j = {\varepsilon_k}^t\left[\hat{\beta}_{m+1}(u_i, v_i)\right]$$

$$\beta_j(u_i, v_i) = [\hat{\beta}_1(u_i, v_i) \dots \hat{\beta}_j(u_i, v_i)]'$$

$$b_i = tr([(\boldsymbol{I} - \boldsymbol{S})^t\boldsymbol{I} - \boldsymbol{S}]^t), i = 1,2$$

$$c_i = tr\left(\frac{1}{n}\boldsymbol{B}_j^t\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\boldsymbol{B}_j\right)^t, i = 1,2$$

$\boldsymbol{J}$ merupakan matrik yang berukuran $n \times n$ dimana semua elemennya yaitu 1 serta ε_k merupakan vektor kolom yang berukuran $k + 1$ dimana bernilai satu pada elemen ke-j dan untuk lainnya bernilai nol.

Nilai statistika uji F mengikuti distribusi F dengan derajat bebas yaitu df_1 dan df_2, dimana $df_1 = \frac{{c_1}^2}{c_2}$ dan $df_2 = \frac{{b_1}^2}{b_2}$. Sehingga pada taraf signifikan sebesar α maka H_0 ditolak apabila $F_{hit} \geq F_{\alpha,df_1,df_2}$ yang berarti variabel prediktor tersebut bersifat lokal. Selain itu, jika nilai statistik uji $F_{hit} < F_{\alpha,df_1,df_2}$, maka variabel prediktor tersebut bersifat global.

1.6.7 *Geographically Weighted Poisson Regression Semiparametric* (GWPRS)

Geographically Weighted Poisson Regression Semiparametric (GWPRS) didefinisikan sebagai metode yang dikembangkan dari model GWPR dimana menggabungkan parameter bersifat lokal serta bersifat global berdasarkan lokasi (Nakaya dkk., 2005). Dengan demikian, model GWPRS merupakan kombinasi antara regresi Poisson dengan GWPR. Dalam model GWPRS, variabel respon diprediksi menggunakan variabel prediktor dimana koefisien regresinya $\left(\beta_j(u_i, v_i)\right)$ bergantung pada lokasi geografis serta β_g bersifat konstan. Lokasi geografis yang dinotasikan dengan (u_i, v_i) merupakan koordinat yang ada pada peta atau koordinat pada lokasi ke-j, dengan begitu menurut Nakaya dkk. (2005) model GWPRS dapat dituliskan:

$y_i \sim$ Poisson (μ_i) dengan

$$\mu_i = \exp\left(\sum_{j=0}^{k^*} \beta_j(u_i, v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig} + \varepsilon_i\right) \tag{10}$$

dimana y_i adalah nilai observasi dari variabel respon ke-i, x_{ij} adalah nilai observasi dari variabel prediktor pada lokasi (u_i, v_i), $\beta_j(u_i, v_i)$ adalah parameter pada model variabel prediktor yang bersifat lokal di lokasi (u_i, v_i) dengan $j = 1,2, \dots, k$ dan $i = 1,2, \dots, n$, (u_i, v_i) merupakan titik koordinat atau lintang, bujur pada lokasi ke-i, γ_g adalah parameter model variabel prediktor global pada lokasi (u_i, v_i), x_{ig} adalah nilai observai variabel prediktor global pada lokasi (u_i, v_i) serta $\varepsilon_i$merupaka nilai *error* regresi ke-i.

Estimasi parameter model GWPRS dapat menggunakan Metode Kemungkinan Maksimum dengan bentuk fungsi *likelihood* yaitu:

$$L\left(\beta_j(u_i, v_i), \gamma_g\right) = \prod_{i=1}^{n} \frac{e^{-\mu_i} {\mu_i}^{y_i}}{y_i!}$$

Persamaan ln *likelihood* :

$$\begin{aligned}
\ln L\left(\beta_j(u_i, v_i), \gamma_g\right) &= \ln \prod_{i=1}^{n} \frac{e^{-\mu_i} {\mu_i}^{y_i}}{y_i!} \\
&= \sum_{i=1}^{n} \ln \frac{e^{-\mu_i} {\mu_i}^{y_i}}{y_i!} \\
&= \sum_{i=1}^{n} (\ln(e^{-\mu_i}) + \ln({\mu_i}^{y_i}) - \ln(y_i!)) \\
&= \sum_{i=1}^{n} (-\mu_i + y_i \ln(\mu_i) - \ln(y_i!))
\end{aligned}$$

Setelah itu substitusikan $\mu_i = \exp(\sum_{j=0}^{k} \beta_j(u_i, v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig})$ maka:

$$\begin{aligned}
&= \sum_{i=1}^{n} \left(-\exp\left(\sum_{j=0}^{k} \beta_j(u_i, v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig}\right)\right) \\
&\quad + y_i \ln\left(\exp\left(\sum_{j=0}^{k} \beta_j(u_i, v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig}\right)\right) - \ln(y_i!)) \\
&= \sum_{i=1}^{n} \left(-\exp\left(\sum_{j=0}^{k} \beta_j(u_i, v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig}\right)\right) \\
&\quad + y_i \left(\sum_{j=0}^{k} \beta_j(u_i, v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig}\right) - \ln(y_i!))
\end{aligned} \tag{11}$$

Lokasi geografis berfungsi sebagai pembobot dalam model GWPRS dan berikutnya pembobot tersebut dimasukkan ke dalam persamaan (11) untuk mendapatkan model GWPRS yaitu:

$$\begin{aligned}\ln L^*(\beta_j(u_i,v_i),\gamma_g) = &\left(-\sum_{i=1}^{n}\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)\right.\\ &+\sum_{i=1}^{n}y_i\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)\\ &-\sum_{i=1}^{n}\ln(y_i!)\Bigg)\,W_{ij}(u_i,v_i)\end{aligned} \tag{12}$$

Selanjutnya yaitu persamaan (12) dideferensialkan terhadap $\beta_j(u_i,v_i)$ dan γ_g untuk memaksimumkan fungsi *ln likelihood.* Hasil dari pendiferensialan tersebut disamadengankan 0, sehingga didapatkan:

$$\begin{aligned}\frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_j(u_i,v_i)} = &-\sum_{i=1}^{n}x_{ij}W_{ij}(u_i,v_i)\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}\right.\\ &\left.+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)+\sum_{i=1}^{n}y_i x_{ij}W_{ij}(u_i,v_i)=0\\ \frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_j} = &-\sum_{i=1}^{n}x_{ig}W_{ij}(u_i,v_i)\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}\right.\\ &\left.+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)+\sum_{i=1}^{n}y_i x_{ig}W_{ij}(u_i,v_i)=0\end{aligned} \tag{13}$$

Proses penyelesaian pada persamaan (13) dapat diselesaikan dengan menggunakan iterasi Newton Raphson. Adapun prosedur dalam penaksiran parameter model GWPRS tersebut yaitu dengan menaksir $\beta_j(u_i,v_i)$ dan γ_g (Nakaya dkk. 2005). Penaksir parameter-parameter model GWPRS memiliki bentuk umum persamaan yaitu:

$$\begin{aligned}\left(\beta^{(t+1)}(u_i,v_i),\gamma^{(m+1)}\right) = &\left(\beta^{(t)}(u_i,v_i),\gamma^{(m)}\right)\\ &-\left(H^{(t)^{-1}}\left(\beta^{(t)}(u_i,v_i),\gamma^{(m)}\right)\right)\left(g^{t}\left(\beta^{(t)}(u_i,v_i),\gamma^{(m)}\right)\right)\end{aligned}$$

dengan:

$$g^t\left(\beta^{(t)}(u_i,v_i),\gamma^{(m)}\right)=\begin{bmatrix}\dfrac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_0(u_i,v_i)}\\ \dfrac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_1(u_i,v_i)}\\ \vdots\\ \dfrac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_{k*}(u_i,v_i)}\\ \dfrac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_{k*+1}}\\ \vdots\\ \dfrac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_k}\end{bmatrix}$$

$$H^{(t)^{-1}}\left(\beta^{(t)}(u_i,v_i),\gamma^{(m)}\right)$$
$$=$$
$$\begin{bmatrix}
\frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_0(u_i,v_i)\beta_0(u_i,v_i)} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_1(u_i,v_i)\beta_0(u_i,v_i)} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_{k*}(u_i,v_i)\beta_0(u_i,v_i)} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_{k*+1}\beta_0(u_i,v_i)} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_k\beta_0(u_i,v_i)}\\
\frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_0(u_i,v_i)\beta_1(u_i,v_i)} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_1(u_i,v_i)\beta_1(u_i,v_i)} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_{k*}(u_i,v_i)\beta_1(u_i,v_i)} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_{k*+1}\beta_1(u_i,v_i)} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_k\beta_1(u_i,v_i)}\\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\
\frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_0(u_i,v_i)\beta_{k*}(u_i,v_i)} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_1(u_i,v_i)\beta_{k*}(u_i,v_i)} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_{k*}(u_i,v_i)\beta_{k*}(u_i,v_i)} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_{k*+1}\beta_{k*}(u_i,v_i)} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_k\beta_{k*}(u_i,v_i)}\\
\frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_0(u_i,v_i)\gamma_{k*+1}} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_1(u_i,v_i)\gamma_{k*+1}} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_{k*}(u_i,v_i)\gamma_{k*+1}} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_{k*+1}\gamma_{k*+1}} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_k\gamma_{k*+1}}\\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\
\frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_0(u_i,v_i)\gamma_k} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_1(u_i,v_i)\gamma_k} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_{k*}(u_i,v_i)\gamma_k} & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_{k*+1}\gamma_k} & \cdots & \frac{\partial^2 \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_k\gamma_k}
\end{bmatrix}$$

Untuk setiap langkah iterasi ke-t berlaku:

$$\frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta_j(u_i,v_i)}=-\sum_{i=1}^{n}x_{ij}W_{ij}(u_i,v_i)\left(y_i-\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)\right)$$
$$=0$$

$$\frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma_g}=-\sum_{i=1}^{n}x_{ig}W_{ij}(u_i,v_i)\left(y_i-\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)\right)$$
$$=0$$

$$\frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\beta^T(u_i,v_i)\beta_j(u_i,v_i)}=-\sum_{i=1}^{n}x_{ij}W_{ij}(u_i,v_i)\left(x_{ij}\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)\right)$$
$$=0$$

$$\frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial\gamma^T\gamma_g}=-\sum_{i=1}^{n}x_{ig}W_{ij}(u_i,v_i)\left(x_{ig}\exp\left(\sum_{j=0}^{k}\beta_j(u_i,v_i)x_{ij}+\sum_{g=k^*+1}^{k}\gamma_g x_{ig}\right)\right)$$
$$=0$$

$$\frac{\partial \ln L^*(\beta_j(u_i,v_i),\gamma)}{\partial \gamma_g \beta_j(u_i,v_i)} = -\sum_{i=1}^{n} x_{ig} W_{ij}(u_i,v_i) \exp\left(\sum_{j=0}^{k} \beta_j(u_i,v_i)x_{ij} + \sum_{g=k^*+1}^{k} \gamma_g x_{ig}\right)$$
$$= 0$$

Penaksir parameter yang bersifat lokal dan bersifat global dapat diperoleh dengan mengulang prosedur iterasi tersebut untuk setiap titik regresi ke-i. Iterasi akan berhenti pada saat konvergen, dimana ketika nilai $\left\|\beta^{(t+1)}(u_i,v_i) - \beta^{(t)}(u_i,v_i)\right\| \leq \varepsilon$ dan $\left\|\gamma^{(t+1)} - \gamma^{(t)}\right\| \leq \varepsilon$.

1.6.8 Pengujian Signifikan Model GWPRS

Pengujian signifikan pada model GWPRS terdiri dari dua pengujian yaitu pengujian serentak atau bersamaan dan pengujian parsial atau individu pada parameter yang bersfat global dan bersifat lokal (Agresti, 2007).

a. Pengujian serentak parameter model GWPRS

Menurut Purhadi dan Yasin (2012), pengujian serentak bertujuan untuk menguji signifikans parameter pada variabel model GWPRS secara bersamaan. Pengujian simultan atau serentak pada parameter variabel model GWPRS mencakup dua jenis, yaitu pengujian pada parameter variabel yang bersifat global dan pengujian pada parameter variabel yang bersifat lokal. Adapun hipotesis yang akan digunakan dalam pengujian simultan pada parameter global model GWPRS adalah sebagai berikut:

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_k = 0$

H_1 : minimal terdapat satu $\beta_g \neq 0, \quad g = 1,2,\dots,k$

Statistik uji yang dapat digunakan yaitu uji *likelihood ratio*, yaitu:

$$G_1 = -2\ln\left[\frac{\ell(\widehat{\Omega}_g)}{\ell(\widehat{\omega}_g)}\right] = -2\left(\ell(\widehat{\Omega}_g) - \ell(\widehat{\omega}_g)\right) \tag{14}$$

dimana $\widehat{\Omega}_g = \{\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k\}$ merupakan *log-likelihood* pada model yang tidak terdapat variabel bebas serta $\widehat{\omega}_g = \hat{\beta}_0$ merupakan *log-likelihood* pada model yang terdapat keseluruhan variabel bebas. Jadi dengan taraf kepercayaan sebesar $\alpha =$ 5%, tolak H_0, apabila $G_1 > \chi^2_{df;\alpha}$ dengan $df = \left[\sum_{j=1}^{k} \beta_g\right] - 1$.

Selanjutnya yaitu pengujian simultan atau serentak pada parameter variabel yang bersifat lokal model GWPRS. Adapun dengan hipotesis yang digunakan yaitu:

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

H_1 : minimal terdapat satu $\beta_l(u_i,v_i) \neq 0, \quad l = 0,1,2,\dots,k$

Statistik uji yang dapat digunakan yaitu uji *likelihood ratio*, dimana:

$$G_2 = -2\ln\left[\frac{\ell(\widehat{\Omega}_l)}{\ell(\widehat{\omega}_l)}\right] = -2\left(\ell(\widehat{\Omega}_l) - \ell(\widehat{\omega}_l)\right) \tag{15}$$

dengan $\widehat{\Omega}_l = \{\hat{\beta}_0(u_i,v_i), \hat{\beta}_1(u_i,v_i), \hat{\beta}_2(u_i,v_i), \dots, \hat{\beta}_k(u_i,v_i)\}, \quad i = 1,2,3,\dots,k$ merupakan *log-likelihood* pada model yang tidak terdapat variabel bebas dan $\widehat{\omega}_l = \{\hat{\beta}_0(u_i,v_i), \quad i = 1,2,3,\dots,k\}$ merupakan *log-likelihood* pada model yang terdapat keseluruhan variabel bebas. Dengan taraf kepercayaan sebesar $\alpha = 5\%$, tolak H_0, apabila $G_2 > \chi^2_{df;\alpha}$ dengan $df = \left[\sum_{j=1}^{k} \beta_l(u_i,v_i)\right] - 1$ (Li dkk, 2013)

b. Pengujian Parsial Parameter Model GWPRS

Menurut Purhadi dan Yasin (2012), pengujian parsial dilakukan untuk mengidentifikasi variabel yang bersifatr global serta bersifat lokal yang memiliki pengaruh signifikan secara parsial atau individu terhadap variabel respon pada model GWPRS. Pengujian ini mencakup pengujian parsial untuk parameter variabel yang bersifat global serta variabel yang bersifat lokal. Pada pengujian parsial parameter variabel yang bersifat global maka digunakan hipotesis sebagai berikut:

H_0 : $\beta_g = 0$ untuk setiap $g = 1,2,\dots,k$

H_1 : minimal terdapat satu $\beta_g \neq 0$ untuk setiap $g = 1,2,\dots,k$

Statistik uji yang digunakan yaitu:

$$t_{g\_hit} = \frac{\hat{\beta}_g}{se(\hat{\beta}_g)} \quad \sim \ t_{\alpha/2}\,; n-(p+1) \tag{16}$$

dengan $\hat{\beta}_g$ nilai estimasi parameter dan $se(\hat{\beta}_g)$ adalah nilai *standard error* dari nilai estimasi parameter. Dimana:

$$se(\hat{\beta}_g) = \sqrt{cov(\hat{\beta}_g)}$$

$$cov(\hat{\beta}_g) = \boldsymbol{C}\,\boldsymbol{A}^{-1}\boldsymbol{C}^T$$

Kriteria pengujiannya yaitu dengan taraf kepercayaan sebesar α, tolak H_0 apabila $|t_{g\_hit}| > t_{\alpha/2}\,; n-(p+1)$ yang dapat disimpulkan bahwa, pengaruh pada variabel prediktor ke-g terhadap variabel respon (y) signifikan pada model.

Uji hipotesisi selanjutnya yaitu guna mengetahui variabel yang bersifat lokal berpengaruh signifikan terhadap variabel respon dalam model GWPRS. Adapun hipotesis yang digunakan yaitu:

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

H_1 : $\beta_l(u_i,v_i) \neq 0$

Dimana $l = 1,2,\dots,k$ dan $i = 1,2,\dots,n$. Statistik uji yang dapat digunakan dalam pengujian hipotesis tersebut yaitu:

$$t_{l\_hit} = \frac{\hat{\beta}_l(u_i,v_i)}{se\left(\hat{\beta}_l(u_i,v_i)\right)} \quad \sim \ t_{\alpha/2}\,; n-(p+1) \tag{17}$$

dengan $\hat{\beta}_l(u_i,v_i)$ nilai estimasi parameter dan $se\left(\hat{\beta}_l(u_i,v_i)\right)$ adalah nilai *standard error* dari nilai estimasi parameter. Dimana:

$$se\left(\hat{\beta}_l(u_i,v_i)\right) = \sqrt{cov\left(\hat{\beta}_l(u_i,v_i)\right)}$$

$$cov\left(\hat{\beta}_j(u_i,v_i)\right) = \boldsymbol{C}(u_i,v_i)\boldsymbol{A}(u_i,v_i)^{-1}\boldsymbol{C}(u_i,v_i)^T$$

Kriteria pengujiannya yaitu dengan taraf kepercayaan sebesar $\alpha = 5\%$, tolak H_0 apabila $|t_{l\_hit}| > t_{\alpha/2}\,; n-(p+1)$ yang dapat disimpulkan bahwa pengaruh pada variabel prediktor ke-l terhadap variabel respon (y) pada lokasi (u_i,v_i) signifikan dalam model (Lu dkk, 2015).

1.6.9 Pencilan

Istilah pencilan merujuk pada pengamatan pola data secara keseluruhan persamaan regresi yang dihasilkan tidak mengikuti bentuk pola umum. Pencilan merupakan pengamatan yang dapat dikatakan jauh dari pusat data dan dapat memiliki dampak yang signifikan pada koefisien regresi. Pencilan dapat terjadi akibat dari kesalahan dalam pencatatan data, analisis, kesalahan pengukuran, ataupun berbagai kesalahan yang lainnya. Meskipun pengamatan pencilan dapat memengaruhi estimasi parameter, pencilan juga dapat memberikan informasi penting yang diperlukan. Oleh karena itu, keputusan untuk menghapus pencilan harus didasarkan pada alasan yang kuat (Arimie dkk., 2020).

Pencilan tidak bisa dihilangkan begitu saja dari pengamatan dikarenakan terkadang pencilan dapat memberikan informasi yang tidak dapat diperoleh dari titik data lainnya. Pencilan secara tidak langsung merupakan hasil observasi yang menyimpang dari pola yang terbentuk oleh sebagian besar data (Ghozali, 2011). Pencilan tersebut dapat diabaikan jika setelah diteliti terbukti bahwa pencilan disebabkan oleh kesalahan dalam menyiapkan peralatan atau kesalahan dalam pencatatan data (Draper & Smith, 1992). Menurut Arimie dkk. (2020), beberapa metode yang dapat digunakan untuk mengidentifikasi adanya pencilan yang berpengaruh terhadap koefisien regresi antara lain:

a. Metode Grafis (*Boxplot*)

Keunggulan dari metode grafis adalah kemudahannya dalam pemahaman, karena data disajikan dalam bentuk grafis atau gambar dan tidak memerlukan perhitungan yang kompleks. Namun, kelemahannya adalah keputusan mengenai apakah suatu data dianggap sebagai pencilan atau tidak tergantung pada penilaian peneliti karena metode ini hanya bergantung pada visualisasi. Salah satu metode grafis yang digunakan dalam penelitian ini adalah *Boxplot*.

Metode *Boxplot* menggunakan nilai kuartil dan jangkauan untuk mendeteksi pencilan. Kuartil 1, 2, dan 3 akan membagi data yang telah diurutkan sebelumnya menjadi empat bagian. Jaungkauan (*IQR*) didefinisikan sebagai selisih kuartil 1 terhadap kuartil 3. Data-data yang merupakan pecilan yaitu nilai yang kurang dari Dalam metode Boxplot untuk mendeteksi pencilan maka dapat menggunakan nilai kuartil dan jangkauan. Kuartil 1, 2, dan 3 membagi data yang telah diurutkan menjadi empat bagian. Jangkauan *Interquartile Rang* (IQR) didefinisikan sebagai selisih antara kuartil 1 dan kuartil 3 atau $IQR = Q_3 - Q_1$. Data yang dianggap sebagai pencilan adalah nilai-nilai yang kurang dari $1{,}5 \times IQR$ terhadap kuartil 1 dan nilai-nilai yang lebih dari $1{,}5 \times IQR$ terhadap kuartil 3.

b. Metode DfFITS

Selain menggunakan metode grafis, metode lain yang dapat digunakan untuk mendeteksi pencilan dalam penelitian ini adalah dengan menggunakan metode *Standardized DfFITS* atau *DfFITS* (*Difference fitted value FITS*). Metode *DfFITS* menunjukkan nilai perubahan dalam predikdi harga ketika suatu kasus atau kondisi tertentu dikeluarkan yang sudah distandardisai. Perhitungan *DfFITS* dapat dilakukan dengan cara berikut:

$$DfFITS_i = t_i \frac{h_{ii}}{1-h_{ii}}^{\frac{1}{2}} \quad (18)$$

dimana t_i merupaka *studentized deleted* pada kasus ke-i serta h_{ii} merupakan nilai *leverage* pada kasus ke-i, dengan

$$t_i = e_i \sqrt{\frac{n-p-1}{JKG(1-h_{ii}-e_i^2)}}$$

dimana e_i merupakan nilai residual ke-i, serta JKG merupakan jumlah kuadrat galat yang terdapat pada matriks yaitu:

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \quad (19)$$

dengan matriks $\mathbf{H}$ merupakan matriks yang berukuran $n \times n$.
Adapun elemen diagonal h_{ii} pada matriks didapatkan secara langsung dari:

$$h_{ii} = \mathbf{X}_i(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_i{}^T \quad (20)$$

dengan $\mathbf{X}_i$ matriks $n \times p$, $(\mathbf{X}^T\mathbf{X})^{-1}$ adalah matriks yang berukuran $p \times p$, dan $\mathbf{X}_i{}^T$ merupakan matriks yang berukuran $p \times n$.

Data yang memiliki nilai *absolute DfFITS* lebih besar dari nilai $2\sqrt{\frac{p}{n}}$, maka diidentifikasikan sebagai pencilan, dimana p merupakan banyaknya variabel prediktor serta n merupakan banyaknya observasi (Zhang & Xue, 2020).

1.6.10 *GM-estimator*

Regresi *robust* adalah metode regresi yang diterapkan pada saat terdapat beberapa pencilan yang dapat mempengaruhi hasil analisi regresi atau ketika distribusi residual tidak normal. Metode ini sangat penting untuk menganalisis data yang terpengaruh oleh pencilan, sehingga model yang dihasilkan menjadi *rubost* atau kekar terhadap pengaruh pencilan (Drapper dan Smith, 1998). Menurut Chen (2002), regresi *robust* mempunyai beberapa metode untuk mengestimasi salah satunya yaitu metode *GM-estimator*. *GM-estimator* adalah metode yang dikembangkan dari metode *M-estimator,* dimana pada saat metode *M-estimator* kurang resisten terhadap pencilan yang terdapat variabel x_i. Ide dasar yang menjadi latar belakang metode *GM-estimator* yaitu untuk membatasi pengaruh pencilan terhadap variabel x_i dengan menggunakan fungsi pembobot (η) dan hanya bergantung pada x_i. Metode *GM-estimator* secara umum dapat dituliskan berikut:

$$GM = \sum_{i=1}^{n} \eta(x_i)\rho(\varepsilon_i{}^*) \quad (21)$$

dimana $\eta(x_i)$ adalah fungsi pembobot untuk variabel x_i serta $\rho(\varepsilon_i{}^*)$ merupakan fungsi objektif dimana nilai $\varepsilon_i{}^* = \frac{\varepsilon_i}{\hat{\sigma}}$. Nilai $\hat{\sigma}$ adalah skala *robust* yang akan diestimasi. Estimasi pada $\hat{\sigma}$ yang akan digunakan yaitu:

$$\hat{\sigma} = \frac{MAD}{0.6745} = \frac{med\{|\varepsilon_i - med(\varepsilon_i)|\}}{0{,}6745}$$

Pemilihan konstanta 0.6745 akan menghasilkan $\hat{\sigma}$ estimator yang mendekati tak bias dari $\hat{\sigma}$ ketika n besar serta residual berdistribusi normal. Pembuktian bahwa $\hat{\sigma} = \frac{MAD}{S}$ maka $S = 0.6745$ dapat diperoleh dari perhitungan sebagai berikut:

$$P\left(|Z| \leq \frac{MAD}{\hat{\sigma}}\right) = \frac{1}{2}$$

$$P\left(-\frac{MAD}{\hat{\sigma}} \leq Z \leq \frac{MAD}{\hat{\sigma}}\right) = \frac{1}{2}$$

$$P\left(Z \leq \frac{MAD}{\hat{\sigma}}\right) = \frac{1}{4}$$

$$P\left(Z > \frac{MAD}{\hat{\sigma}}\right) = \frac{3}{4}$$

$$\frac{MAD}{\hat{\sigma}} = 0.6745$$

$$\frac{MAD}{0.6745} = \frac{MAD}{0.6745}$$

$$S = 0.6745$$

sehingga demikian persamaan (21) dapat dituliskan menjadi sebagai berikut:

$$GM = \sum_{i=1}^{n} \eta(x_i)\rho\left(\frac{\varepsilon_i}{\hat{\sigma}}\right) \tag{22}$$

Dimisalkan $\rho({\varepsilon_i}^*)$ memiliki turunan $\frac{\partial\rho({\varepsilon_i}^*)}{\partial({\varepsilon_i}^*)} = \psi({\varepsilon_i}^*, \beta)$, maka untuk meminimumkan persamaan (22) akan digunakan turunan parsial pertama dari $\rho({\varepsilon_i}^*)$ terhadap β setelah itu disamadengankan nol sehingga menghasilkan persamaan sebagai berikut:

$$\frac{\partial\rho({\varepsilon_i}^*)}{\partial({\varepsilon_i}^*)} = \frac{\partial\rho({\varepsilon_i}^*)}{\partial({\varepsilon_i}^*)} . \frac{\partial\rho({\varepsilon_i}^*)}{\partial\beta_l(u_i, v_i)} = 0$$

$$= \sum_{i=1}^{n} x_i\eta(x_i)\psi\left(\frac{{\varepsilon_i}^*}{\hat{\sigma}}\right) = 0 \tag{23}$$

dimana ψ adalah fungsi pengaruh yang akan digunakan untuk menghasilkan fungsi pembobot *robust*. Fungsi pembobot *GM-estimator* $\eta(x_i)$ secara umum didefinisikan berdasarkan nilia *leverge* di mana $\eta(x_i) = \sqrt{1 - h_{ii}}$. Nilia *leverge* (h_{ii}) dibatasi pada $0 \leq h_{ii} \leq 1$ yang dihasilkan dari matriks diagonal pada *hat* matriks simetris yang berukuran $n \times n$ pada persamaan (20). Berdasarkan persamaan (23) Draper dan Smith (1998) memberikan solusi dengan fungsi pembobot $\boldsymbol{W_i}^*$ didefinisikan berikut:

$$W_i^* = W_i({\varepsilon_i}^*) = \frac{\psi\left(\frac{\varepsilon_i}{\hat{\sigma}}\right)}{\frac{\varepsilon_i}{\hat{\sigma}}} \tag{24}$$

Setelah itu kedua ruas pada persamaan (24) dikalikan dengan $\left(\frac{\varepsilon_i}{\hat{\sigma}}\right)$ sehingga didapatkan:

$$W_i^*\left(\frac{\varepsilon_i}{\hat{\sigma}}\right) = \psi\left(\frac{\varepsilon_i}{\hat{\sigma}}\right)$$

maka dari itu, persamaan (16) dapat dituliskan sebagai berikut:

$$\sum_{i=1}^{n} x_i \eta(x_i) W_i^* \left(\frac{\varepsilon_i}{\hat{\sigma}}\right) = 0 \tag{25}$$

Persamaan (25) dapat diselesaikan dengan menggunakan metode iterasi yaitu metode IRLS atau *Iteratively Reweighted Least Square.* Denga menggunakan metode IRLS maka prosedur estimasinya dapat dilakukan dengan langkah-langkah diantaranya yaitu:

1) Mengetimasi parameter regresi dengan menggunakan *least square*, sehingga didapatkan $\hat{y}_i^{\,0}$ serta didapatkan residual $\varepsilon_i^{\,0}$ dengan $\varepsilon_i^{\,0} = y_i - \hat{y}_i^{\,0}$, $i = 1,2,3,\dots,n$ yang dibutuhkan sebagai nilai awal.
2) Menentukan estimator skala *robust* $\hat{\sigma}^0$yang berdasarkan nilai residual yang didapatkan dari langkah (1) serta fungsi pembobot awal $W_i^{*0} = \frac{\psi(\varepsilon_i^{*0})}{\varepsilon_i^{*0}}$, dimana W_i^{*0} dihitung menggunakan fungsi pembobot *Tukey Bisquare.*
3) Mencari nilai estimasi parameter pada iterasi pertama yaitu $m = 1$ dengan menggunakan nilai residual yaitu ε_i^{*0} serta W_i^{*0}.
$$\hat{\beta}^{m+1} = (\boldsymbol{X}^T \boldsymbol{\eta} W_i^{*m} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{\eta} W_i^{*m} \ln(y_i)$$
4) Menghitung $\sum_{i=1}^{n} |y_i - \hat{y}_i^{(1)}|$ atau $\sum_{i=1}^{n} |\varepsilon_i^{(1)}|$.
5) Mengulang kemabali langkah (2) sampai dengan langkah (4) sehingga didapatkan nilai $\sum_{i=1}^{n} |\varepsilon_i^{(m)}|$ yang konvergen kesuatu nilai yaitu pada saat selisih antara $\hat{\beta}^m$ dan $\hat{\beta}^{m+1}$ nilainya mendekati nol.
$$|\hat{\beta}^m - \hat{\beta}^{m+1}| < \varepsilon$$
dimana ε merupakan nilai bilangan positif kecil.

1.6.11 Fungsi Objektif

Fungsi yang digunakan dalam menentukan fungsi pembobot dalam regresi *robust* yaitu fungsi objektif. Dimana ketika fungsi objektif ini diturunkan terhadap e_i, maka akan dihasilkan suatu fungsi yang dikenal sebagai fungsi pengaruh. Pembagian fungsi pengaruh dengan $\psi(e_i)$ akan menghasilkan fungsi pembobot yang digunakan dalam perhitungan pembobot *robust*. Salah satu fungsi pembobot yang diterapkan adalah *Tukey Bisquare*. Berikut adalah fungsi objektif yang diberikan:

$$\rho'(\varepsilon_i) = \begin{cases} \dfrac{c^2}{\sigma}\left\{1 - \left[1 - \left(\dfrac{\varepsilon_i}{c}\right)^2\right]^3\right\} & , |\varepsilon_i| \le c \\ \dfrac{c^2}{\sigma} & , |\varepsilon_i| > c \end{cases}$$

dengan:

$$\rho'(\varepsilon_i) = \psi(\varepsilon_i) = \frac{\partial(\rho(\varepsilon_i))}{\partial(\varepsilon_i)} = \begin{cases} \varepsilon_i \left[1 - \left(\dfrac{\varepsilon_i}{c}\right)^2\right]^2 & , |\varepsilon_i| \le c \\ 0 & , |\varepsilon_i| > c \end{cases}$$

Setelah diperoleh nilai $\rho'(e_i)$, maka diperoleh nilai dari fungsi pembobot yaitu:

$$W_i^* = W_i^*(\varepsilon_i) = \frac{\psi(\varepsilon_i)}{\varepsilon_i} = \begin{cases} \left[1 - \left(\dfrac{\varepsilon_i}{c}\right)^2\right]^2 & , |\varepsilon_i| \le c \\ 0 & , |\varepsilon_i| > c \end{cases}$$

dengan $c = 4.685$ merupakan konstanta yang menghasilkan efisiensi tinggi ketika residual berdistribusi normal serta dapat memberikan perlindungan terhadap pencilan. (Fox, 2002).

1.6.12 Kematian Ibu

Angka Kematian Ibu (AKI) adalah jumlah kematian yang dialami oleh perempuan selama kehamilan atau dalam waktu 42 hari setelah terminasi kehamilan, yang disebabkan oleh kehamilan itu sendiri atau pengelolaannya, dan bukan oleh faktor lain. AKI merupakan salah satu indikator yang digunakan untuk mencerminkan kualitas pelayanan kesehatan di suatu wilayah. Oleh karena itu, pemerintah daerah, khususnya di Jawa Timur, seharusnya memberikan perhatian lebih terhadap Angka Kematian Ibu. Penanganan Angka Kematian Ibu dapat dilakukan dengan memperhatikan dan mengatasi faktor-faktor yang menyebabkan kematian (BPS, 2021). Untuk menghitung angka kematian ibu, dapat dilakukan dengan cara:

$$AKI = \frac{Dhamil}{JHL} \times 1000 \ldots$$

dimana:

$Dhamil$: jumlah kematian ibu selama tahap kehamilan atau saat melahirkan

JHL : jumlah kelahiran hidup

Adapun faktor-faktor yang mempengaruhi kematian ibu, dapat dilihat dari penyebab kematian serta faktor risiko yang meliputi: faktor pelayanan kesehatan rujukan yaitu cara masuk ke dalam rumah sakit dan metode persalinan, faktor reproduksi yaitu umur ibu dan paritas, serta faktor sosial ekonomi yaitu pendidikan ibu dan jenis pekerjaan. Berikut merupakan penjelasan mengenai faktor-faktor yang diduga berpengaruh terhadap angka kematian ibu (Royston, 1994):

1) Persalinan komplikasi ditangani tenaga kesehatan

   Persalinan yang mengalami komplikasi dan ditangani oleh tenaga kesehatan adalah persalinan yang dilakukan oleh tenaga terampil yang memenuhi standar seperti bidan, dokter, serta tenaga paramedis di fasilitas kesehatan. Persentase ibu hamil yang mengalami komplikasi mencakup kondisi seperti hemoglobin di bawah 8 g%, tekanan darah tinggi (sistolik > 140 mmHg, diastolik > 90 mmHg), ketuban pecah dini, pendarahan pervaginam, edema nyata, eklampsia, letak lintang pada usia kehamilan lebih dari 32 minggu, letak sungsang pada primigravida, infeksi berat atau sepsis, serta persalinan prematur.

2) Pelayanan kesehatan ibu hamil (K1) dan (K4)

   Kunjungan ibu hamil dengan pelayan kesehatan program K1 dan K4 adalah perbandingan antara jumlah ibu hamil yang menerima pelayanan K1 dan K4 di suatu daerah dengan total jumlah ibu hamil di wilayah tersebut. Pelayanan K1 mengacu pada ibu hamil yang pertama kali mendapatkan pelayanan antenatal dalam tiga bulan pertama kehamilan oleh tenaga kesehatan. Sementara itu, pelayanan K4 adalah ibu hamil yang menerima pelayanan antenatal sesuai standar setidaknya empat kali, dengan distribusi pelayanan yang dianjurkan minimal satu kali pada triwulan pertama, satu kali pada triwulan kedua, dan satu kali pada triwulan ketiga.

3) Ibu hamil memperoleh tablet Fe1 dan Fe3
   Ibu hamil yang menerima Fe1 dan Fe3 mendapatkan suplemen tambahan zat besi sebagai bentuk upaya dalam pencegahan dan penanggulangan anemia akibat kekurangan gizi besi.
4) Wanita menikah dibawah 17 tahun
   Wanita yang menikah dengan tujuan membentuk keluarga atau rumah tangga yang bahagia pada usia dibawah 17 tahun.
5) Ibu nifas memperoleh vitamin A
   Ibu hamil pada masa nifas yang mendapatkan kapsul vitamin A sebanyak dua kali (2 x 24 jam) bertujuan untuk memperkecil resiko kelainan bahkan resiko kematian pada ibu nifas.

1.6.13 Kerangka Konsep

Menurut Notoatmodjo (2005), kerangka konsep penelitian merupakan hubungan atau keterkaitan antara berbagai konsep atau variabel yang akan diamati melalui penelitian tersebut. Dalam menyusun kerangka konsep, peneliti perlu memahami variabel yang akan diukur karena kerangka konsep ini memberikan dasar konseptual untuk penelitian. Selain itu, kerangka konsep juga berfungsi untuk mengidentifikasi hubungan antara variabel yang dianggap relevan dengan studi mengenai masalah tertentu. Oleh karena itu, pemahaman mengenai makna variabel dan berbagai jenis variabel yang ada sangatlah penting. Berikut kerangka konsep pada penelitian ini:

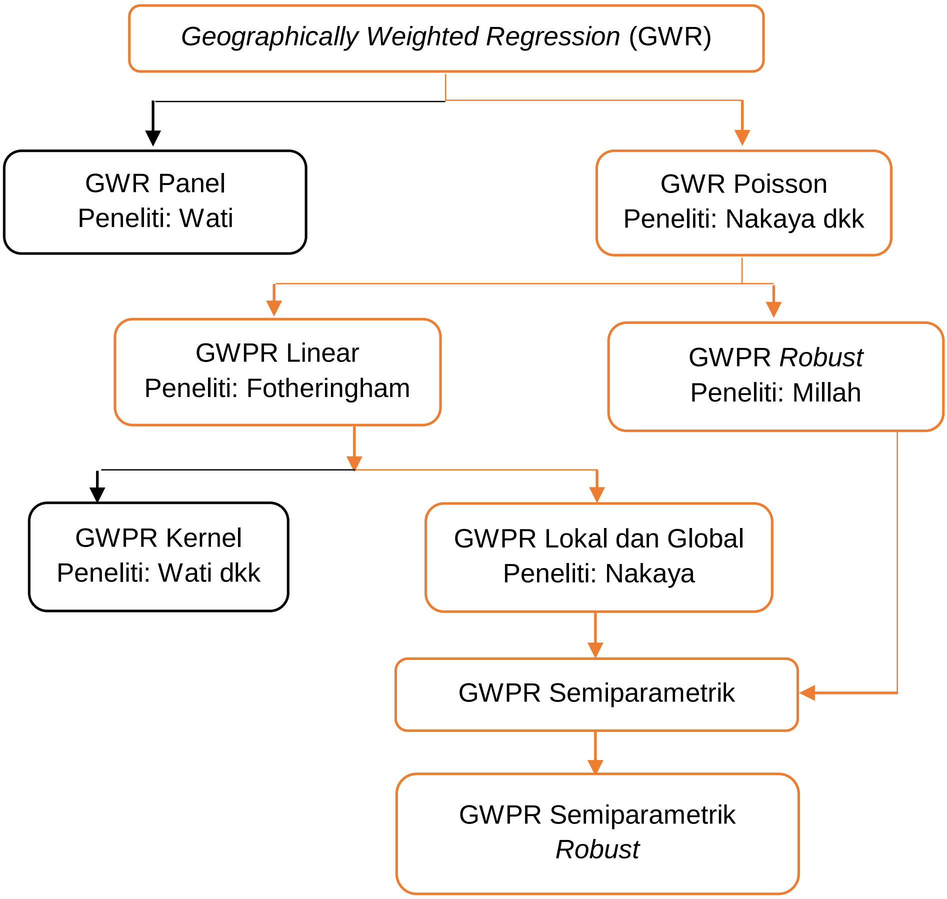


Figure 1. Kerangka Konsep

BAB II
METODOLOGI PENELITIAN

2.1 Sumber Data

Pada penelitian ini data yang akan digunakan merupakan data sekunder, yaitu data Angka Kematian Ibu di Jawa Timur tahun 2021. Data tersebut diperoleh dari website resmi Dinas Kesahatan (Dinkes) Provinsi Jawa Timur (dinkes.jawatimurprov.go.id). Data didapatkan dari Profil Dinas Kesehatan Provinsi Jawa Timur tahun 2022 yang berisi data tahun 2021. Unit observasi pada penelitian ini berjumlah 38 yang terdiri dari 29 kabupaten dan 9 kota yang terdapat di Provinsi JawaTimur.

2.2 Variabel Penelitian

Variabel penelitian pada penelitian ini dibagi menjadi dua diantaranya variabel respon (dependen) (y) dan variabel prediktor (independen) (x). Dimana variabel respon (y) dalam penelitian ini terdiri dari satu variabel yaitu angka kematian ibu, sedangkan variabel prediktor (x) terdiri dari lima variabel yaitu faktor-faktor yang diduga mempengaruhi angka kematian ibu seperti pada Tabel 1 berikut:

Table 1. Variabel Penelitian

Variabel Penelitian	Kode	Nama Variabel
Variabel Respon	(y)	Angka Kematian Ibu
Variabel Prediktor	(x_1)	Persentase Persalinan komplikasi ditangani tenaga kesehatan
	(x_2)	Persentase Pelayanan kesehatan ibu hamil (K1) dan (K4)
	(x_3)	Persentase Ibu hamil memperoleh tablet Fe1 dan Fe3
	(x_4)	Persentase Wanita menikah dibawah 17 tahun
	(x_5)	Persentase Ibu nifas memperoleh vitamin A

2.3 Metode Analisis

2.3.1 Estimasi Parameter Model GWPRS yang Mengandung Pencilan

Dalam mengestimasi parameter model GWPRS yang mengandung pencilan langkah-langkahnya adalah sebagai berikut:

a. Mengestimasi parameter yang bersifat global dengan menggunakan metode *GM-estimator* langkah-langkahnya yaitu:
   1) Menentukan model regresi bersifat global yang mengandung pencilan.
   2) Mengestimasi parameter model dengan menggunakan metode *GM-estimator*, dimana langkah-langkahnya yaitu:
      (a) Mengestimasi parameter β_g dengan OLS.
      (b) Mencari nilai fungsi pembobot W_i.

(c) Mencari estimasi baru dengan menggunakan WLS.

(d) Melakukan penyelesaian estimasi dengan metode IRLS (*Iteratively Reweighted Least Square*), sebagai berikut:

(1) Menentukan ${\hat{\beta}_g}^0$ sebagai estimator awal

(2) Mencari fungsi pembobot baru berdasarkan estimator awal.

(3) Mencari nilai ${\hat{\beta}_g}^{m+1}$ sebagai estimator yang konvergen.

b. Mengestimasi parameter yang bersifat lokal dengan menggunakan metode *GM-estimator* langkah-langkahnya yaitu:

1) Menentukan model regresi lokal yang mengandung pencilan.

2) Mengestimasi parameter model dengan menggunakan metode *GM-estimator* dimana langkah-langkahnya yaitu:

(a) Mengestimasi parameter $\beta_l(u_i, v_i)$ dengan OLS.

(b) Mencari nilai fungsi pembobot W_i.

(c) Mencari estimasi baru dengan menggunakan WLS.

(d) Melakukan penyelesaian estimasi dengan metode IRLS (*Iteratively Reweighted Least Square*), dengan cara sebagai berikut:

(1) Menentukan $\beta_l(u_i, v_i)$ sebagai estimator awal.

(2) Mencari fungsi pembobot baru berdasarkan estimator awal.

(3) Mencari nilai ${\hat{\beta}_l}^{m+1}$ sebagai estimator yang konvergen

2.3.2 Penerapan Pada Angka Kematian Ibu di Provinsi Jawa Timur

Selanjutnya yaitu menerapkan model GWPRS yang mengandung pencilan pada data Angka Kematian Ibu di Provinsi Jawa Timur tahun 2021:

1. Melakukan analisis deskriptif data yang digunakan sebagai gambaran awal guna mengetahui keadaan angka kematian ibu.
2. Melakukan pengujian asumsi pada data.
3. Mendeteksi adanya pencilan pada data.
4. Menganalisis data dengan menggunakan model GWPRS pada data yang mengandung pencilan.
5. Membuat interpretasi dan kesimpulan

2.4 Diagram Alir

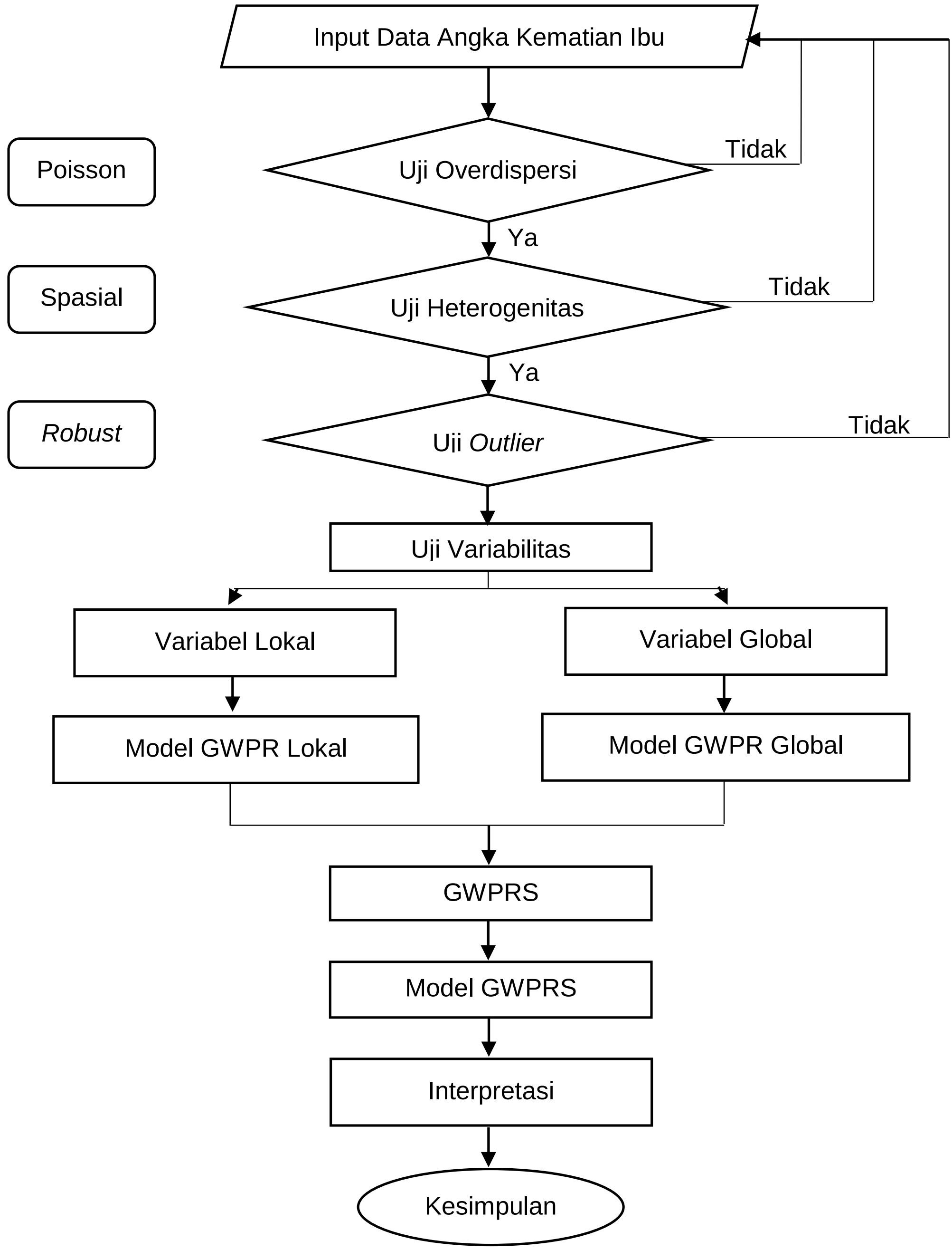


Figure 2. Diagram Alir