PEMODELAN *GEOGRAPHICALLY WEIGHTED REGRESSION* (GWR) DENGAN FUNGSI PEMBOBOT KERNEL GAUSSIAN DAN BI-SQUARE

Skripsi

disajikan sebagai salah satu syarat

untuk memperoleh gelar Sarjana Sains

Program Studi Matematika

oleh

Nurul Lutfiani

4111412072

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI SEMARANG

2017

PERNYATAAN

Dengan ini saya menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya yang diterbitkan oleh orang lain, kecuali yang secara tertulis dirujuk dalam skripsi ini dan disebutkan dalam daftar pustaka.

Semarang, 15 Agustus 2017

Nurul Lutfiani

NIM. 4111412072

PENGESAHAN

Skripsi yang berjudul

Pemodelan Geographically Weighted Regression (GWR) dengan Fungsi Pembobot Kernel Gaussian dan Bi-square

Disusun oleh

Nama : Nurul Lutfiani

NIM : 4111412072

Telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA Unnes pada

Hari : Jum'at

Tanggal : 11 Agustus 2017

Panitia:

Ketua

Prof. Dr. Zaenuri, S.E., M.Si. Akt.
NIP. 196412231988031001

Sekretaris

Drs. Arief Agoestanto, M.Si
NIP. 196807221993031005

Ketua Penguji

Dra. Sunarmi, M.Si.
NIP. 195506241988032001

Anggota Penguji

Pembimbing Utama

Drs. Sugiman, M.Si.
NIP. 196401111989011001

Anggota Penguji

Pembimbing Pendamping

Dr. Scolastika Mariani, M.Si.
NIP. 196502101991022001

MOTTO DAN PERSEMBAHAN

MOTTO

- Allah akan meninggikan derajat orang-orang yang beriman diantara kamu dan orang-orang yang memiliki ilmu pengetahuan (Al-Mujadillah : 12)
- Kebaikan tidak bernilai selama diucapkan, akan tetapi bernilai sesudah dikerjakan
- Lakukan yang terbaik, bersikaplah yang baik maka kau akan menjadi orang yang terbaik

PERSEMBAHAN

- Untuk Allah SWT, Tuhan Semesta Alam
- Untuk Dosen Jurusan Matematika
- Untuk Dosen Pembimbing, Drs. Sugiman, M.Si. dan Dr. Scolastika Mariani, M.Si.
- Untuk kedua orang tua tercinta Abdul Gofur dan Umi Farida
- Untuk adikku, Saifuddin Rizki
- Untuk keluarga besar
- Untuk sahabatku Ratih Permatasari, Anika Liansari Dewi, Wahyu Zuli Astutik
- Untuk teman-teman matematika 2012

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pemodelan *Geographically Weighted Regression* (GWR) dengan Fungsi Pembobot Kernel Gaussian dan Kernel Bi-square”.

Penulisan skripsi ini dapat terselesaikan karena adanya bimbingan, bantuan, dan dukungan dari berbagai pihak baik secara langsung maupun tidak langsung. Oleh karena itu, penulis mengucapkan terimakasih kepada:

1. Prof. Dr. Fathur Rokhman, M. Hum., Rektor Universitas Negeri Semarang.
2. Prof. Dr. Zaenuri, S.E., M.Si., Akt., Dekan FMIPA Universitas Negeri Semarang.
3. Drs. Arief Agoestanto, M. Si., Ketua Jurusan Matematika FMIPA Universitas Negeri Semarang.
4. Drs. Mashuri, M.Si. koordinator Program Studi Matematika FMIPA Universitas Negeri Semarang serta selaku dosen wali sekaligus orang tua yang telah memberikan arahan dan bimbingannya selama masa kuliah hingga selesai.
5. Drs. Sugiman, M.Si., pembimbing pertama yang telah memberikan bimbingan, motivasi, dan pengarahan sehingga skripsi ini dapat terselesaikan.
6. Dra. Scolastika Mariani, M.Si., pembimbing pendamping yang telah memberikan bimbingan, motivasi, dan pengarahan sehingga skripsi ini dapat terselesaikan.
7. Dra. Sunarmi, M.Si., Dosen penguji yang telah memberikan inspirasi, kritik, saran, dan motivasi kepada penulis, sehingga penulis dapat menyelesaikan skripsi.

8. StaffDosen Matematika Universitas Negeri Semarang yang telah membekali penulis dengan berbagai ilmu selama mengikuti perkuliahan sampai akhir penulisan skripsi ini.
9. Staff Tata Usaha Universitas Negeri Semarang yang telah membantu penulis selama mengikuti perkuliahan dan penulisan skripsi ini.
10. Orangtua dan keluarga tercinta yang senantiasa mendoakan serta memberikan dukungan baik secara moral maupun spiritual.
11. Sahabat-sahabat penulis yang telah memberikan banyak motivasi, kritik, usulan, yang menjadikan terselesaikannya penulisan skripsi ini.
12. Mahasiswa matematika angkatan 2012 yang telah memberikan dorongan dan motivasi.
13. Semua pihak yang telah membantu terselesaikannya penulisan skripsi ini.

Hanya ucapan terimakasih dan doa, semoga apa yang telah diberikan tercatat sebagai amal baik dan mendapatkan balasan dari Allah SWT.

Semoga skripsi ini bisa membawa manfaat bagi penulis sendiri khususnya dan bagi para pembaca pada umumnya.

ABSTRAK

Lutfiani, Nurul. 2017. *Pemodelan Geographically Weighted Regression (GWR) dengan Fungsi Pembobot Kernel Gaussian dan Bi-square*. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang. Pembimbing Utama Drs. Sugiman, M.Si. dan Pembimbing Pendamping, Dra. Scolastika Mariani, M.Si.

Kata Kunci : Regresi Spasial, Geographically Weighted Regression, Kernel Gaussian, Kernel Bi-square

Model spasial *Geographically Weighted Regression* (GWR) adalah salah satu metode statistika yang dapat digunakan untuk menganalisis faktor risiko secara spasial dengan pendekatan titik. Dalam penelitian ini, fungsi pembobot yang digunakan untuk model GWR adalah fungsi kernel normal (gaussian) dan kernel kuadrat ganda (bi-square). Menurut bank dunia salah satu sebab kemiskinan adalah karena kurangnya pendapatan dan aset (*Lact of income and assets*) untuk memenuhi kebutuhan dasar. Kemiskinan suatu wilayah dipengaruhi oleh kemiskinan di wilayah sekitarnya. Tobler (Tobler's first law of geography) dalam Schabenberger dan Gotway (2005) mengatakan *"everything is related to everything else, but near things are more related than distant things"*.

Langkah analisis yang dilakukan yaitu melakukan pengujian dengan metode OLS. Dalam pengujian diperoleh 2 variabel yang signifikan. Selanjutnya melakukan pengujian menggunakan metode *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel gaussian dan bi-square. Membandingkan nilai R^2 dan AIC antara model GWR dengan fungsi pembobot kernel gaussian dan kernel bi-square dengan Program R.

Berdasarkan hasil penelitian yang diperoleh bahwa dengan ANOVA untuk menguji kebaikan GWR secara global, model GWR lebih efektif daripada OLS. Diperoleh model *Geographically Weighted Regression* dengan fungsi pembobot gaussian di Kabupaten Cilacap $\hat{y}_i^* = 0{,}017574 - 0{,}714742X_{i1}^* + 0{,}812049X_{i3}^*$ dan model *Geographically Weighted Regression* dengan fungsi pembobot bi-square di Kabupaten Cilacap $\hat{y}_i^* = -0{,}024805 - 0{,}716867X_{i1}^* + 0{,}832846X_{i3}^*$. Kebaikan model dapat dilihat dari nilai R^2 dan AIC. Nilai R^2 yang diperoleh pada kernel gaussian sebesar 77,47% dan nilai AIC sebesar 53,44198. Sedangkan nilai R^2 yang diperoleh pada kernel bi-square sebesar 76,19% dan nilai AIC sebesar 54,64947. Nilai R^2 terbesar dan nilai AIC terkecil dimiliki oleh model GWR dengan kernel gaussian.

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

DAFTAR SIMBOL

Y	: variabel respon
X	: matriks yang mengandung vektor kolom
β	: vektor parameter regresi
k	: banyaknya variabel prediktor
n	: jumlah pengamatan
ε	: vektor error
(u_i, v_i)	: koordinat titik lokasi pengamatan ke-i (longitude, latitude)
W	: matriks pembobot
d_w	: nilai durbin-watson
d_L	: nilai batas bawah
d_u	: nilai batas atas
CV	: cross validation
d_{ij}	: jarak euclidean
b	: bandwidth
I	: matriks identitas
S_z	: matriks topi untuk model OLS atau GWR
tr	: trace matrik (penjumlahan elemen diagonal suatu matriks)
SE	: standart error

α : taraf signifikansi

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Model regresi disebut juga model regresi global karena pada regresi OLS diasumsikan bahwa nilai estimasi parameter bernilai sama untuk semua lokasi di dalam wilayah penelitian (Permai, 2016). Namun, pada kenyataannya terkadang kondisi data pada lokasi yang satu tidak sama dengan kondisi yang lain. Kondisi yang dipengaruhi oleh aspek spasial yaitu parameter regresi bervariasi secara parsial atau disebut juga nonstasioneritas spasial pada parameter regresi, sehingga jika regresi OLS diterapkan maka asumsi kehomogenan ragam error sulit untuk dipenuhi akibatnya kesimpulan yang didapat dari hasil pengujian untuk model regresi maupun variabel independen yang ada dalam model tidak tepat (LeSage, 2011).

Regresi spasial merupakan pengembangan dari metode linear klasik. Pengembangan ini karena adanya pengaruh tempat atau spasial pada data yang dianalisis (Wang, 2016). Apabila terdapat data dengan efek spasial maka analisis yang digunakan adalah analisis regresi spasial. Karena apabila menggunakan regresi linear sederhana maupun regresi berganda maka hasil kurang akurat dan kesimpulan kurang tepat karena asumsi error saling bebas tidak terpenuhi. Tobler (Tobler's first law of geography) dalam Schabenberger dan Gotway (2005) mengatakan *"everything is related to everything else, but near things are more* *related than distant things"*. Segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang dekat lebih mempunyai pengaruh daripada sesuatu yang jauh. (Day, 2013) menerangkan mengenai pentingnya peranan posisi lokasi yaitu pengetahuan mengenai lokasi dari suatu aktifitas memungkinkan hubungannya dengan aktifitas lain atau elemen lain dalam daerah yang sama atau lokasi yang berdekatan. Fenomena – fenomena yang termasuk data spasial diantaranya adalah penyebaran penyakit, penentuan harga jual rumah, pertanian, kedokteran, kriminalitas, kemiskinan, anak tidak bersekolah, dan lain – lain (Yasir, 2016).

Parameter model diperkirakan untuk setiap lokasi. Hal ini menunjukkan bahwa setiap lokasi memiliki berbeda parameter, maka setiap lokasi memiliki model regresi yang berbeda. Data spasial memuat suatu informasi lokasi yang memiliki ketergantungan antar lokasi. Data spasial mengindikasikan adanya ketergantungan antara pengukuran data dengan lokasi pengamatan atau sering disebut dengan efek spasial (Mullen et al, 2015). Efek spasial merupakan fenomena dimana pengamatan yang dilakukan di suatu lokasi memiliki ketergantungan cukup kuat dengan pengamatan yang dilakukan di lokasi yang berdekatan. Efek spasial dibedakan menjadi 2 yaitu dependensi spasial dan keragaman spasial. Penyelesaian jika terjadi dependensi spasial dengan pendekatan area sedangkan keragaman spasial yang dapat diselesaikan menggunakan pendekatan titik. Adanya efek spasial menyebabkan estimasi tidak tepat karena melanggar asumsi identik dan independen (Septiana dan Wulandari, 2012).

Berdasarkan tipe data, pemodelan spasial dapat dibedakan menjadi pemodelan dengan pendekatan titik dan area (Salamah, 2012). Jenis pendekatan titik yaitu *Geographically Weighted Regression (GWR)*, *Geographically Weighted Poisson Regression (GWPR)*, *Geographically Weighted Logistic Regression (GWLR)*, *Space-Time Autoregressive (STAR)* dan *Generalized Space Time Autoregressive (GSTAR)*. Menurut LeSage (2011), jenis pendekatan area diantaranya *Mixed Regressive-Autoregressive* atau *Spasial Autoregressive Model (SAR), Spasial Error Models (SEM), Spasial Durbin Model (SDM), Conditional Autoregressive Models (CAR), Spasial Autoregressive Moving Average (SARMA),* dan panel data.

Salah satu metode statistika yang dapat digunakan menganalisis faktor risiko secara spasial dengan pendekatan titik adalah model spasial *Geographically Weighted Regression* (GWR). Metode ini memperluas kerangka model regresi global menjadi model regresi lokal yang memungkinkan estimasi parameter secara lokal (Brunton, 2016). Setiap parameter regresi diestimasi di setiap titik lokasi geografis sehingga hubungan antara variabel respon (Y) dan variabel penjelas (X) bervariasi (tidak sama) di sepanjang lokasi. Penggunaan data spasial dan data temporal pada pemodelan berbasis GWR secara simultan dapat menghasilkan model yang lebih informatif dibandingkan dengan hanya menggunakan cross-sectional data.

Pemilihan matriks pembobot adalah salah satu langkah utama dalam GWR karena akan sangat mempengaruhi model GWR yang dihasilkan (Lin & Wen, 2011). Hal yang paling penting dalam model GWR adalah pembobotan karena beratnya merupakan nilai untuk setiap lokasi. Lokasi yang dekat memiliki pengaruh kuat dalam estimasi dari lokasi jauh (Permai, 2016). Metode yang dapat digunakan dalam menentukan berat model GWR menggunakan fungsi kernel, termasuk fungsi Gaussian jarak, fungsi eksponensial, fungsi Bi-Square dan fungsi kernel Tricube. Ada beberapa cara dalam menentukan unsur – unsur matriks pembobot W_i dalam GWR diantaranya pembobot yang mengadopsi fungsi sebaran kernel. Fungsi kepekatan kernel seringkali digunakan dalam pemulusan data dengan memberikan pembobotan sesuai lebar jendela (bandwidth) optimal yang nilainya tergantung pada kondisi data. Fungsi kernel yang digunakan dalam matriks pembobot GWR pada penelitian ini adalah bentuk kernel normal (Gaussian) dan fungsi kernel kuadrat ganda (bi-square) yang menggunakan jarak antara lokasi dalam fungsinya (Sinaga, 2013). Kedua fungsi ini dipilih karena keduanya menggunakan unsur jarak antar lokasi pengamatan yang nilainya bersifat kontinu, sehingga diharapkan hasil analisis akan lebih baik. Model GWR tidak dapat digunakan untuk menduga parameter selain parameter di lokasi pengamatan.

Indonesia merupakan salah satu negara berkembang. Sejak bangsa indonesia dinyatakan merdeka, sudah menjadi cita – cita bangsa untuk mensejahterakan seluruh rakyatnya karena pada kenyataannya yang dihadapi adalah kemiskinan. Menurut bank dunia salah satu sebab kemiskinan adalah karena kurangnya pendapatan dan aset (*Lact of income and assets*) untuk memenuhi kebutuhan dasar seperti makanan, pakaian, perumahan. Di samping itu kemiskinan juga berkaitan dengan keterbatasan lapangan pekerjaan (pengangguran) serta tingkat pendidikan dan kesehatan yang tidak memadai. Setiap tahunnya Badan Pusat Statistik (BPS) melakukan penghitungan kemiskinan dan analisis deskriptif untuk kemiskinan pedesaan maupun perkotaan.

Permasalahan kemiskinan merupakan salah satu persoalan mendasar yang terus di hadapi di sejumlah daerah di Indonesia, tidak terkecuali Provinsi Jawa Tengah. Tingkat kemiskinan di Provinsi Jawa Tengah masih relatif tinggi, dari tahun 2013 sebesar 4834,95 ribu jiwa, pada tahun 2014 sebesar 4836,45 ribu jiwa dan di tahun 2015 sebesar 4577,04 ribu jiwa. Perhitungan dilihat dari hasil Survei Sosial Ekonomi Nasional (Susenas). Langkah untuk mengatasi kemiskinan lebih efektif dengan pendekatan geografis yang berhubungan dengan sumber daya alam dan manusia. Kemiskinan suatu wilayah dipengaruhi oleh kemiskinan di wilayah sekitarnya. Data kemiskinan yang baik dapat digunakan untuk mengevaluasi kebijakan pemerintah terhadap kemiskinan, membandingkan kemiskinan antar waktu dan daerah, serta menentukan target penduduk miskin dengan tujuan untuk memperbaiki kondisi mereka.

Ada beberapa penelitian yang membahas tentang GWR antara lain yang dilakukan oleh Ayunin dan Sutikno (2012) menganalisis pemodelan balita gizi buruk di Kabupaten Ngawi dengan GWR. Hasil penelitian tersebut menunjukkan bahwa faktor geografis mempengaruhi kejadian gizi buruk di Kabupaten Ngawi, sehingga tiap kecamatan berbeda-beda faktor yang mempengaruhinya. Penelitian juga dilakukan oleh Rahmawati dkk (2011) menganalisis penggunaan GWR dengan pembobot gauss kernel untuk klasifikasi desa miskin di Kabupaten Jember. Hasil penelitian tersebut menunjukkan bahwa faktor spasial sangat mempengaruhi hasil pengukuran data dengan menggunakan faktor geografis. Selanjutnya penelitian yang dilakukan oleh Ardhanacitri dan Ratnasari (2013) yang menganalisis pemodelan dan pemetaan pendidikan di Provinsi jawa timur dengan GWR. Hasil penelitian tersebut menunjukkan bahwa berdasarkan GWR didapatkan empat kelompok pemetaan pendidikan yang berpengaruh atau sejalan linear dengan variabel respon.

Berdasarkan uraian di atas, hal yang menjadi fokus dalam penelitian ini adalah **"Pemodelan *Geographically Weighted Regression* (GWR) dengan Fungsi Pembobot Kernel Gaussian dan Bi-square".**

1.2 Rumusan Masalah

Berdasarkan uraian di atas, maka permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut :

1. Bagaimana model *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel gaussian?
2. Bagaimana model *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel bi-square?
3. Model manakah diantara model *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel gaussian dan *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel bi-square yang terbaik?

1.3 Batasan Masalah

Batasan masalah yang dilakukan pada penelitian ini sebagai berikut :

1. Pemodelan spasial yang digunakan adalah pendekatan titik
2. Pendekatan titik yang digunakan adalah *Geographically Weigted Regression* (GWR)
3. Fungsi pembobot model GWR dengan menggunakan kernel gaussian dan kernel bi-square.
4. Data yang digunakan adalah data jumlah penduduk miskin di Provinsi Jawa Tengah dan faktor yang mempengaruhinya pada tahun 2014

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Membentuk model *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel gaussian
2. Membentuk model *Geographically Weighted Regression* (GWR) dengan fungsi pembobot kernel bi-square
3. Menentukan model terbaik antara model *Geographically Weighted Regression* dengan fungsi pembobot kernel gaussian dan model *Geographically Weighted Regression* dengan fungsi pembobot kernel bi-square.

1.5 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

1.5.1 Bagi Penulis

a. Menerapkan ilmu yang telah diperoleh dari perkuliahan sehingga dapat menunjang persiapan untuk persaingan di dunia kerja.

b. Menambah dan menerapkan ilmu pengetahuan statistik yang berhubungan dengan pengembangan model egresi (*Geographically Weighted Regression*).

c. Dapat menguji apakah kemampuan pribadi yang diperoleh selama perkuliahan mampu digunakan dalam berhubungan dengan masyarakat di dalam dunia kerja.

1.5.2 Bagi Pembaca

a. Sebagai bahan bacaan yang dapat menambah ilmu pengetahuan.

b. Dapat menambah informasi dan referensi bacaan serta bahan masukan yang bermanfaat untuk melakukan penelitian selanjutnya.

1.6 Sistematika Penulisan Skripsi

Secara garis besar skripsi dibagi menjadi tiga bagian (bab) yaitu bagian awal skripsi, bagian isi skripsi dan bagian akhir skripsi. Berikut ini menjelaskan masing – masing bagian skripsi.

1. Bagian awal skripsi

Bagian awal skripsi meliputi halaman judul, pernyataan keaslian tulisan, pengesahan, motto dan persembahan, kata pengantar, abstrak, daftar isi, daftar gambar, daftar tabel dan daftar lampiran.

2. Bagian isi skripsi

Bagian skripsi secara garis besar terdiri dari lima bab, yaitu :

BAB 1 PENDAHULUAN

Bab ini berisi mengenai latar belakang, rumusan masalah, batasan masalah, manfaat penelitian dan sistematika penulisan skripsi

BAB 2 TINJAUAN PUSTAKA

Bab ini berisi tentang kajian teori yang mendasari dan berhubungan dengan pemecahan masalah. Teori – teori tersebut digunakan untuk memecahkan masalah yang diangkat dalam skripsi ini. Teori yang digunakan adalah Model *Geographically Weighted Regression* dengan fungsi pembobot kernel gaussian dan fungsi pembobot kernel bi-square dan kerangka berfikir.

BAB 3 METODE PENELITIAN

Bab ini mengulas metode yang digunakan dalam penelitian yang berisi langkah – langkah yang dilakukan untuk memecahkan masalah yaitu identifiksi masalah, fokus permasalahan, metode pengumpulan data, analisis data, pemecahan masalah dan kesimpulan.

BAB 4 HASIL DAN PEMBAHASAN

Bab ini berisi mengenai penyelesaian dari permasalahan yang di ungkapkan.

BAB 5 PENUTUP

Bab ini berisi tentang kesimpulan dari pembahasan dan saran yang berkaitan dengan simpulan.

3. Bagian Akhir skripsi

Bagian akhir skripsi ini meliputi daftar pustaka yang memberikan informasi tentang buku sumber literatur yang digunakan dan lampiran – lampiran yang mendukung skripsi.

BAB 2

TINJAUAN PUSTAKA

2.1 Model Regresi

Analisis regresi berkaitan dengan studi mengenai ketergantungan satu variabel, yaitu variabel dependen terhadap satu atau lebih variabel independen. Bentuk model regresi dengan k variabel prediktor dan jumlah pengamatan n adalah sebagai berikut :

(2.1)

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_{ij} + \varepsilon_i \qquad \text{dengan } i = 1,2,\dots,n$$

Atau dalam bentuk matriks :

(2.2)

$$Y = X\beta + \varepsilon$$

Dimana :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Vektor – vektor Y merupakan variabel respon, β adalah vektor parameter regresi, sedangkan X adalah matriks yang mengandung vektor kolom 1 dan vektor kolom nilai – nilai variabel bebas X untuk setiap variabel di dalam model regresi dan ε adalah error yang diasumsikan identik, independen dan berdistribusi normal dengan mean nol dan varians konstan σ^2.

2.1.1 Pengujian Parameter Model

Pengujian parameter dalam model regresi bertujuan untuk mengetahui apakah parameter tersebut telah menunjukkan hubungan yang nyata antara variabel dependen dan variabel independen serta untuk mengetahui kelayakan parameter dalam menerangkan model. Dalam pengujian parameter regresi, terdapat dua uji yang perlu dilakukan yaitu uji individu dan uji serentak.

a. Uji individu

Uji individu merupakan pengujian yang bertujuan untuk mengetahui apakah nilai koefisien regresi mempunyai pengaruh yang signifikan. Hipotesis dari pengujian secara individu sebagai berikut :

$H_0 : \beta_i = 0$ (tidak ada pengaruh yang signifikan antar variabel)

$H_1 : \beta_i \neq 0, i = 1,2, \dots, p$ (ada pengaruh yang signifikan antar variabel)

Statistik uji yang digunakan adalah

(2.3)

$$t_{hitung} = \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)}$$

Dimana $\hat{\beta}_i$ adalah nilai dugaan β_i dan $se(\hat{\beta}_i)$ adalah simpangan baku bagi β_i. Berikutnya nilai t_{hitung} dibandingkan nilai $t_{(\frac{\alpha}{2}, n-p)}$ dengan keputusan sebagai berikut :

1. Apabila nilai $t_{hitung} > t_{(\frac{\alpha}{2}, n-p)}$ maka nilai H_0 ditolak pada tingkat signifikansi α, artinya variabel independen X_i memberikan pengaruh yang signifikan terhadap variabel dependen Y.

2. Apabila nilai $t_{hitung} < t_{(\frac{\alpha}{2}, n-p)}$ maka nilai H_0 diterima pada tingkat signifikansi α, artinya variabel independen X_i tidak memberikan pengaruh yang signifikan terhadap variabel dependen Y.

b. Uji serentak

Analisa pendekatan varians (ANOVA) digunakan untuk menguji signifikansi keseluruhan dari observasi regresi berganda. Hipotesis dari pengujian secara serentak ini adalah sebagai berikut

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

(H_0 : peubah respon bersama-sama tidak berpengaruh dengan peubah penjelas ke-1 s.d ke-p)

$H_1 : Minimal\ terdapat\ satu\ \beta_i \neq 0, i = 1,2, \ldots, p$

(H_1 : peubah respon bersama-sama berpengaruh dengan minimal 1 peubah penjelas ke-1 s.d ke-p)

Pengujian signifikansi model secara serentak pada model regresi dapat menggunakan tabel ANOVA sebagai berikut :

Tabel 2.1 ANOVA

Sumber Variansi	Jumlah kuadrat	Derajat bebas	Rata – rata kuadrat	Rasio
Regresi	JKR	$p-1$	$RKR = \frac{JKR}{P-1}$	$F = \frac{RKR}{RKS}$
Sisa	JKS	$n-p$	$RKS = \frac{JKS}{n-p}$	
Total	JKT	$n-1$		

Dengan

$$JKT = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 = Y'Y - n\bar{Y}^2$$

$$JKR = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 = b'X'Y' - n\bar{Y}^2$$

$$JKS = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = Y'Y - b'X'Y'$$

Statistik uji yang digunakan adalah

(2.4)

$$F_{hitung} = \frac{RKR}{RKS}$$

Apabila $F_{hitung} > F_{\sigma(p-1,n-p),}$ maka H_0 ditolak artinya minimal terdapat satu variabel bebas yang memiliki pengaruh signifikan terhadap variabel dependen. Pengambilan keputusan juga dapat diperoleh dengan melihat nilai p value untuk mendapatkan nilai F. Jika p value $< \alpha$ maka H_0 ditolak.

2.1.2 Pengujian Asumsi Regresi

Menurut Gujarati (2003) asumsi pada model regresi linear berganda adalah sebagai berikut :

1. Model regresinya linear dalam parameter
2. Nilai rata – rata dari error adalah nol
3. Variansi dari error adalah konstant (homokedastik)
4. Tidak terjadi autokorelasi pada error
5. Tidak terjadi multikolinearitas pada variabel prediktor
6. Error berdistribusi normal

Asumsi tersebut kadang – kadang tidak dipenuhi, maka untuk mendeteksi dan mengatasi adanya masalah pelanggaran asumsi dilakukan :

Tabel 2.2 Asumsi Klasik Regresi

No.	Masalah	Deteksi	Solusi
1	Residual tak berdistribusi normal	Normal probability plot Uji kenormalan : kolmogorof smirnov	Transformasi variabel Regresi bootstrap
2	Heteroskedastisitas $var(\varepsilon_i) \neq {\sigma_t}^2$	Plot e dengan $\hat{y}$ Uji Glester, white Uji Golfeld, Quadt	Transformasi variabel Weight Least Squared
3	Multikolinearitas	$r(X_i, X_j)$ tinggi, $VIF > 10$ $\lvert X'X \rvert \approx 0$ R^2 tinggi, tetapi tidak signifikan	Stepwise Principal component regression Ridge regression
4	Autokorelasi $cov(\varepsilon_i, \varepsilon_j) \neq 0$ untuk $i \neq j$	Plot e dengan $\hat{y}$ Uji Durbin Watson ACF plot	Regresi beda, Regresi ratio memasukkan trend Cochrane Orcutt, Hildtrethlu, Durbin, Prais-Winsten

Uji asumsi klasik digunakan untuk memastikan bahwa multikolinearitas dan autokorelasi tidak terdapat dalam model yang digunakan dan data yang dihasilkan terdistribusi normal. Jika keseluruhan syarat tersebut terpenuhi, berarti model analisis telah layak digunakan. Jika model analisis dikatakan telah layak digunakan maka hasil dari regresi yang dilakukan akan dinilai valid. Uji penyimpangan klasik, dapat dijabarkan sebagai berikut :

a. Uji Normalitas

Uji normalitas bertujuan untuk menguji apakah dalam model regresi, variabel dependen dan variabel independen memiliki data yang terdistribusi normal atau tidak. Data yang terdistribusi normal menunjukkan bahwa tidak terdapat nilai ekstrem yang nantinya dapat mengganggu hasil data penelitian. Model regresi yang baik adalah yang memiliki data normal atau mendekati normal. Untuk mendeteksi normalitas data maka dilakukan analisis statistik yang salah satunya dapat dilihat melalui uji Kolmogorov-Smirnov (K-S). Dalam uji KS terdapat nilai asymp. Sig (2 tailed) yang dapat digunakan sebagai alat bantu untuk mengetahui apakah data terdistribusi normal atau tidak.

Hipotesis :

H_0 : data residual berdistribusi normal

H_1 : data residual tidak berdistribusi normal

Pedoman yang digunakan dalam pengambilan keputusan yaitu :

1. Jika nilai $asymp.sig\ (2\ tailed) > 0{,}05$ maka data berdistribusi normal
2. Jika nilai $asymp.sig\ (2\ tailed) > 0{,}05$ maka data tidak berdistribusi normal

b. Uji Multikolinearitas

Uji Multikolinearitas bertujuan menguji adanya korelasi antar variabel bebas pada model regresi. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel bebas. Jika variabel bebas saling berkorelasi, maka variabel – variabel ini tidak ortogonal atau memiliki koefisien korelasi yang tidak sama dengan nol terhadap variabel bebas lainnya. Adanya multikolinearitas menyebabkan suatu model regresi memiliki varian yang besar, sehingga mengakibatkan sulit mendapatkan estimasi yang tepat, interval konsistensi cenderung lebih lebar dan nilai hitung statistik uji t akan lebih kecil.

Pedoman yang digunakan dalam pengambilan keputusan yaitu :

1. Jika nilai $tolerance > 0{,}1$ dan nilai $VIF < 10$ maka data tidak mengandung multikolinearitas
2. Jika nilai $tolerance > 0{,}1$ dan nilai $VIF < 10$ maka data mengandung multikolinearitas.

c. Uji heteroskedastisitas

Uji heteroskedastisitas digunakan untuk melihat apakah terdapat ketidaksamaan varians dari residual satu ke pengamatan-pengamatan lainnya. Untuk mendeteksi ada atau tidaknya gejala heteroskedasrisitaas dalam model penelitian yang dianalisis dapat dilakukan menggunakan uji gletser. Dengan menggunakan uji gletser, gejala heteroskedastisitas dapat diketahui melalui koefisien regresi dari masing-masing variabel independen terhadap nilai absolute residualnya (abresid) sebagai variabel dependen. Model dipastikan tidak mengandung unsur heteroskedastisitas jika nilai probabilitasnya dan nilai $sig >$

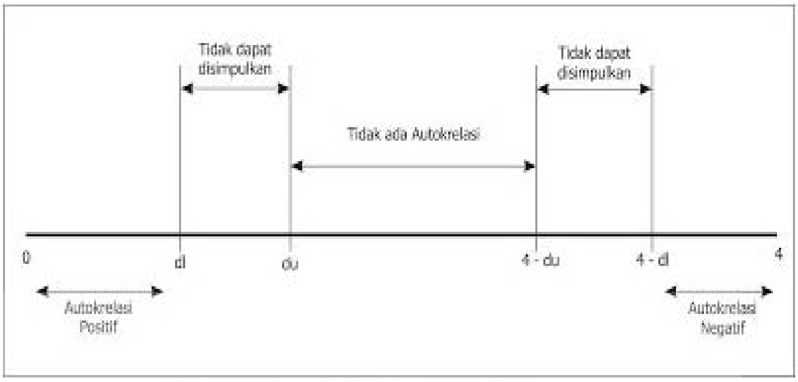
$\alpha(0{,}05)$. Selain itu, medeteksi heteroskedastisitas dapat dilakukan dengan metode scatter plot dengan memplotkan nilai ZPRED (nilai prediksi) dengan SRESID (nilai residualnnya). Model yang baik didapatkan jika tidak terdapat pola tertentu pada grafik, seperti mengumpul ditengah, menyempit, kemudian melebar atau sebaliknya.

Pedoman grafik scatterplot untuk menganalisis apakah terjadi heteroskedastisitas dengan penyebaran titik

1. Jika ada pola tertentu, seperti titik-titik yang ada membentuk pola yang teratur (bergelombang, melebar lalu menyempit) maka mengindikasi terjadi heteroskedastisitas
2. Jika tidak ada pola jelas, seperti titik-titik menyebar diatas dan dibawah angka 0 pada sumbu Y, maka tidak terjadi heteroskedastisitas.

d. Uji Autokorelasi

Autokorelasi dapat diartikan sebagai adanya korelasi antara observasi satu dengan observasi lain yang berlainan waktu. Dalam kaitannya dengan asumsi metode kuadrat terkecil (OLS), autokorelasi merupakan korelasi antara satu residual dengan residual yang lain. Sedangkan satu asumsi penting metode OLS berkaitan dengan residual adalah tidak adanya hubungan antara residual satu dengan residual lain. Cara yang digunakan untuk mendeteksi masalah autokorelasi adalah dengan menggunakan uji Durbin Watson. Menurut Imam Ghozali (2009), uji Durbin Watson merupakan uji paling baik digunakan pada penelitian dengan jumlah data ≤ 100. Gambaran secara garis besar mengenai uji Durbin Watson, sebagai berikut :



Gambar 2.1 Garis besar uji Durbin Watson

Tabel 2.3 Ketentuan Autokorelasi

Kesimpulan	Daerah pengujian
Terjadi autokerelasi positif	$d_w < d_L$
Terjadi autokorelasi negatif	$d_w > (4 - d_L)$
Tidak ada kesimpulan	$d_L \leq d_w \leq d_u$ atau $4 - d_u \leq d_w \leq 4 - d_L$
Tidak terjadi autokorelasi	$d_u < d_w < 4 - d_u$

2.2 Model *Geographically Weighted Regression* (GWR)

Model *Geographically Weighted Regression* (GWR) merupakan pengembangan dari metode regresi. Hanya saja pada model GWR parameter persamaan untuk setiap lokasi pengamatan berbeda dengan lokasi lainnya sehingga banyaknya vektor parameter yang diduga adalah sebanyak lokasi pengamatan yang digunakan dalam data. Dalam analisis GWR, model yang dihasilkan juga tidak dapat digunakan untuk menduga paremeter selain parameter di lokasi pengamatan (Walter et al.2005).

GWR adalah salah satu analisis yang bersifat lokal dan regresi merupakan contoh analisis global. Secara garis besar, perbedaan analisis regresi dan GWR dapat dirumuskan seperti Tabel 2.3. Dalam regresi, nilai parameter diasumsikan sama untuk semua titik lokasi pengamatan, sehingga penduga parameter yang dihasilkan juga bersifat tunggal dan diberlakukan untuk semua lokasi. Sedangkan dalam GWR, nilai parameter tiap lokasi berbeda dengan lokasi lainnya sehingga penduga parameter yang dihasilkan juga banyak sesuai jumlah lokasi pengamatan data yang digunakan (multi-valued statistics). Berbeda dengan regresi yang tidak memperhatikan faktor lokasi (tempat), dalam GWR sangat memperhatikan lokasi (space) sehingga analisis ini seringkali dilanjutkan dengan pemetaan dan dapat didekati dengan Sistem Informasi Geografis atau Geographic Information System (GIS).

Tabel 2.4 Perbedaan regresi biasa dan GWR

	Regresi biasa	GWR
Nilai parameter	Sama untuk semua lokasi, tidak bisa dipetakan	Berbeda untuk setiap lokasi, sehingga bisa dipetakan
Nilai statistik	Tunggal (hanya satu)	Banyak (sebanyak lokasi)
GIS	Tidak ada	Ada
Faktor lokasi	Tidak diperhatikan	Diperhatikan

Pendekatan yang dilakukan dalam GWR adalah pendekatan titik. Setiap nilai parameter ditaksir pada setiap titik lokasi pengamatan, sehingga setiap titik lokasi pengamatan mempunyai nilai parameter yang berbeda – beda.

Model dari *Geographically Weighted Regression* (GWR) dapat ditulis sebagai berikut :

(2.5)

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k\,(u_i, v_i)X_{ik} + \varepsilon_i$$

dimana

Y_i : nilai variabel respon pada titik lokasi pengamatan ke-i

X_{ik} : nilai variabel prediktor ke-k pada titik lokasi pengamatan ke-i

(u_i, v_i) : koordinat titik lokasi pengamatan ke-i (longitude, latitude)

$\beta_0(u_i, v_i)$: koordinat / intercept GWR

$\beta_k(u_i, v_i)$: koefisien regresi ke-k pada titik lokasi pengamatan ke-i

ε_i : error pada titik lokasi ke-i yang diasumsikan dengan rata-rata nol dan varians σ^2

Persamaan dapat ditulis dalam bentuk matriks sebagai berikut :

(2.6)

$$Y = X\beta(u_i, v_i) + \varepsilon$$

dengan $\varepsilon \sim N(0, I\sigma^2)$

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}, \beta(u_i, v_i) = \begin{pmatrix} \beta_0(u_i, v_i) \\ \beta_1(u_i, v_i) \\ \vdots \\ \beta_p(u_i, v_i) \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Jika terdapat n data pengamatan dan p buah variabel prediktor, maka hasil kali antara X dengan β akan menghasilkan matriks berurukan $n * (p + 1)$ dan 1 adalah vektor satuan berukuran $(p + 1) * 1$.

2.3 Estimasi Parameter Model GWR

Pengestimasian parameter $\beta(u_i, v_i)$ pada lokasi ke-i, dapat dilakukan dengan menggunakan metode weighted least squares atau WLS (Brunsdon et al., 1996). Dalam pengestimasian parameter di suatu titik lokasi, metode WLS memberikan pembobot yang tidak sama pada semua amatan. Besarnya pembobot tersebut didasarkan pada jarak antar lokasi amatan. Semakin dekat jarak terhadap amatan yang diestimasi parameternya, semakin besar bobot tersebut dalam estimasi $\beta(u_i, v_i)$.

Tahap awal dari metode WLS adalah membentuk matriks pembobot spasial yang berdimensi mxm untuk lokasi i :

$$W(u_i, v_i) = \begin{pmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{im} \end{pmatrix}$$

Atau dapat ditulis

$$W_i = diag[w_1(u_i, v_i), w_2(u_i, v_i), \dots, w_n(u_i, v_i)]$$

dengan w_{ij} adalah pembobot lokasi ke-j untuk estimasi parameter lokasi ke-i (Fotheringham et al., 2002). Setiap titik lokasi memiliki satu matriks pembobot $W(u_i, v_i)$ (Brunsdon et al., 1996).

Misalkan pembobot untuk setiap lokasi (u_i, v_i) adalah $w_k(u_i, v_i)$ dengan $k = 1,2, \dots, p$ maka parameter pada lokasi pengamatan (u_i, v_i) diestimasi dengan menambahkan *Sum Square Residual (SSR)* dari persamaan (2.5), yaitu :

(2.7)

$$\sum_{j=1}^{n} w_j(u_i, v_i)\varepsilon_j{}^2 = \sum_{j=1}^{n} w_j\,(u_i, v_i)\left[y_j - \beta_0(u_i, v_i) - \sum_{k=1}^{p} \beta_k\,(u_i, v_i)x_{jk}\right]^2$$

Atau dalam bentuk matriks SSR adalah

$$\varepsilon^T W_l \varepsilon = (y - X\beta_l)^T W_l (y - X\beta_l)$$

$$= (y^T - \beta_l^T X^T) W_l (y - X\beta_l)$$

$$= y^T W_l y - W_l y^T X\beta_l - \beta_l^T X^T W_l y + \beta_l^T X^T W_l X\beta_l$$

$$= y^T W_l y - W_l (y^T X\beta_l)^T - \beta_l^T X^T W_l y + \beta_l^T X^T W_l X\beta_l$$

$$= y^T W_l y - \beta_l^T X^T W_l y - \beta_l^T X^T W_l y + \beta_l^T X^T W_l X\beta_l$$

(2.8)

$$= y^T W_l y - 2\beta_l^T X^T W_l y + \beta_l^T X^T W_l X\beta_l$$

dengan

$$\beta_i = \begin{pmatrix} \beta_0(u_i, v_i) \\ \beta_1(u_i, v_i) \\ \vdots \\ \beta_p(u_i, v_i) \end{pmatrix} \text{ dan } W_i = diag[w_1(u_i, v_i), w_2(u_i, v_i), \dots, w_n(u_i, v_i)]$$

Untuk mendapatkan penaksir parameter $\beta(u_i, v_i)$ yang efisien, yaitu dengan menurunkan persamaan (2.8) terhadap $\beta^T(u_i, v_i)$ sebagai berikut :

$$\frac{\partial \varepsilon^T W_l \varepsilon}{\partial \beta^T} = \frac{\partial (y^T W_l y - 2\beta_l^T X^T W_l y + \beta_l^T X^T W_l X\beta_l}{\partial \beta^T}$$

$$= 0 - 2X^T W_l y + X^T W_l X\beta_l + W_l (X^T \beta_l^T X)^T$$

$$= -2X^T W_l y + X^T W_l X\beta_l + X^T W_l X\beta_l$$

$$= -2X^T W_l y + 2X^T W_l X\beta_l$$

$$2X^T W_l y = 2X^T W_l X\beta_l$$

$$X^T W_l y \;= X^T W_l X\beta_l$$

$\beta_l \quad = X^T W_l X^{-1} X^T W_l y$

Sehingga diperoleh estimator parameter model GWR berikut :

(2.9)

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y$$

dengan X adalah matriks rancangan data variabel penjelas X. Matriks rancangan data dapat dinyatakan sebagai $X = \begin{bmatrix} 1 & X_{1;1} & \cdots & X_{p;1} \\ 1 & X_{1;2} & \cdots & X_{p;2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1;m} & \cdots & X_{p;m} \end{bmatrix} = \begin{bmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_m^T \end{bmatrix}$, $W(u_i, v_i)$ adalah matriks pembobot spasial untuk lokasi i dan Y adalah vektor kolom data variabel respon Y (Fortheringham et al., 2002).

2.4 Bandwidth GWR

Dalam fungsi pembobot kernel di atas, terdapat parameter bandwidth yang nilainya tidak diketahui. Sehingga, perlu dilakukan penaksir terhadap parameter bandwidth tersebut. Bandwidth dapat di analogikan sebagai radius (b) suatu lingkaran, sehingga sebuah titik lokasi pengamatan yang berada dalam radius lingkaran masih dianggap berpengaruh dalam membentuk parameter di titik lokasi pengamatan ke-i. Pemilihan bandwidth optimum dalam GWR merupakan hal yang penting karena akan mempengaruhi ketepatan model terhadap data.

Nilai bandwidth yang sangat kecil akan mengakibatkan penaksiran parameter di lokasi pengamatn ke-i semakin bergantung pada titik lokasi pengamatan lain yang memiliki jarak terdekat dengan lokasi pengamatan ke-i, sehingga varians yang dihasilkan akan semakin besar. Sebaliknya, jika nilai bandwidth sangat besar maka akan mengakibatkan bias yang semakin besar, sehingga model yang diperoleh terlalu halus (Dwinata, 2012).

Validasi silang (cross validation) merupakan salah satu cara yang dapat digunakan sebagai kriteria untuk mendapatkan nilai lebar jendela optimum. Lebar jendela optimum yang digunakan adalah yang menghasilkan nilai koefisien validasi silang minimum, dengan rumus koefisiennya adalah :

(2.10)

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

dengan $\hat{y}_{\neq i}(b)$ adalah nilai dugaan y_i (fitting value) dengan pengamatan di lokasi ke- i dihilangkan dari proses prediksi (Fortheringham et al, 2002). Lebar jendela optimum diperoleh dengan proses iterasi hingga didapatkan CV minimum.

2.5 Pembobot Model GWR

Peran pembobot dalam GWR merupakan aspek penting. Pembobot tersebut bergantung pada jarak antar titik lokasi pengamatan. Pembobot berupa matriks diagonal dimana elemen-elemen diagonalnya merupakan sebuah fungsi pembobot dari setiap titik lokasi pengamatan. Fungsi dari matriks pembobot adalah untuk menentukan atau menaksir parameter yang berbeda pada setiap titik lokasi pengamatan.

Matriks pembobot pada GWR merupakan matriks pembobot berbasis pada kedekatan titik lokasi pengamatan ke-i dengan titik lokasi pengamatan lainnya. Pengamatan terdekat ke titik lokasi pengamatan ke-i umumnya diasumsikan memiliki pengaruh paling besar terhadap penaksiran parameter di titik lokasi pengamatan ke-i. Oleh karena itu, matriks pembobot $W(u_i, v_i)$ akan semakin besar seperti jarak yang semakin dekat.

Terdapat beberapa cara yang dapat digunakan untuk menentukan nilai pembobot. Salah satu cara yang paling sederhana adalah dengan memberikan bobot sebesar 1 untuk setiap titik lokasi pengamatan i dan j sebagai berikut :

(2.11)

$$w_{ij} = 1 \text{ , } \forall\ i \text{ dan } j$$

Sehingga model yang dihasilkan apabila menggunakan fungsi pembobot ini adalah model regresi linear klasik.

2.5.1 Fungsi Invers Jarak

Pembobot GWR juga dapat ditentukan dengan menggunakan fungsi invers jarak sebagai berikut :

$$w_{ij} = \begin{cases} 1 & , \ jika\ d_{ij} < b \\ 0 & , \ jika\ d_{ij} \geq b \end{cases}$$

d_{ij} adalah jarak euclidean antara titik lokasi pengamatan ke-i dengan titik lokasi pengamatan ke-j (Fortheringham et al, 2002).

(2.12)

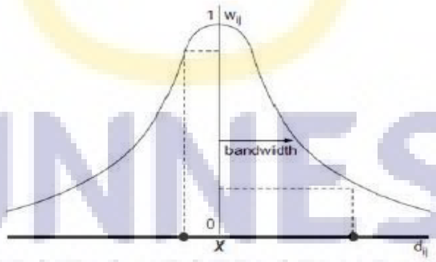
$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$$

dan b merupakan bandwidth atau lebar jendela yang dianalogikan sebagai radius (b) suatu lingkaran, sehingga sebuah titik lokasi pengamatan yang berada dalam radius lingkaran masih dianggap berpengaruh dalam membentuk parameter di titik lokasi pengamatan ke-i.

Satu kelemahan menggunakan pembobot ini adalah tidak bisa digunakan sebagai pembobot untuk dirinya sendiri karena akan menghasilkan nilai yang tak terhingga (unlimited). Fungsi invers jarak akan memberikan bobot = 1, jika titik lokasi ke-j berada di dalam radius b. Sedangkan jika titik lokasi ke-j berada di luar radius b dari titik lokasi ke-i, maka fungsi invers jarak akan memiliki bobot = 0.

2.5.2 Fungsi Kernell

Matriks pembobot $W(u_i, v_i)$ dapat ditentukan menggunakan fungsi kernel. Fungsi kernel memberikan pembobot sesuai bandwidth optimum yang nilainya bergantung pada kondisi data. Fungsi kernel digunakan untuk mengestimasi parameter dalam model GWR jika fungsi jarak adalah fungsi yang kontinu dan monoton turun.



Gambar 2.2 Fungsi kernel spasial

dengan :

X : titik lokasi pengamatan ke-i (regression point)

· : titik lokasi pengamatan lainnya (data point)

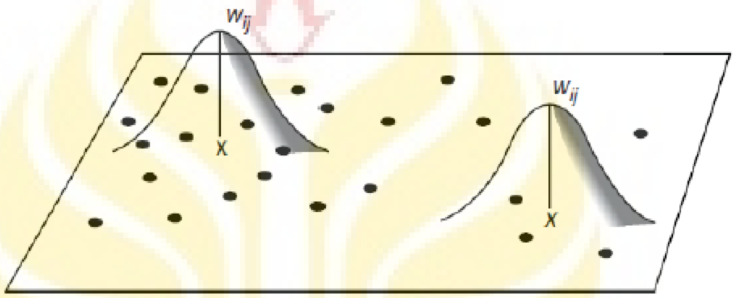
w_{ij} : pembobot dari titik lokasi pengamatan ke-j terhadap titik lokasi pengamatan ke-i

d_{ij} : jarak antara titik lokasi pengamatan ke-i terhadap titik lokasi pengamatan ke-j

Terdapat dua jenis fungsi kernel dalam GWR yaitu fungsi kernel tetap atau fixed kernel dan fungsi kernel adaptif atau adaptive kernel (Wheeler dan Antonio, 2010)

a. Fungsi Kernel Tetap

Fungsi kernel tetap memiliki bandwidth yang sama pada setiap titik lokasi



Gambar 2.3 GWR dengan kernel tetap

dengan :

X : titik lokasi pengamatan ke-i (regression point)

· : titik lokasi pengamatan lainnya (data point)

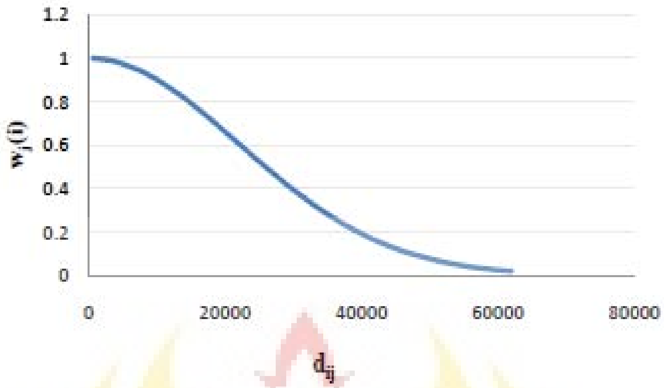
Dua jenis fungsi kernel tetap yang digunakan dalam GWR adalah :

1. Fungsi Kernel Gaussian

(2.13)

$$w_{ij} = e^{\left[-1/2\left(\frac{d_{ij}}{b}\right)^2\right]}$$

Ilustrasi bobot fungsi ini sebagai berikut :



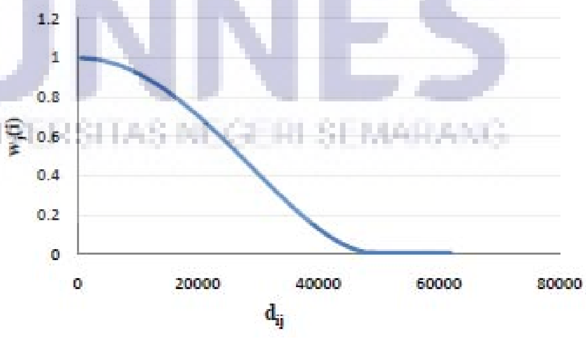
Gambar 2.4 kernel gaussian

2. Fungsi Kernel Bi-square

(2.14)

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b}\right)^2\right]^2 & , jika\ d_{ij} < b \\ 0 & , lainnya \end{cases}$$

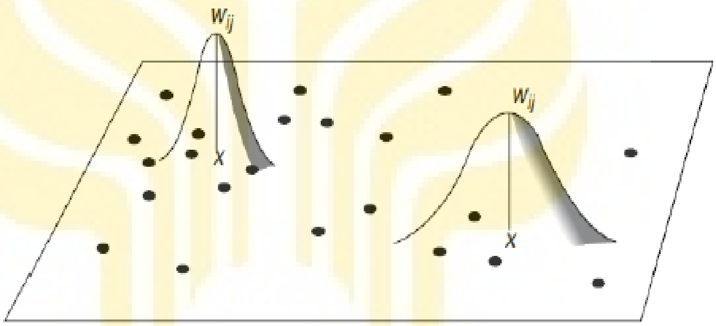
Ilustrasi bobot fungsi ini sebagai berikut :



Gambar 2.5 Kernel Bi-square

b. Fungsi Kernel Adaptif (Adaptive Kernel)

Fungsi kernel adaptif memiliki bandwidth yang berbeda untuk setiap titik lokasi pengamatan. Hal ini disebabkan kemampuan fungsi kernel adaptif yang dapat disesuaikan dengan kondisi titik-titik pengamatan. Bila titik-titik lokasi pengamatan terbesar secara padat disekitar lokasi pengamatan ke-i maka bandwidth yang diperoleh relatif sempit. Sebaliknya jika titik-titik lokasi pengamatan memiliki jarak yang relatif jauh dari titik lokasi pengamatan ke-i maka bandwidth yang diperoleh akan semakin luas (Dwinata, 2012)



Gambar 2.6 GWR dengan kernel adaptif

Dengan :

X : titik lokasi pengamatan ke-i (regression point)

· : titik lokasi pengamatan lainnya (data point)

Dua jenis fungsi kernel adaptif yang digunakan dalam GWR adalah :

1. Fungsi Kernel adaptif Gaussian

(2.15)

$$w_{ij} = e^{\left[-1/2\left(\frac{d_{ij}}{b_{i(q)}}\right)^2\right]}$$

2. Fungsi Kernel adaptif Bi-square

(2.16)

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_{i(q)}}\right)^2\right]^2 & ,jika\ d_{ij} < b \\ 0 & ,lainnya \end{cases}$$

Dengan $b_{i(q)}$ adalah bandwidth adaptif yang menetapkan q sebagai jarak tetangga terdekat dari titik lokasi pengamatan ke-i.

2.6 Uji Hipotesis Model GWR

2.6.1 Pengujian Kesesuaian Model (Goodness of Fit)

Pengujian dilakukan untuk menguji signifikansi dari faktor geografis yang merupakan inti dari model GWR.

Pengujian ini dilakukan dengan hipotesis sebagai berikut :

H_0: $\beta_k(u_i, v_i) = \beta_k, k = 1,2, \dots, p$ (tidak ada pengaruh faktor geografis pada model, untuk setiap $i, i = 1, \dots, n$

H_1: Paling sedikit ada satu $\beta_k(u_i, v_i)$ yang berhubungan dengan lokasi (u_i, v_i) (ada pengaruh faktor geografis pada model).

Penentuan statistik uji untuk uji keberartian model GWR didasarkan pada jumlah kuadrat residual atau residual sum of square yang diperoleh dari model OLS dan GWR.

Misalkan $\hat{Y} = (\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n)'$ adalah vektor dari taksiran Y dan $\hat{\varepsilon} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n)'$ adalah vektor dari nilai residual. Maka diperoleh nilai taksiran untuk variabel respon Y sebagai berikut :

(2.17)

$$\hat{Y} = S_z Y$$

dan nilai taksiran residual dari model OLS atau GWR dinyatakan sebagai berikut :

(2.18)

$$\hat{\varepsilon} = Y - \hat{Y} = Y - S_Z Y = (I - S_Z)Y$$

I merupakan matriks identitas dan S_z merupakan matriks topi dengan z bernilai 0 atau 1 yang masing-masing menunjukkan model OLS atau GWR.

Jumlah kuadrat residual dari kedua model tersebut dapat ditulis sebagai berikut :

(2.19)

$$\hat{\varepsilon}'\hat{\varepsilon} = Y'(I - S_Z)'(I - S_Z)Y$$

Ketika $z = 0$, maka diperoleh S_0 yaitu matriks topi untuk model OLS yang dinyatakan sebagai berikut :

(2.20)

$$S_0 = X(X'X)^{-1}X'$$

Sehingga, jumlah kuadrat residual atau residual sum of square untuk model OLS adalah :

(2.21)

$$JK(S)_{OLS} = Y'(I - S_0)'(I - S_0)Y$$

dan ketika $z = 1$, maka diperoleh S_1 yaitu merupakan matriks topi untuk model GWR yang dinyatakan sebagai berikut :

$$S_1 = \begin{bmatrix} x_1^t[X'W(u_1, v_1)X]^{-1}X'W(u_1, v_1) \\ x_2^{t\prime}[X'W(u_2, v_2)X]^{-1}X'W(u_2, v_2) \\ \vdots \\ x_n^{t\prime}[X'W(u_n, v_n)X]^{-1}X'W(u_n, v_n) \end{bmatrix}$$

Sehingga, jumlah kuadrat residual atau residual sum of square untuk model GWR adalah :

(2.22)

$$JK(S)_{GWR} = Y'(I - S_1)'(I - S_1)Y$$

Selisih antara jumlah kuadrat residual model OLS dan model GWR disebut sebagai GWR improvement dan dinyatakan sebagai berikut :

(2.23)

$$GWR_{IMP} = JK(S)_{OLS} - JK(S)_{GWR}$$

Adapun statistik uji yang digunakan untuk uji keberartian model GWR adalah :

$$F = \frac{JK(S)_{OLS} - JK(S)_{GWR}/df_1}{JK(S)_{GWR}/df_2}$$

dengan n adalah banyak lokasi pengamatan

Kriteria uji yang digunakan yaitu, jika $F \geq F_{\alpha;(dk1,dk2)}$, maka H_0 ditolak. Artinya, ada perbedaan yang signifikan antara model OLS dan model GWR dalam memodelkan data. Nilai $F_{\alpha;(dk1,dk2)}$, diperoleh dari Tabel Distribusi F dengan taraf signifikan α, dk pembilang = $dk_1 = n - p - 1$ dan dk penyebut = $dk_2 = n - 2tr(S_1) + tr(S_1'S_1)$ (Brusdon, Fortheringham & Charlton, 2002, 91-92).

2.6.2 Pengujian Parameter Model

Pengujian parameter model GWR dilakukan dengan menguji parameter secara parsial. Pengujian dilakukan untuk mengatahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Bentuk hipotesisnya adalah sebagai berikut :

$H_0: \beta_k(u_i, v_i) = 0$ (tidak ada pengaruh antar variabel bebas terhadap variabel tak bebas)

$H_1: \beta_k(u_i, v_i) \neq 0; k = 1,2, \dots p$ (minimal terdapat satu variabel bebas yang berpengaruh terhadap variabel tak bebas)

Statistik uji yang digunakan adalah :

(2.24)

$$t_{hit} = \frac{\hat{\beta}_k(u_i, v_i)}{SE[\hat{\beta}_k(u_i, v_i)]}$$

Dimana $SE[\hat{\beta}_k(u_i, v_i)]$ merupakan standar error yang diperoleh dari akar positif varians $\hat{\beta}_k(u_i, v_i)$

Persamaan (estimasi parameter WLS) yaitu penaksir parameter lokal untuk model GWR. Persamaan tersebut dapat ditulis dalam bentuk sederhana sebagai berikut :

(2.25)

$$\hat{\beta}(u_i, v_i) = CY$$

Varians dari penaksir parameter GWR adalah :

(2.26)

$$Var[\hat{\beta}_k(u_i, v_i)] = CC'\sigma^2$$

dengan σ^2 adalah jumlah kuadrat residual normal dari regresi lokal dan di definisikan sebagai berikut :

(2.27)

$$\sigma^2 = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{(n - 2v_1 + v_2)}$$

Dengan

(2.28)

$$v_1 = tr(S_1)$$

dan

(2.29)

$$v_2 = tr(S_1'S_1)$$

Sehingga,

(2.30)

$$SE\left[\hat{\beta}_k(u_i, v_i)\right] = \sqrt{Var\left[\hat{\beta}_k(u_i, v_i)\right]} = \sqrt{CC'\sigma^2}$$

Kriteria pengujian yang digunakan yaitu, jika $|t_{hit}| > t_{1-\frac{\alpha}{2},(n-p-1)}$ maka H_0 ditolak. Artinya $\hat{\beta}_k(u_i, v_i) \neq 0$ atau dengan kata lain koefisien regresi lokal $\hat{\beta}_k(u_i, v_i)$ yang diperoleh untuk model GWR tersebut berarti. Nilai $t_{1-\frac{\alpha}{2},(n-p-1)}$ diperoleh dari Tabel Distribusi t-Student dengan taraf signifikansi $\alpha = 5\%$ dan $dk = (n - p - 1)$.

2.7 Pemilihan Model Terbaik

2.7.1 Koefisien Determinasi (R^2)

Dalam model regresi global (model regresi linear klasik) koefisien determinasi digunakan untuk mengukur proporsi dari variasi dalam data pengamatan yang dapat dijelaskan oleh model. Nilai R^2 yang kecil atau mendekati nol berarti kemampuan variabel bebas dalam menjelaskan variabel tak bebas sangat terbatas, sedangkan nilai R^2 mendekati satu berarti kemampuan dari variabel bebas dalam menjelaskan variabel tak bebas sangat kuat, sehingga mengidentifikasikan bahwa model mampu menjelaskan variabilitas data (Putri, 2013). Sedangkan dalam GWR, koefisien determinasi (R^2) dapat ditentukan dengan menggunakan persamaan sebagai berikut:

(2.31)

$$R_i^2 = \frac{JKT_{GWR} - JKS_{GWR}}{JKT_{GWR}}$$

dengan :

JKT_{GWR} : jumlah kuadrat total model GWR yang dinyatakan sebagai berikut:

(2.32)

$$JKT_{GWR} = \sum_{j}^{n} = 1\, w_{ij}(Y_j - \bar{Y})$$

JKS_{GWR} : jumlah kuadrat residual model GWR yang dinyatakan sebagai berikut :

(2.33)

$$JKS_{GWR} = \sum_{j}^{n} = 1\, w_{ij}(Y_j - \bar{Y}_j)$$

w_{ij} : Pembobot pada titik lokasi pengamatan ke-i dari titik lokasi pengamatan ke-i dengan $i, j = 1,2, \dots, n$ (Maulani, 2013)

2.7.2 Akaike Information Criterion (AIC)

Pemilihan model terbaik merupakan proses evaluasi dari model untuk mengetahui seberapa besar peluang masing – masing model yang terbentuk sudah sesuai dengan data. (Fortheringham, et al, 2002) menuliskan bahwa selain dapat digunakan untuk menentukan bandwidth optimum, AIC juga dapat digunakan dalam pemilihan model untuk menentukan model mana yang terbaik. Penentuan nilai AIC dilakukan dengan perhitungan sebagai berikut :

(2.34)

$$AIC = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n + tr(S)$$

dimana $\hat{\sigma}$ adalah nilai duga standar deviasi residual dan S adalah hat matrix. Model yang didapatkan dari perhitungan dengan nilai AIC terkecil.

2.8 Program R

R adalah suatu kesatuan software yang terintegrasi dengan beberapa fasilitas untuk manipulasi, perhitungan dan penampilan grafik yang handal. R berbasis pada bahasa pemrograman S, yang dikembangkan oleh AT&T Bell Laboratories (sekarang Lucent Technologies) pada akhir tahun '70 an. R merupakan versi gratis dari bahasa S dari software (berbayar) yang sejenis yakni S-PLUS yang banyak digunakan para peneliti dan akademisi dalam melakukan kegiatan ilmiahnya.

Pada awalnya, versi pertama R dibuat oleh Ross Ihaka dan Robert Gentleman dari Universitas Auckland, namun selanjutnya R dikembangkan oleh tim yang disebut tim inti. Tim inti (core team) terdisi dari ahli statistik, ahli komputer dan pemrograman, geografi, ekonomi dari institusi yang berbeda dari seluruh dunia yang mencoba membangun sebuah sistem (software) yang handal namun dengan biaya yang sangat murah. Menurut kutipan dari penghargaan Association for Computing Machinery Software bagi John Chamber 1998, menyatakan bahwa (bahasa pemrograman) S telah merubah orang dalam memanipulasi, visualisasi dan menganalisis data untuk selamanya. R dibuat searah dengan ide yang ada pada bahasa pemrograman S. Banyak projek lainnya yang berkaitan / berbasis / perluasan dari R, seperti geoR, Rattle, R Commander, SciViews R GUI dan lainnya.

2.8.1 Kelebihan progrsm R

R mempunyai karakteristik tersendiri, dimana selalu dimulai dengan prompt “>” pada console-nya. R mempunyai beberapa kelebihan, diantaranya :

1. Efektif dalam pengelolaan data dan fasilitas penyimpanan. Ukuran file yang disimpan jauh lebih kecil dibanding software lainnya.
2. Lengkap dalam operator perhitungan array
3. Lengkap dan terdiri dari koleksi tools statistik yang terintegrasi untuk analisis data, diantaranya mulai statistik deskriptif, fungsi probabilitas, berbagai macam uji statistik hingga time series.
4. Tampilan grafik yang menarik dan fleksibel ataupun costumized
5. Dapat dikembangkan sesuai keperluan dan kebutuhan dan sifatnya yang terbuka, setiap orang dapat menambahkan fitur – fitur tambahan dalam bentuk paket ke dalam software R.

2.8.2 Fungsi Penting dalam R

2.6.1.1 Fungsi Dasar Matematika

Beberapa fungsi dasar telah didefinisikan secara internal di dalam R. Fungsi – fungsi tersebut dapat dilihat pada Tabel 2.4. terhadap matriks atau vektor, operasi tersebut bekerja perunsur. Berikut beberapa contoh penerapan dari fungsi – fungsi tersebut :

Tabel 2.5 Daftar Beberapa Fungsi Matematika Penting Dalam R

No	Nama Fungsi	Notasi Matematika	Fungsi R
1	Harga mutlak	‖	$abs()$
2	Geometri	Sin, cos, tan	$sin(), cos(), tan()$
3	Invers geometri	⟦sin⟧ $\wedge(-1)$, ⟦co	$asin(), acos(), atan()$
4	Hiperbolikus	$Sinh, cosh, tanh$	$sinh(), cosh(), tanh()$
5	Exponensial dan log	exp, ln	$exp(), log()$
6	Logaritma 10	$ln10log$	$log10()$
7	Gamma	$\Gamma()$	$gamma()$

2.8.3 Operasi Vektor dan Matriks

Untuk matriks atau vektor yang berdimensi sama, operasi hitung biasa dapat dilakukan dan itu akan dikerjakan berdasarkan unsur – unsur yang bersesuaian. Khusus untuk operasi vektor dan matriks, R memiliki operasi dasar seperti yang ditunjukkan dalam Tabel 2.6

Tabel 2.6 Daftar Operasi vektor dan Matriks dalam R

No.	Nama Fungsi / Operasi	Notasi Matematika	Fungsi R
1	Pembentukan matriks	X	$matrix\ (data, nbaris, nkolom)$; $seq(awal, akhir, kenaikan)$;
2	Pembentukan barisan		$seq(awal, akhir, length = n)$ $t(x)$
3	Transpose matriks	xT	$det(x)$
4	Determinan matriks	$det(x)$	$diag(data)$
5	Matriks diagonal	□	$diag(matriks)$
6	Diagonal matriks		$x \% * \% y$
7	Perkalian matriks	Xy	$solve(x)$
8	Inverse matriks	$x - 1$	$sum(f(x))$

2.6.1.2 Fungsi Dasar Statistika

Selain fungsi dasar dalam matematika, R juga mempunyai sekumpulan fungsi dasar yang biasa dipergunakan dalam bidang statistika. Variabel dalam fungsi statistika ini berupa vektor data. Fungsi ini dirangkum pada Tabel 2.7

Tabel 2.7 Fungsi dasar statistika

No	Nama Fungsi	Notasi Statistika	Fungsi perintah dalam R
1	Minimum, maksimum	Min, max	$min(\,), max(\,)$
2	Range	$Range$	$range(\,)$
3	Mean, median	$\bar{x}, median$	$mean(\,), median(\,)$
4	Variance	$S2$	$var(\,)$
5	Correlation	ρ_{xy}	$cor(x, y)$
6	Ringkasan data		$summary(\,)$
7	contoh		$sample(\,)$

2.6.1.3 Fungsi Pembangkit Data Peubah Acak

Fungsi ini merupakan fungsi untuk membangkitkan data dari peubah acak dengan berbagai distribusi yang banyak dijumpai, seperti normal, poisson dan gamma dengan jumlah/ukuran sampel n. Pada dasarnya ada empat jenis fungsi terkait dengan distribusi peubah acak yaitu sebagai berikut :

Rdistribusi untuk membangkitkan data acak / random dari suatu distribusi dengan parameter tersebut

Ddistribusi untuk mencari nilai fungsi kepadatan $f(x)$ pada suatu nilai x tertentu.

Pdistribusi untuk mencari luas daerah (nilai peluang) suatu distribusi yang dibatasi oleh nilai x tertentu

Qdistribusi untuk mencari nilai x yang membatasi luas daerah (nilai peluang) tertentu dari suatu distribusi.

Dalam istilah diatas, distribusi merupakan nama-nama distribusi yang tersedia pada R diantaranya beberapa yang penting yang banyak dipakai adalah $norm\ (normal), gamma(gamma), t\ (t), F(F), chisq(\chi^2), pois\ (Poisson)$. Sebagai daftar fungsi-fungsi ini dapat dilihat pada tabel 2.8

Tabel 2.8 Fungsi Pembangkit Data Peubah Acak

No	Nama distribusi	parameter	Perintah dalam R
1	Normal $N(\mu, \sigma^2)$	Mean = μ Varians = σ^2	$rnorm(n, mean, sigma)$; $dnorm(x, mean, sigma)$; $pnorm(x, mean, sigma)$; $qnorm(p, mean, sigma)$
2	Gamma $G(\alpha, \beta)$	$\mu = \alpha/\beta$ $\sigma^2 = \alpha/\beta^2$	$rgamma(n, alpha, beta)$; $dgamma(x, alpha, beta)$; $pgamma(x, alpha, beta)$; $qgamma(p, alpha, beta)$
3	Poisson $(\boldsymbol{\lambda})$	$\mu = \sigma^2 = \boldsymbol{\lambda}$	$rpois(n, lamda)$; $dpois(x, lamda)$; $ppois(x, lamda)$; $qpois(p, lamda)$
4	Binomial (s, π)	$\mu = s\pi$ $\sigma^2 = s\pi(1-\pi)$	$rbinom(n, s, pi)$; $dbinom(x, s, pi)$; $pbinom(x, s, pi)$; $qbinom(p, s, pi)$
5	Chi-kuadrat (χ_v^2)		$rchisq(n, nu)$; $dchisq(x, nu)$; $pchisq(x, nu)$; $qchisq(p, nu)$

2.6.1.4 Fungsi untuk Menangani Grafik

Untuk mengangani grafik, R memiliki beberapa fungsi seperti ditunjukkan pada Tabel 2.9. Dokumentasi yang lebih lengkap dapat diperoleh dengan menggunakan lay out lembaran grafik yang dibagi menjadi matriks sublembaran kecil (a x b).

Tabel 2.9 Fungsi Grafik

No	Tujuan	Perintah R	Keterangan
1	Membuat layout multigrafik (banyak layar)	$par(mfrow = c(b,k))$	b=banyak baris k= banyak kolom
2	Membuat diagram (grafik pencaran= p, dan garis= l)	$plot(x, y, type = 'l/p/b', xlab = " ", ylab = " ", lty = 0, ylim = c(,))$	l=line (grafik garis) p=point (grafik titik) b= keduanya

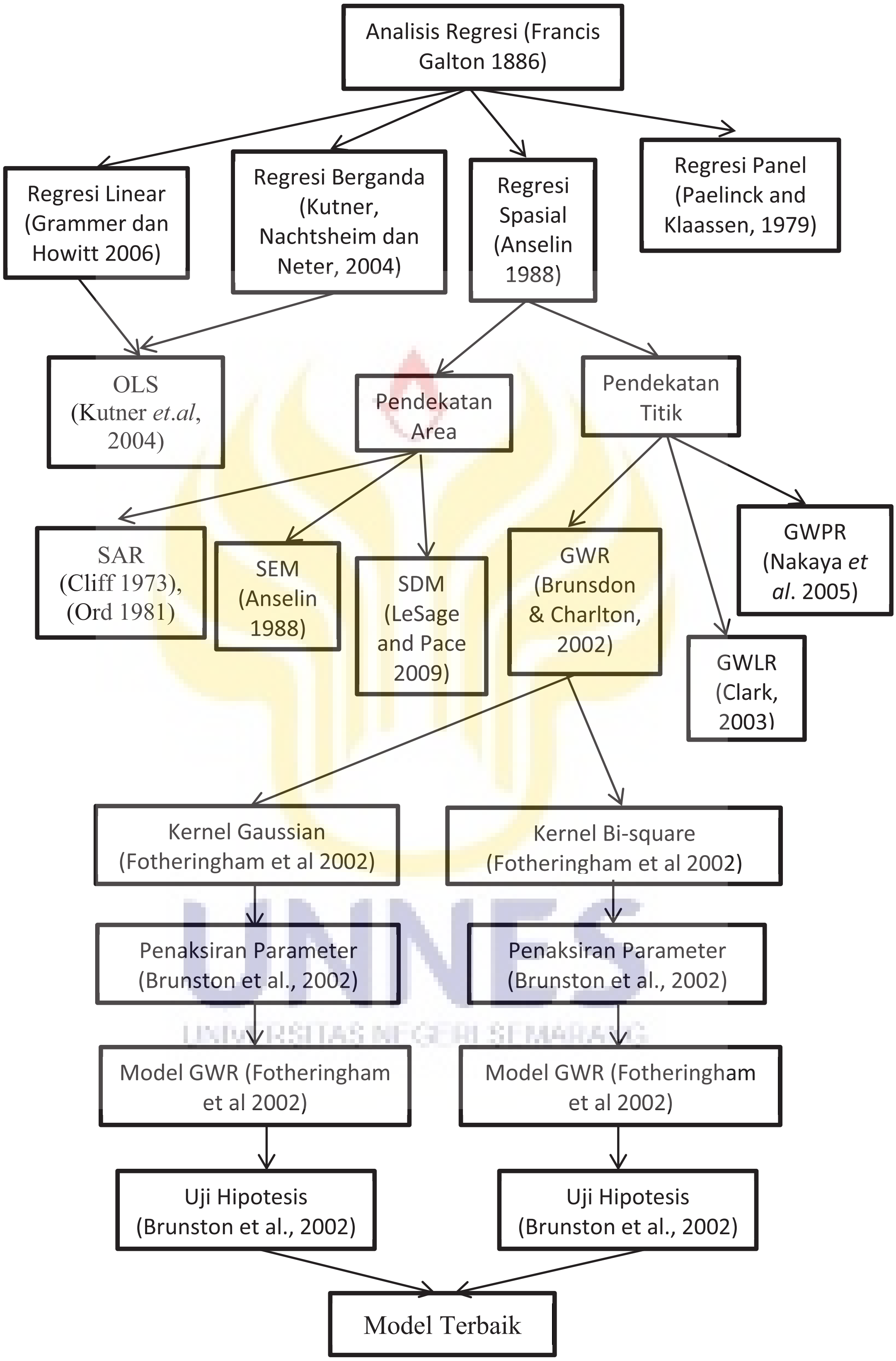
2.9 Kerangka berpikir

Analisis regresi dapat dibedakan menjadi 3 bagian yaitu analisis regresi sederhana (linear dan berganda), analisis regresi spasial, dan analisis data panel. Penelitian dalam skripsi ini hanya akan membahas tentang analisis regresi spasial. Regresi spasial merupakan pengembangan dari metode linear klasik. Pengembangan berdasarkan adanya pengaruh tempat atau spasial pada data yang dianalisis (Anselin, 1988). Data spasial adalah data yang mengacu pada posisi, objek, dan hubungan diantaranya dalam ruang bumi.

Pendekatan keruangan mendasarkan pada perbedaan sifat penting lokasi seperti struktur, pola, dan proses. Struktur keruangan berkaitan dengan elemen pembentuk ruang berupa kenampakan titik, kenampakan garis, dan kenampakan area. Data lokasi yang dibutuhkan dalam analisis keruangan meliputi data titik, garis, dan area. Data spasial dalam penelitian ini lebih difokuskan pada pendekatan area sehingga pemodelan yang termasuk dalam pendekatan titik diantaranya *Geographically Weighted Regression* (GWR), *Geographically Weighted Logistic Regression* (GWLR), *Geographically Weighted Poisson Regression* (GWPR). Dalam metode *Geographically Weighted Regression* digunakan unsur yang sangat penting yaitu matriks pembobot karena nilai pembobot ini mewakili letak data observasi satu dengan lainnya. Pembobot memiliki fungsi untuk memberikan hasil pendugaan parameter yang berbeda pada setiap lokasi.

Fungsi pembobot yang dipilih merupakan fungsi pembobot yang menggunakan unsur jarak yang bersifat kontinu dan diharapkan dapat menghasilkan model dengan tingkat pemulusan yang lebih baik. Fungsi kernel digunakan untuk mengestimasi parameter dalam model GWR jika fungsi jarak adalah fungsi yang kontinu dan monoton menurun. Fungsi pembobot fixed kernel digunakan karena keduanya melibatkan unsur jarak antar lokasi amatan yang nialainya kontinu dalam membangun matriks pembobot , sehingga setiap lokasi akan mendapatkan bobot sesuai dengan jarak lokasi tersebut dengan lokasi amatan. Fungsi pembobot fixed kernel dibagi menjadi dua bagian yaitu Fungsi Pembobot kernel gaussian, dan Fungsi Pembobot kernel bi-square. Untuk pengamatan yang dekat dengan lokasi i maka akan lebih berpengaruh dalam membentuk parameter model lokasi ke i. Karena itu pengamatan – pengamatan yang terletak didalam radius h masih dianggap berpengaruh terhadap model pada lokasi tersebut, sehingga akan diberi bobot yang akan bergantung pada fungsi yang digunakan. Metode pemilihan bandwidth sangat penting digunakan untuk pendugaan fungsi kernel yang tepat. Dengan adanya bandwidth dan fungsi kernel yang ditentukan maka selanjutnya melakukan estimasi parameter (penaksir parameter) terhadap model *Geographically Weighted Regression* (GWR), estimasi dilakukan menggunakan metode Weighted Least Square (WLS). Pembobot pada model GWR memiliki peran penting karena nilai pembobot mewakili letak data observasi satu dengan yang lainnya. Pembobot yang diperoleh dapat dilanjutkan untuk menentukan koefisien regresi GWR dalam tiap lokasi pengamatan yang digunakan untuk menentukan model GWR.

Pengujian hipotesis dalam model GWR terdiri dari 2 macam, yaitu uji keberartian model GWR dan uji signifikansi parameter model GWR. Uji keberartian dilakukan untuk menentukan apakah model GWR baik secara signifikan dalam memodelkan data atau tidak. Sedangkan uji signifikansi parameter dilakukan untuk mengetahui faktor – faktor apasaja yang mempengaruhi jumlah penduduk miskin di setiap kota / Kabupaten di Jawa Tengah.


Analisis Regresi (Francis Galton 1886)
Regresi Linear (Grammer dan Howitt 2006)
Regresi Berganda (Kutner, Nachtsheim dan Neter, 2004)
Regresi Spasial (Anselin 1988)
Regresi Panel (Paelinck and Klaassen, 1979)
OLS (Kutner et.al, 2004)
Pendekatan Area
Pendekatan Titik
SAR (Cliff 1973), (Ord 1981)
SEM (Anselin 1988)
SDM (LeSage and Pace 2009)
GWR (Brunsdon & Charlton, 2002)
GWPR (Nakaya et al. 2005)
GWLR (Clark, 2003)
Kernel Gaussian (Fotheringham et al 2002)
Kernel Bi-square (Fotheringham et al 2002)
Penaksiran Parameter (Brunston et al., 2002)
Penaksiran Parameter (Brunston et al., 2002)
Model GWR (Fotheringham et al 2002)
Model GWR (Fotheringham et al 2002)
Uji Hipotesis (Brunston et al., 2002)
Uji Hipotesis (Brunston et al., 2002)
Model Terbaik

BAB 5

PENUTUP

5.1 Simpulan

Berdasarkan uraian yang telah dijelaskan pada bab sebelumnya, maka dapat disimpulkan sebagai berikut :

1. Model *Geographically Weighted Regression* dengan fungsi pembobot kernel gaussian sebagai contoh di Kabupaten Cilacap adalah sebagai berikut :

$$\hat{y}_i^* = 0{,}017574 - 0{,}714742X_{i1}^* + 0{,}812049X_{i3}^*$$

2. Model *Geographically Weighted Regression* dengan fungsi pembobot kernel bi-square sebagai contoh di Kabupaten Cilacap adalah sebagai berikut :

$$\hat{y}_i^* = -0{,}024805 - 0{,}716867X_{i1}^* + 0{,}832846X_{i3}^*$$

Dengan :

$\hat{y}_i^*$: nilai JPM Kabupaten/Kota yang telah distandarisasi

X_{i1}^* : nilai IPM Kabupaten/Kota yang telah distandarisasi

X_{i3}^* : nilai JFK Kabupaten/Kota yang telah distandarisasi

3. Model terbaik antara model *Geographically Weighted Regression* dengan fungsi pembobot kernel gaussian dan model *Geographically Weighted Regression* dengan pembobot kernel bi-square ditentukan oleh nilai R^2 dan AIC. Nilai R^2 yang diperoleh pada kernel gaussian sebesar 77,47% dan nilai AIC yang diperoleh sebesar 53,44198. Sedangkan nilai R^2 yang diperoleh pada kernel bi-square

Sebesar 78,19% dan nilai AIC yang diperoleh sebesar 54,64947. Nilai R^2 terbesar dan nilai AIC terkecil dimiliki oleh model GWR dengan kernel gaussian. Sehingga model GWR dengan kernel gaussian lebih baik daripada model GWR dengan kernel bi-square untuk pemodelan jumlah penduduk miskin di Provinsi Jawa Tengah di tahun 2014.

5.2 Saran

Berdasarkan hasil penelitian dan pembahasan, saran yang dapat diberikan peneliti adalah sebagai berikut :

1. Faktor-faktor yang berpengaruh signifikan terhadap jumlah penduduk miskin terdiri dari 2 faktor yaitu Indeks Pembangunan Manusia (IPM) dan Fasilitas Kesehatan. Diharapkan faktor tersebut dapat dijadikan pertimbangan bagi Pemerintah Daerah Provinsi Jawa Tengah dalam pengambilan keputusan untuk penanganan kemiskinan agar menjadi lebih efektif dan efisien.
2. Peneliti menggunakan regresi spasial dengan pendekatan titik yaitu *Geographically Weighted Regression* (GWR). Dengan penambahan faktor lain dalam penelitian selanjutnya memungkinkan model regresi spasial dengan pendekatan titik lainnya seperti *Geographically Weighted Poisson Regression* (GWPR) atau *Geographically Weighted Logistic Regression* (GWLR).
3. Fungsi pembobot yang digunakan dalam penelitian ini adalah dengan fungsi kernel gaussian dan fungsi kernel bi-square, penelitian selanjutnya dapat menggunakan fungsi pembobot tricube atau fungsi pembobot adaptive.

DAFTAR PUSTAKA

Anselin, L. 1999. *Spatial Econometrics*. Bruton Center. School of Social Science, University of Texas at Dallas.

Astuti R.D.K, Yasin H, dan Sugito. 2013. *Aplikasi Model Spatial Autoregressive Untuk Pemodelan Angka Partisipasi Murni Jenjang Pendidikan SMA Sederajat di Provinsi Jawa Tengah Tahun 2011*. Prosiding Seminar Nasional Statistika, Universitas Diponegoro

Aulele S.N, Lesnussa Y.A. 2013. *Pendekatan Model Geographically Weighted Regression Untuk Menganalisis Jumlah Penduduk Miskin : Upaya Penurunan Jumlah Penduduk Miskin di Provinsi Maluku*. Prosiding FMIPA Universitas Pattimura.

Brunsdon Ch, Fotheringham S, and Charlton M. 1996. Geographically Weighted Regression : A Method For Exploring Spatial Nonstationarity. *Geogr. Anal*. 28, 281-298.

Cardozo O.D, Carlos J, and Gutierrez J. 2012. *Application of Geographically Weighted Regression to The Direct Forecasting of Transit Ridership at Station-leve*l. Journal of Applied Geography 34, 548-558.

Chen D-R, Truong, and Khoa., 2012. *Using Multilevel Modeling And Geographically Weighted Regression To Identify Spatial Variations In The Relationship Place-Level Disadvantages And Ibesity In Taiwan*. Journal of Applied Geography 32, 737-745.

Cisse B, Ndiaye J.L, and Bathiery O. 2015. *Application of Geographically Weighted Regresson Analysis to Asses Risk Factors for Malaria Hotspots In Keur Soce Health And Demographic Surveillance Site*. Ndiath et al. Malaria Journal 14 : 463

Day J, Lewis B. 2013. *Beyond Univariate Measurement of Spatial Autocorrelation : disaggregated spillover effects for Indonesia*. Faculty of Architecture, The University of Melbourne, vol. 19 no.3 page 169-185

Dwinata, A. 2012. *Model Regresi Logistik Terboboti Geografis (studi kasus : pemodelan kemiskinan di Provinsi Jawa Timur)*. Tesis. Institut Pertanian Bogor.

Fotheringham AS, Brunsdon C, Chalton M . 2002. *Geographically Weighted Regression : The Analysis of Spatially Varying Relationships* Vol.13. John Wiley & Sons, Chichester.

Ghozali, Imam. 2009. *Aplikasi Analisis Multivariate dengan Program SPSS*, Semarang : Badan Penerbit Universitas Diponegoro.

Gollini I, Lu Binbin, and Charlton M. 2015. *Gwmodel : An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models*. Journal of Statistical Software, volume 63, issue 17.

Gurajati, Damodar. 2003. *Ekonometrika Dasar* : Edisi Keenam, Jakarta : Erlangga

Lin C-H, Wen T-H. 2011. *Using Geographically Weighted Regression (GWR) to Explore Spatial Varying Relationships of Immature Mosquitoes and Human Densities with the Incidence of Dengue*. International Journal of Environmental Research and Public Health 8, 2798-2815

Lu Binbin, Charlton M, and Fortheringham S. 2011. *Geographically Weighted Regression Using a Non-Euclidean Distance Metric with a Study on London House Price Data*. Procedia Environment Science 7, 92-97.

Matuszewska K.L, Urbanski J.A. 2014. *Prediction of near-bottom water salinity in the Baltic Sea using Ordinary Least Squares and Geographically Weighted Regression models*. Estuarine, Coastal and Shelf Science, 255-263

Miranti Ita, Djuraidah A, dan Indahwati. 2015. *Modeling of Malaria Prevalence in Indonesia with Geographically Weighted Regression*. Departement of Statistics, Bogor Agricultural University, Vol.9 No.2, pp 109-118.

Paramita, Asharina Dwi. 2014. *Estimasi Model Mixed Geographically Weighted Regression (MGWR) Menggunakan Fungsi Pembobot Fixed Kernel Pada Data Spasial*. Skripsi. FMIPA, Universitas Brawijaya.

Pasculli A, Sarra Annalina, and Miccadei E. 2014. *A Modelling Methodology For The Analysis of Radon Potential Based on Environment Geology and Geographically Weighted Regression*. Environmental Modelling & Software, 54, 165-181.

Permai S.D, Tanty H, dan Rahayu, A. 2016. *Geographically Weighted Regression Analysis for Human Development Index.* Departement of Mathematics and Statistics, School of Computer Science, Bina Nusantara University, Jakarta. Published by the American Institute of Phisics.

Putri A, Salamah M. 2013. *Pemodelan Kasus Balita Gizi Buruk di Kabupaten Bojonegoro dengan Geographically Weighted Regression*. FMIPA, Institut Teknologi Sepuluh Nopember (ITS). Jurnal Sains dan Seni Pomits Vol. 2, No.1

Rahmawati R, Djuraidah A. 2010. *Regresi Terboboti Geografis Dengan Pembobot Kernel Kuadrat Ganda Untuk Data Kemiskinan di Kabupaten Jember*. Forum Statistika dan Komputasi, Vol. 15 No.2 page : 32-37.

Salamah M, Pertiwi L.D, dan Sutikno. 2012. *Spatial Durbin Model untuk Mengidentifikasi Faktor-Faktor yang Mempengaruhi Kematian Ibu di Jawa Timur*. FMIPA, Institut Teknologi Sepuluh Nopember. Jurnal Sains dan Seni Vol. 1, No.1

Septiana L, Wulandari S.P. 2012. *Pemodelan Remaja Putus Sekolah Usia SMA di Propinsi Jawa Timur dengan Menggunakan Metode Regresi Spasial*. [Skripsi]. Surabaya : Program Sarjana Jurusan Statistika.

Schabenberger O, Gotway C.A. 2005. *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC.

Sheng J, Han X, and Zhou H. 2016. *Spatially varying patterns of afforestation / reforestation and socioeconomic factors in China : a geographically weighted regression approach*. Journal of Cleaner Production XXX, pg : 1-10.

Walter J, Carsten R and Jeremy W Lichstein. 2005. *Local and Global Approaches to Spatial Data Analysis in Ecology*. Global Ecology and Biogeography 14, 97-98

Wang, C. 2016. *The Impact of car ownership and public transport usage on cancer screening coverage : Empirical evidence using a spatial analysis in England.* Journal of transport geography. University of London, page 15-22.

Yasir A, Fariqa N, dan Ramadhan F. 2016. *Model Regresi Spasial untuk Analisis Persentase Penduduk Miskin di Provinsi Nanggroe Aceh Darussalam*. Jurnal Statistik Industri dan Komputasi, Jurusan Statistik IST AKPRIND Yogyakarta, vol.1., pp 53-61.

Zhang H, Zhang J, and Lu Shaojing. 2011. *Modeling Hotel Room Price with Geographically Weighted Regression*. International Journal of Hospitality Management 30, 1036-1043.