PEMODELAN *LOCALLY COMPENSATED RIDGE GEOGRAPHICALLY WEIGHTED REGRESSION* (LCR-GWR) DAN *GEOGRAPHICALLY WEIGHTED LASSO* (GWL) DALAM MENGATASI MULTIKOLINIERITAS PADA *GROSS DOMESTIC PRODUCT* (GDP) AMERIKA SERIKAT TAHUN 2022

(Skripsi)

Oleh

NUFUS AULIA

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024

ABSTRACT

MODELING *LOCALLY COMPENSATED RIDGE GEOGRAPHICALLY WEIGHTED REGRESSION* (LCR-GWR) AND *GEOGRAPHICALLY WEIGHTED LASSO* (GWL) IN OVERCOMING MULTICOLLINEARITY IN *GROSS DOMESTIC PRODUCT* (GDP) OF THE UNITED STATES IN 2022

By

NUFUS AULIA

Geographically Weighted Regression (GWR) is a multiple regression to model spatial data to overcome the problem of spatial heterogeneity. Another problem that often arises in spatial data is the problem of multicollinearity, which the GWR method cannot handle. To handle multicollinearity problems, several methods can be used including Ridge regression and LASSO regression, both of which are able to shrink highly correlated and insignificant coefficients to zero. While in dealing with multicollinearity in spatial data, appropriate methods are needed, including the *Locally Compensated Ridge-Geographically Weighted Regression* (LCR-GWR) and *Geographically Weighted Lasso* (GWL) methods. This study aims to see the ability between LCR-GWR and GWL in overcoming multicollinearity. The results obtained are the GWL method is able to provide a better solution in handling multicollinearity and spatial heterogeneity in the United States GDP data in 2022 seen from the lower AIC and RMSE values compared to the LCR-GWR.

Keywords: GWR Regression, Multicollinearity, LCR-GWR Regression, GWL Regression

ABSTRAK

PEMODELAN *LOCALLY COMPENSATED RIDGE GEOGRAPHICALLY WEIGHTED REGRESSION* (LCR-GWR) DAN *GEOGRAPHICALLY WEIGHTED LASSO* (GWL) DALAM MENGATASI MULTIKOLINIERITAS PADA *GROSS DOMESTIC PRODUCT* (GDP) AMERIKA SERIKAT TAHUN 2022

Oleh

NUFUS AULIA

Geographically Weighted Regression (GWR) merupakan regresi berganda untuk memodelkan data spasial untuk mengatasi masalah heterogenitas spasial. Permasalahan lain sering muncul pada data spasial adalah masalah multikolinieritas, dimana metode GWR tidak dapat menanganinya. Untuk menangani masalah multikolinieritas dapat digunakan beberapa metode diantaranya regresi Ridge dan regresi LASSO, kedua metode tersebut mampu menyusutkan koefisien yang memiliki korelasi tinggi dan tidak signifikan menjadi nol. Sedangkan dalam menangani multikolinieritas pada data spasial, diperlukan metode yang sesuai di antaranya adalah metode *Locally Compensated Ridge-Geographically Weighted Regression* (LCR-GWR) dan *Geographically Weighted Lasso* (GWL). Penelitian ini bertujuan untuk melihat kemampuan antara LCR-GWR dan GWL dalam mengatasi multikolinieritas. Hasil yang diperoleh adalah metode GWL mampu memberikan solusi yang lebih baik dalam menangani multikolinieritas dan heterogenitas spasial pada data GDP Amerika Serikat tahun 2022 dilihat dari nilai AIC dan RMSE yang lebih rendah dibandingkan dengan LCR-GWR.

Kata Kunci: Regresi GWR, Multikolinieritas, Regresi LCR-GWR, Regresi GWL.

PEMODELAN *LOCALLY COMPENSATED RIDGE GEOGRAPHICALLY WEIGHTED REGRESSION* (LCR-GWR) DAN *GEOGRAPHICALLY WEIGHTED LASSO* (GWL) DALAM MENGATASI MULTIKOLINIERITAS PADA *GROSS DOMESTIC PRODUCT* (GDP) AMERIKA SERIKAT TAHUN 2022

Oleh

NUFUS AULIA

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024

Judul Skripsi : PEMODELAN *LOCALLY COMPENSATED RIDGE GEOGRAPHICALLY WEIGHTED REGRESSION* (LCR-GWR) DAN *GEOGRAPHICALLY WEIGHTED LASSO* (GWL) DALAM MENGATASI MULTIKOLINIERITAS PADA *GROSS DOMESTIC PRODUCT* (GDP) AMERIKA SERIKAT TAHUN 2022

Nama Mahasiswa : **Nufus Aulia**

Nomor Pokok Mahasiswa : **2017031089**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Prof. Ir. Netti Herawati, M.Sc., Ph.D.	**Dra. Dorrah Azis, M.Si.**
NIP. 196501251990032001	NIP. 196101281988112001

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Prof. Ir. Netti Herawati, M.Sc., Ph.D.**	....
Sekretaris	: **Dra. Dorrah Azis, M.Si.**	....
Penguji Bukan Pembimbing	: **Drs. Nusyirwan, M.Si.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi: **7 Mei 2024**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama Mahasiswa	: **Nufus Aulia**
Nomor Pokok Mahasiswa	: **2017031089**
Jurusan	: **Matematika**
Judul Skripsi	: **PEMODELAN *LOCALLY COMPENSATED RIDGE GEOGRAPHICALLY WEIGHTED REGRESSION* (LCR-GWR) DAN *GEOGRAPHICALLY WEIGHTED LASSO* (GWL) DALAM MENGATASI MULTIKOLINIERITAS PADA *GROSS DOMESTIC PRODUCT* (GDP) AMERIKA SERIKAT TAHUN 2022**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri, bukan hasil orang lain, dan semua hasil yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 7 Mei 2024
Penulis

METERAI TEMPEL
10000
A9674ALX160044338

Nufus Aulia
NPM. 2017031089

RIWAYAT HIDUP

Penulis bernama lengkap Nufus Aulia, anak pertama dari dua bersaudara yang dilahirkan di Banjar Agung pada tanggal 6 Agustus 2001 dari pasangan Bapak Hambali dan Ibu Narti.

Penulis menempuh pendidikan di TK Dharma Wanita Hadimulyo yang diselesaikan pada tahun 2008, kemudian melanjutkan pendidikam di SD Negeri 01 Hadimulyo yang diselesaikan pada tahun 2014. Kemudian menempuh pendidikan di SMP Negeri yang diselesaikan pada tahun 2017, dan kemudian melanjutkan pendidikan di SMA Negeri 2 Way Serdang yang diselesaikan pada tahun 2020.

Pada tahun 2020, penulis diterima sebagai mahasiswa S1 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN). Selama menjadi mahasiswi penulis pernah bergabung menjadi anggota organisasi Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) sebagai anggota Bidang Keilmuan pada periode 2021.

Pada tahun 2023, di bulan Januari-Februari penulis melaksanakan Kerja Praktik (KP) di BPS Kabupaten Mesuji dan di bulan Juni-Agustus penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Bina Karya Buana, Kecamatan Rumbia, Kabupaten Lampung Tengah.

KATA INSPIRASI

"Sesungguhnya bersama kesulitan ada kemudahan."
(Q. S Al-Insyirah: 6)

"Hanya orang-orang yang bersabarlah uang disempurnakan pahalanya tanpa batas."
(Q.S Az-Zumar:10)

"Jangan pikirkan hari esok, karena jika kita memperbaiki hari ini niscaya hari esok akan lebih baik."
(Dr. Aidh Al-Qarni)

PERSEMBAHAN

Dengan mengharap rahmat dan ridho Allah SWT, kupersembahkan karya sederhana ini kepada:

Bapak dan Ibu tercinta

Yang tak pernah lelah menyayangi, merawat, dan mendidikku hingga saat ini dengan sepenuh hati. Yang selalu mendukung semua kegiatanku dan senantiasa berdoa untuk keberhasilanku. Terima kasih karena selalu ada untukku di saat suka maupun duka. Semoga karya ini menjadi langkah awal kesuksesanku agar dapat membuat Bapak dan Ibu bangga.

Kakek dan Nenekku tersayang

Yang senantiasa menyayangiku, memberikan dukungan, dan motivasi kepadaku agar aku lebih bersemangat dalam menyelesaikan pendidkan.

Adikku Tersayang

Yang selalu membantu dan mengerti keadaanku. Yang selalu memberikan semangat dan motivasi kepadaku agar aku bisa menjadi kakak yang baik.

Dosen Pembimbing dan Penguji

Yang telah meluangkan waktu untuk memberikan bimbingan dan ilmu yang bermanfaat.

Almamater Tercinta, Universitas Lampung

SANWACANA

Alhamdulillahirobbil'alaamiin puji dan syukur penulis panjatkan kehadirat Allah subhanahu wata'ala. Tuhan Semesta Alam atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul "**Pemodelan *Locally Compensated Ridge Geographically Weighted Regression* (LCR-GWR) dan *Geographically Weighted Lasso* (GWL) dalam Mengatasi Multikolinieritas pada *Gross Domestic Product* (GDP) Amerika Serikat Tahun 2022**" dengan baik dan tepat pada waktunya.

Skripsi ini dapat penulis susun dengan baik dan lancar tidak lain karena dukungan dan bantuan dari banyak pihak. Oleh karena itu, penulis mengucapkan rasa terima kasih kepada:

1. Ibu Prof. Ir. Netti Herawati, M.Sc., Ph.D. selaku dosen pembimbing I dan dosen pembimbing akademik yang telah bersedia memberikan bimbingan, arahan, dan saran dalam proses menyelesaikan skripsi.
2. Ibu Dra. Dorrah Azis, M.Si. selaku dosen pembimbing II atas kesediaan waktu dalam memberikan bimbingan, arahan, dan saran dalam proses penyusunan skripsi ini.
3. Bapak Drs. Nusyirwan, M.Si. selaku dosen penguji yang telah memberikan kritik dan saran sehingga skripsi ini dapat terselesaikan.
4. Bapak Dr. Aang Nuryaman, S.Si, M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.
5. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

6. Seluruh dosen, staf, dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang telah memberikan ilmu dan bantuan kepada penulis.
7. Ayah, ibu, adik tercinta dan keluarga besar yang selalu mendoakan, memberi dukungan dan motivasi untuk selalu semangat dan fokus dalam mengejar cita-cita.
8. Teman-teman jurusan matematika angkatan 2020 yang memberikan rasa kekeluargaan dan dukungan baik
9. Semua pihak yang tidak bisa penulis sebutkan satu persatu yang telah membantu dalam menyelesaikan laporan ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Kritik dan saran sangat penulis harapkan agar dapat digunakan sebagai bahan perbaikan ke depannya. Semoga skripsi ini dapat memberikan manfaat kepada kita.

Bandar Lampung, 7 Mei 2024
Penulis

Nufus Aulia

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi merupakan suatu teknik untuk menyelidiki dan membentuk model hubungan antar variabel. Aplikasi regresi sangat luas dan banyak terjadi di hampir setiap bidang kehidupan, termasuk teknik, ilmu fisika dan kimia, ilmu ekonomi, manajemen, ilmu hayati dan biologi, dan ilmu sosial. Model regresi linier sederhana adalah model regresi yang melibatkan hubungan antara satu variabel X dengan variabel respon Y (Montgomery, *et al*., 2012).

Menurut Yan & Su (2009), regresi linier berganda bertujuan pada hubungan linier antara satu variabel respon dengan lebih dari satu variabel prediktor. Regresi linier berganda memiliki lebih banyak kemungkinan masalah yang muncul dibandingkan regresi linier sederhana seperti kolinearitas, inflasi varians, grafis diagnostik regresi, deteksi pencilan, dan pengamatan yang mempengaruhinya. Dalam regresi linier berganda diasumsikan bahwa variabel respon merupakan fungsi linier dari parameter model dan di dalam model regresi tersebut terdapat lebih dari satu variabel prediktor (Harlan, 2018).

Metode regresi dapat pula digunakan pada data yang mengandung efek spasial. Data spasial merupakan data yang mempunyai sistem koordinat khusus dan berorientasi geografis sehingga dapat dimungkinkan dalam pembentukan sebuah peta (Yulita, 2016). Analisis spasial dilakukan jika data memiliki karakteristik spasial yaitu data mempunyai sifat error yang saling berkorelasi atau data memiliki heterogenitas spasial (Anselin, 1998). Heterogenitas spasial didefinisikan sebagai suatu kondisi dimana terdapat variasi karakteristik antar satu wilayah dengan wilayah pengamatan lainnya sehingga model tidak dapat memodelkan hubungan antara variabel respon dan variabel prediktor dengan tepat (Fotheringham, *et al*., 2002).

Pada kasus data spasial, ada kalanya bahwa pengamatan di suatu wilayah berpengaruh atau bergantung pada pengamatan di wilayah lain yang berdekatan atau bertetangga, sehingga metode analisis regresi tidak cukup memberikan informasi mengenail lokasi yang diamati. Dalam penanganan data spasial, metode regresi dapat dibedakan berdasarkan pada dua pendekatan, yaitu pendekatan titik dan pendekatan area. Pada pendekatan titik digunakan informasi jarak sebagai sebagai matriks pembobotnya. Sedangkan pada pendekatan area metode yang digunakan berdasarkan pada persinggungan antar lokasi yang bertetanggaan atau berdekatan (Hidayah & Indrasetianingsih, 2019). *Geographically Weighted Regression* (GWR) merupakan salah satu metode yang digunakan untuk menganalisis data spasial berdasarkan pendekatan titik (Lutfiani, *et al*., 2019).

Menurut Klar (2021), *Geographically Weighted Regression* (GWR) adalah metode regresi berganda yang digunakan untuk menganalisis data spasial. Untuk menangani masalah heterogenitas spasial tersebut, model lokal digunakan untuk memodelkan hubungan yang berbeda-beda secara spasial antara variabel respon dengan variabel prediktor untuk setiap lokasi dengan memberikan bobot yang lebih besar pada titik data yang terdekat daripada titik data yang jauh. Sebagian besar studi yang menggunakan pendekatan ini yang dikenal sebagai *Geographically Weighted Regression* (GWR) menyimpulkan bahwa metode GWR mampu menghasilkan hasil prediksi yang lebih akurat daripada metode regresi kuadrat terkecil biasa (Yang, *et al*., 2020).

Masalah yang mungkin muncul ketika variabel prediktor lebih dari satu adalah multikolinieritas. Multikolinearitas dapat didefinisikan sebagai keberadaan hubungan yang linier antara beberapa atau seluruh variabel prediktor dalam suatu model regresi yang menyebabkan hasil pendugaan parameter memiliki ragam yang besar sehigga hasil analisis mungkin tidak valid. Multikolinieritas dapat ditangani dengan metode penyusutan dengan menyusutkan keofisien sampai ke no. Metode yang sering digunakan adalah regresi Ridge dan *Least Absolute Shrinkage and Selection Operator* (LASSO) dengan solusi yang digunakan adalah algoritma *Least Angle Regression* (LARS) yaitu dengan menyusutkan koefisien dugaan parameter yang tidak berpengaruh secara signifikan sampai ke nol (Yulita, 2016).

Metode *Locally Compensated Ridge Geographically Weighted Regression* (LCR-GWR) merupakan pengembangan regresi ridge dengan menggunakan metode WLS pada model GWR dan menggunakan satu koefisien bias ridge untuk wilayah pengamayan tertentu. Setiap wilayah akan menghasilkan koefisien ridge secara lokal dan parameter dibiarkan bervariasi setiap wilayah dengan menyesuaikan efek multikolinieritas antarvariabel prediktor sehingga model diharapkan menjadi lebih akurat (Fadliana, *et al*., 2019). Metode *Geographically Weighted Lasso* (GWL) merupakan metode yang digunakan untuk menangani masalah multikolinieritas lokal dan heterogenitas spasial pada model GWR. Metode GWL merupakan salah satu metode yang dapat digunakan untuk menangani keragaman wilayah yang disebabkan oleh perbedaam letak dan kondisi antar wilayah serta untuk menangani adanya multikolinieritas lokal pada model GWR (Wang & Zuo, 2020). Dengan menggunakan metode ini setiap wilayah akan memiliki model regresi yang berbeda sesuai dengan karakteristik spasialnya masing-masing.

Menurut Sanusi, *et al.* (2021), *Gross Domestic Product* (GDP) merupakan total nilai semua jenis barang dan jasa yang dihasilkan oleh suatu negara dalam suatu periode tertentu (biasanya per tahun), sedangkan *Gross Regional Domestic Product* (GRDP) merupakan GDP di suatu wilayah. Urrutia & Tampis (2017), mendefinisikan GDP sebagai nilai pasar semua barang dan jasa yang diproduksi di dalam batas-batas negara yang menjadi faktor pertumbuhan ekonomi suatu negara. GDP dianggap sebagai salah satu indikator utama dalam menentukan kesehatan ekonomi. GDP paling umum diukur dengan menggunakan metode pengeluaran, dimana GDP dihitung dengan menambahkan pengeluaran untuk barang konsumsi baru, pengeluaran investasi baru, pengeluaran pemerintah, dan nilai ekspor dikurangi dengan nilai impor.

Setiyorini, *et.al*. (2017), menyimpulkan bahwa metode GWL merupakan metode yang lebih baik dibandingkan metode GWR dalam menagani masalah multikolinieritas pada data spasial dalam penelitiannya mengenai kemiskinan di Indonesia. Penelitian mengenai metode LCR-GWR mengenai stunting di Nusa Tenggara Timur menyimpulkan bahwa LCR-GWR mampu mengatasi multikolinieritas (Fadliana, *et. al*., 2019). Penelitian mengenai pemodelan GWL dan LCR-GWR untuk mengatasi multikolinieritas pada GDP Amerika Serikat belum pernah dilakukan sebelumnya. Untuk itu, peneliti tertarik mengangkat penelitian ini dengan judul "Pemodelan *Locally Compensated Ridge Geographically Weighted Regression* (LCR-GWR) dan *Geographically Weighted Lasso* (GWL) dalam Mengatasi Multikolinieritas pada *Gross Domestic Product* (GDP) Amerika Serikat Tahun 2022".

1.2 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengetahui performa LCR-GWR dan GWL dalam mengatasi multikolinieritas pada GDP Amerika Serikat tahun 2022 menurut variasi wilayah dengan pendekatan spasial.
2. Mendapatkan nilai duga koefisien regresi yang diperoleh menggunakan metode LCR-GWR dan GWL.
3. Melihat model terbaik berdasarkan nilai RMSE dan AIC.

1.3 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Sebagai pengetahuan dan wawasan mengenai penerapan metode LCR-GWR dan GWL dalam mengatasi multikolinieritas pada data GDP Amerika Serikat Tahun 2022.
2. Dapat digunakan sebagai referensi bagi penelitian selanjutnya.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi Linier

Analisis regresi merupakan teknik statistika yang untuk menyelidiki dan memodelkan hubungan antar variabel. Biasanya X merupakan variabel independen dan Y disebut variabel dependen. Untuk mengurangi kebingungan akan konsep statistik mengenai independensi, variabel X sering disebut sebagai variabel prediktor dan variabel Y disebut dengan variabel respon. Model regresi linier sederhana adalah model regresi yang melibatkan hubungan antara satu variabel prediktor dengan variabel respon (Montgomery, *et al.,* 2012).

Menurut Yan & Su (2009), analisis regresi adalah metode untuk menemukan hubungan antara satu atau lebih variabel respon (juga disebut variabel dependen, variabel yang dijelaskan, variabel yang diprediksi, atau variabel regresi, biasanya dilambangkan dengan Y) dan prediktor (juga disebut variabel independen, variabel prediktor, variabel penjelas, variabel kontrol, atau regressor, biasanya dilambangkan dengan $X_1, X_2, \ldots, X_k$). Sedangkan regresi linier berganda bertujuan pada hubungan linier antara satu variabel respon dengan lebih dari satu variabel prediktor. Regresi linier berganda memiliki lebih banyak kemungkinan masalah yang muncul dibandingkan regresi linier sederhana seperti kolinearitas, inflasi varians, grafis diagnostik regresi, deteksi pencilan, dan pengamatan yang mempengaruhinya

Menurut Mongomery, *et al*. (2012), model regresi linier berganda merupakan model regresi yang melibatkan lebih dari satu variabel prediktor. Secara umum, variabel respon Y dapat berhubungan dengan k buah variabel prediktor, $X_1, X_2, \dots, X_k$, dan seterusnya dan dituliskan dalam bentuk:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \tag{2.1}$$

dengan:

Y	= Variabel respon
β_0	= Intersep
$\beta_1, \beta_2, \dots, \beta_k$	= Koefisien regresi
$X_1, X_2, \dots, X_k$	= Variabel prediktor
ε	= Galat

Sedangkan model regresi dalam bentuk matriks dituliskan dalam persamaan sebagai berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.2}$$

dengan:

$\boldsymbol{Y}$ = Vektor variabel respon dengan ukuran $n \times 1$

$\boldsymbol{X}$ = Matriks variabel prediktor dengan ukuran $n \times k$

$\boldsymbol{\beta}$ = Vektor parameter dengan ukuran $k \times 1$

$\boldsymbol{\varepsilon}$ = Vektor galat dengan ukuran $n \times 1$

2.2 Multikolinieritas

Multikolinieritas adalah kondisi bahwa adanya hubungan linier dari kolom-kolom matriks **X** yang menyebabkan $|\boldsymbol{X}^T\boldsymbol{X}|$ bernilai nol pada solusi model regresi. Multikolinieritas dapat juga didefinisikan sebagai adanya hubungan linier antara beberapa atau seluruh variabel prediktor yang ada pada model regresi. Hal ini biasanya menyebabkan estimasi koefisien regresi, yang kemudian memiliki varian dan kovarian yang besar. (Draper & Smith, 1998). Menurut Montgomery, *et al*. (2012), multikolinieritas dapat dideteksi dengan cara sebagai berikut:

a. Pemeriksaan korelasi variabel prediktor

Pemeriksaan multikolinieritas dapat dilakukan dengan melihat korelasi antar variabel prediktor. Adanya korelasi antar variabel prediktor yang tinggi menunjukkan adanya masalah multikolinieritas.

b. Menggunakan *Variance Inflation Factor* (VIF)

Untuk menghitung nilai VIF digunakan rumus sebagai berikut:

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

di mana R_j^2 adalah koefisien determinasi berganda yang didapatkan dari regresi X_j terhadap variabel-variabel prediktor lainnya. Jelasnya, jika X_j hampir bergantung secara linier pada beberapa regresi lainnya, maka R_j^2 akan mendekati satu dan nilai VIF_j akan besar. VIF yang lebih besar dari 10 mengimplikasikan masalah serius dengan multikolinieritas. Hal ini merupakan indikasi koefisien regresi yang terkait diestimasi dengan buruk karena adanya multikolinieritas. Sedangkan nilai VIF pada model GWR, dihitung berdasarkan masing-masing variabel prediktor pada setiap lokasi pengamatan dengan menggunakan matriks pembobot spasial. Nilai VIF pada model GWR dihitung dengan rumus sebagai berikut:

$$VIF_k(u_i, v_i) = \frac{1}{1 - R_k^2(u_i, v_i)} \tag{2.4}$$

dengan $R_k^2(u_i, v_i)$ merupakan koefisien determinasi dari regresi X_k terhadap variabel-variabel prediktor lainnya di setiap lokasi pengamatan (Wheeler, 2009).

Menurut Yan & Su (2009), multikolinieritas dapat ditangani dengan metode regresi Ridge atau regresi *Least Absolute Shrinkage and Selection Operator* (LASSO), yaitu dengan menyusutkan dan menyeleksi variabel. LASSO dengan menggunakan algoritma LARS menyusutkan koefisien sampai ke nol.

2.3 Pemusatan dan Penskalaan

Pemusatan dan penskalaan merupakan elemen dalam menstandarisasi variabel. Modifikasi yang sederhana dari standarisasi ini adalah transformasi korelasi. Pemusatan adalah nilai yang memperlihatkan perbedaan antara setiap pengamatan dengan rata-rata dari seluruh pengamatan pada variabel tersebut. Sementara itu, penskalaan mencakup penempatan pengamatan dalam satuan standar deviasi dari pengamatan pada variabel tersebut. (Kutner, *et al*., 2005). Pada konteks ini model regresi linier akan distandardisasi. Prosedur standardisasi variabel respon dan variabel prediktor adalah sebagai berikut:

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_Y} \text{ dengan } S_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}} \tag{2.5}$$

$$X_k^* = \frac{X_{ik} - \bar{X}}{S_{X_k}} \text{ dengan } S_{X_k} = \sqrt{\frac{\sum_{i=1}^{n}(X_{ik} - \bar{X})^2}{n-1}} \tag{2.6}$$

Untuk $i = 1, 2, 3, .., n$ dan $k = 1, 2, 3, .., p$

dengan:

Y_i^* = variabel Y dalam bentuk baku

X_k^* = variabel X dalam bentuk baku

$\bar{Y}$ = rata-rata dari Y

$\bar{X}$ = rata-rata dari X

S_Y = standar deviasi dari Y

S_{X_k} = standar deviasi dari X_k

Model regresi yang telah distandardisasi adalah sebagai berikut:

$$Y_i^* = \beta_1^* X_{1i}^* + \beta_2^* X_{2i}^* + \cdots + \beta_k^* X_{ki}^* + \varepsilon_i^* \tag{2.7}$$

2.4 Data Spasial

Menurut Boot (2001), data spasial terdiri dari pengukuran yang dilakukan di lokasi tertentu dalam ruang geografis. Biasanya, data tersebut disajikan secara visual dalam bentuk peta meskipun dapat juga direpresentasikan dalam bentuk analog lainnya seperti foto udara dan gambar yang diindera dari jarak jauh. Lokasi relatif dari nilai data dalam ruang geografis menentukan pola spasial. Analisis pola spasial melibatkan penggambaran dan penjelasan pola-pola tersebut. Data spasial merupakan data yang mempunyai sistem koordinat khusus dan berorientasi geografis sehingga dapat dimungkinkan dalam pembentukan sebuah peta. Data spasial umumnya dapat dicirikan oleh adanya ketergantungan spasial dan heterogenitas (struktur spasial) yang kemudian disebut sebagai efek spasial. Karakteristik data spasial yang cenderung menyebabkan heterogenitas yang akan berdampak pada dan kesalahan pengukuran (Anselin, 1998).

2.3.1 Dependensi Spasial

Menurut dasar hukum Tobler I (1979), yang mengatakan bahwa segala hal yang saling berkaitan atau berhubungan akan memiliki pengaruh yang lebih besar jika hal tersebut saling berdekatan. Dependensi spasial atau disebut juga dengan ketergantungan spasial merupakan kondisi dimana pengamatan pada suatu lokasi berpengaruh terhadap lokasi pengamatan lain jika letak lokasi tersebut saling berdekatan (Anselin, 1988). Menurut Amin, *et al*. (2023), untuk menguji adanya ketergantungan spasial digunakan Indeks Moran. Hipotesis yang digunakan untuk uji dependensi spasial adalah sebagai berikut:

H_0 : Tidak ada ketergantungan spasial

H_1: Ada ketergantungan spasial

Taraf signifikansi: α = 5% sehingga $Z = 1{,}96$.

Statistik Uji: Moran's I Test

$$Z_{hit} = \frac{I - E(I)}{\sqrt{Var(I)}} \tag{2.8}$$

dengan:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{2.9}$$

$$E(I) = -\frac{1}{(n-1)} \tag{2.10}$$

$$Var(I) = \frac{n^2 S_1 - nS_2 + 3S_0^2}{(n^2 - 1)S_0^2} - \left(E(I)\right)^2 \tag{2.11}$$

dengan:

I = Nilai Moran's I

$E(I)$ = Nilai rata-rata dari I

$Var(I)$ = Nilai varians dari I

S_0 $= \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$

S_1 $= \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(w_{ij} + w_{ji}\right)^2}{2}$

S_2 $= \sum_{i=1}^{n} \sum_{j=1}^{n} \left(w_{ij} + w_{ji}\right)^2$

$\bar{y}$ = Rata-rata variabel y

Pengambilan keputusan dari uji dependensi spasial adalah tolak H_0, jika $Z_{hit} > Z_{tabel}$ atau *p-value* < 0,05. Nilai I berada pada interval -1 dan 1. Data mengandung autokorelasi positif jika $I > E(I)$ (Nadya, 2017).

2.3.2 Heterogenitas Spasial

Heterogenitas spasial terjadi disebabkan oleh keberagaman karakteristik suatu lokasi serta letak geografisnya (Fotheringham, *et al.*, 2002). Heterogenitas spasial dapat diketahui dengan menggunakan Uji Breusch-Pagan dengan rumusan hipotesis sebagai berikut:

H_0: Tidak ada heterogenitas spasial

H_1: Ada heterogenitas spasial

Statistik Uji: Breusch-Pagan test

$$Breusch - Pagan = \left(\frac{1}{2}\right)\boldsymbol{f^T Z}(\boldsymbol{Z^T Z})^{-1}\boldsymbol{Z^T f} \sim \chi_p^2 \tag{2.12}$$

Dengan $\boldsymbol{f} = (f_1, f_2, \dots, f_n)^T$ adalah:

$$\boldsymbol{f_i} = \frac{e_i^2}{\sigma^2} - 1 \tag{2.13}$$

dengan:

e_i^2 = Residu untuk pengamatan ke-i dengan matriks berukuran $(n \times 1)$

$\boldsymbol{f}$ = Vektor dengan ukuran $(n \times 1)$

n = Jumlah lokasi pengamatan

σ^2 = Varians dari e_i^2

$\boldsymbol{Z}$ = Matriks dengan ukuran $n \times (p + 1)$ yang merupakan vektor dari matriks **X**.

p = Banyaknya variabel prediktor

Dengan pengambilan keputusan tolak H_0 jika nilai Breusch-Pagan > $\chi^2_{(p)}$ atau jika diperoleh *p-value* < 0,05 (Anselin, 1998).

2.5 *Geographically Weighted Regression* (GWR)

Menurut Klar (2021), *Geographically Weighted Regression* (GWR) adalah metode regresi berganda yang digunakan untuk menganalisis data spasial. Untuk menangani masalah heterogenitas spasial tersebut, model lokal digunakan untuk memodelkan hubungan yang berbeda-beda secara spasial antara variabel respon dengan variabel prediktor untuk setiap lokasi dengan memberikan bobot yang lebih besar pada titik data yang terdekat daripada titik data yang jauh. Sebagian besar studi yang menggunakan pendekatan ini yang dikenal sebagai *Geographically Weighted Regression* (GWR) menyimpulkan bahwa metode GWR mampu menghasilkan hasil prediksi yang lebih akurat daripada metode regresi kuadrat terkecil biasa (Yang, *et al*., 2020).

2.5.1 Pemodelan GWR

Menurut Fotheringham, *et al*. (2002), GWR memungkinkan hubungan lokal diukur dan dipetakan. GWR dapat dapat dianggap sebagai 'mikroskop spasial' karena merupakan peralihan dari pendekatan model regresi global ke pendekatan model regresi lokal untuk analisis data spasial. Hal ini kemudian menghasilkan pembentukan model regresi yang bervariasi di setiap lokasi pengamatannya. Model GWR ditulis dalam bentuk sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \tag{2.14}$$

dengan:

y_i = Nilai variabel respon pada lokasi pengamatan ke- i

x_{ik} = Nilai variabel prediktor k pada lokasi pengamatan ke- i

$\beta_0(u_i, v_i)$ = Intersep pada lokasi pengamatan ke- i

$\beta_k(u_i, v_i)$ = Nilai koefisien regresi lokal ke- k di lokasi ke- i

ε_{ik} = Galat pengamatan ke- i

(u_i, v_i) = Titik koordinat dari lokasi pengamatan ke- i

i = $1, 2, 3, \dots, n$

k = $1, 2, 3, \dots, p$

2.5.2 Pendugaan Parameter Model GWR

Pada model regresi linier parameter diduga menggunakan metode *Ordinary Least Square* (OLS). Sedangkan pada model GWR, di mana parameter dalam persamaan (2.11) diestimasi dengan prosedur *Weighted Least Square* (WLS), membuat sistem pembobotan bergantung pada lokasi dalam ruang geografis. Oleh karena itu, memungkinkan parameter lokal dan bukan global untuk diestimasi. Hasil keluaran khas dari model GWR adalah sekumpulan parameter yang dapat dipetakan dalam ruang geografis untuk mewakili nonstasioneritas atau penyimpangan parameter (Leung, *et al*., 2000). Pada model GWR suatu pengamatan diberi bobot berdasarkan kedekatannya dengan lokasi pengamatan i sehingga pembobotan suatu pengamatan tidak lagi konstan melainkan bervariasi dengan i. Data dari lokasi pengamatan yang dekat dengan i diberikan bobot lebih besar daripada data dari lokasi pengamatan yang lebih jauh. Diberikan bahwa pembobot untuk setiap lokasi pengamatan ke-i adalah w_{ij} dimana $j = 1, 2, 3, \dots, n$, kemudian parameter untuk lokasi pengamatan ke-i diestimasi dengan meminimumkan jumlah kuadrat galat sebagai berikut:

$$\sum_{j=1}^{n} w_j\,(u_i, v_i\,)\varepsilon_j^2 = \sum_{j=1}^{n} w_j\,(u_i, v_i\,)\left[y_j - \beta_0(u_i, v_i) - \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{jk}\right]^2 \tag{2.15}$$

Sehingga estimasi parameter menggunakan metode WLS diperoleh dari persamaan dalam bentuk matriks berikut:

$$\boldsymbol{\varepsilon}^T \boldsymbol{W}(u_i, v_i)\boldsymbol{\varepsilon} = [\boldsymbol{y} - \boldsymbol{X\beta}(u_i, v_i)]^T \boldsymbol{W}(u_i, v_i)[\boldsymbol{y} - \boldsymbol{X\beta}(u_i, v_i)] \tag{2.16}$$

Misalkan $\boldsymbol{W}(u_i, v_i) = \boldsymbol{W}$ maka:

$$\begin{aligned}\boldsymbol{\varepsilon}^T \boldsymbol{W\varepsilon} &= \boldsymbol{y}^T\boldsymbol{Wy} - \boldsymbol{y}^T\boldsymbol{WX\beta}(u_i, v_i) - \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{Wy} \\ &\quad + \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{WX\beta}(u_i, v_i) \\ &= \boldsymbol{y}^T\boldsymbol{Wy} - 2(\boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{Wy}) + \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{WX\beta}(u_i, v_i)\end{aligned} \tag{2.17}$$

Karena $\boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{Wy} = \boldsymbol{y}^T\boldsymbol{WX\beta}(u_i, v_i)$ maka persamaan (2.17) menjadi:

$$\begin{aligned}\boldsymbol{\varepsilon}^T \boldsymbol{W\varepsilon} &= \boldsymbol{y}^T\boldsymbol{Wy} - \boldsymbol{y}^T\boldsymbol{WX\beta}(u_i, v_i) - 2\boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{Wy} \\ &\quad + \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{X}^T\boldsymbol{WX\beta}(u_i, v_i)\end{aligned}$$

Kemudian didifrensialkan terhadap $\boldsymbol{\beta}^T(u_i, v_i)$ dan disamakan nilainya dengan nol maka diperoleh solusi sebagai berikut:

$$\begin{aligned}\frac{\partial(\boldsymbol{\varepsilon}^T\boldsymbol{W\varepsilon})}{\partial\boldsymbol{\beta}^T(u_i,v_i)} = -2\boldsymbol{X}^T\boldsymbol{Wy} + 2\boldsymbol{X}^T\boldsymbol{WX\beta}(u_i, v_i) &= 0 \\ 2\boldsymbol{X}^T\boldsymbol{WX\beta}(u_i, v_i) &= 2\boldsymbol{X}^T\boldsymbol{Wy} \\ \boldsymbol{X}^T\boldsymbol{WX\beta}(u_i, v_i) &= \boldsymbol{X}^T\boldsymbol{Wy} \\ \boldsymbol{\beta}(u_i, v_i) &= (\boldsymbol{X}^T\boldsymbol{WX})^{-1}\boldsymbol{X}^T\boldsymbol{Wy}\end{aligned}$$

Sehingga estimasi parameter model GWR pada persamaan (2.14) untuk setiap lokasi pengamatan adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = (\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} \tag{2.18}$$

dengan:

$\boldsymbol{y}$ = Variabel respon $n \times 1$

$\boldsymbol{X}$ = Matriks variabel prediktor

$\widehat{\boldsymbol{\beta}}(u_i, v_i)$ = Koefisien regresi lokal di lokasi ke- i

$\boldsymbol{W}(u_i, v_i)$ = Diagonal matriks pembobot $n \times n$ yang dihitung setiap lokasi ke- i

Matriks $\boldsymbol{W}(u_i, v_i)$ dituliskan ke dalam bentuk sebagai berikut:

$$\begin{bmatrix} w_1(u_i, v_i) & \cdots & 0 & 0 \\ 0 & w_2(u_i, v_i) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n(u_i, v_i) \end{bmatrix}$$

$\boldsymbol{W}(u_i, v_i)$ adalah matriks pembobot berdasarkan kedekatan titik regresi i dengan titik-titik data di sekitar i. Dalam membentuk matriks pembobot dibutuhkan suatu fungsi pembobot yang dipengaruhi oleh ukuran ketetanggaan yang disebut *bandwidth*. Selanjutnya, fungsi pembobot disesuaikan dengan kedekatan titik lokasi ke- i (Yulita, 2016). Menurut Fadillah (2015), *bandwidth* adalah ukuran radius lingkaran di suatu titik untuk menentukan titik mana saja berpengaruh untuk membentuk parameter model di lokasi pengamatan i. Fungsi pembobot dihitung berdasarkan suatu fungsi kernel yang mengakibatkan lokasi pengamatan yang memiliki jarak yang dekat dengan titik lokasi ke- i memiliki bobot yang lebih besar. Fungsi kernel yang digunakan pada penelitian ini adalah fungsi exponential. Fungsi ini memiliki nilai *bandwidth* yang sama pada setiap pengamatan dan dirumuskan sebagai berikut:

$$w_j(u_i, v_i) = exp\left[\frac{-d_{ij}}{h}\right] \tag{2.19}$$

$$i = 1, 2, 3, \dots, n;\ j = 1, 2, 3, \dots, m$$

Dengan d_{ij} adalah jarak antara titik lokasi pengamatan ke- i dan lokasi pengamatan ke- j yang dihitung dengan jarak euclidean berdasarkan koordinat dari data spasial untuk menghasilkan bobot antar pengamatan. Jarak Euclidean pada model GWR dituliskan dalam bentuk sebagai berikut:

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{2.20}$$

Sedangkan h merupakan *bandwidth* optimum yang diperoleh melalui metode *Cross Validation* (CV). CV adalah proses berulang yang bertujuan untuk menemukan *bandwidth* kernel yang meminimalkan kesalahan prediksi dari semua variabel hasil yang diamati (Klar, 2021). CV didefinisikan sebagai berikut:

$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(h)]^2 \tag{2.21}$$

Dimana $\hat{y}_{\neq i}(h)$ merupakan nilai duga untuk nilai pengamtan y_i dengan menghilangkan pengamatan dari titik lokasi ke- i dalam proses prediksi dan *bandwidth* optimum (h). *Bandwidth* yang meminimalkan nilai CV adalah *bandwidth* yang paling sesuai untuk memaksimalkan kekuatan prediksi model.

2.5.3 Pengujian Hipotesis Model GWR

Dalam pengujian hipotesis model GWR, terdapat 2 (dua) macam pengujian, yaitu uji untuk melihat kesesuaian model GWR dan uji untuk melihat signifikansi parameter model GWR.

2.5.3.1 Uji Kesesuaian Model GWR

Uji kesesuaian model GWR berfungsi untuk menjelaskan apakah model GWR dapat mendeskripsikan data lebih baik dibandingkan model regresi OLS (Leung, *et. al.,* 2000). Uji kesesuaian model GWR dirumuskan dengan hipotesis berikut
H_0: Tidak ada perbedaan signifikan antara model regresi OLS dan model GWR
H_1: Terdapat perbedaan yang signifikan antara model regresi OLS dan model GWR

Statistik Uji

$$F_{hitung} = \frac{\frac{(SSE(H_0) - SSE(H_1)}{v}}{\frac{SSE(H_1)}{\delta_1}} \tag{2.22}$$

dengan

$SSE(H_0)$ $= \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{y}$

$\boldsymbol{H}$ $= \boldsymbol{X}(\boldsymbol{X^T X})^{-\boldsymbol{1}}\boldsymbol{X^T}$

$SSE(H_1)$ $= \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{y}$

$\boldsymbol{I}$ = Matriks identitas

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{x}_1^T(\boldsymbol{X^T W}(u_1, v_1)\boldsymbol{X})^{-1}\boldsymbol{X^T W}(u_1, v_1) \\ \boldsymbol{x}_2^T(\boldsymbol{X^T W}(u_2, v_2)\boldsymbol{X})^{-1}\boldsymbol{X^T W}(u_2, v_2) \\ \vdots \\ \boldsymbol{x}_n^T(\boldsymbol{X^T W}(u_n, v_n)\boldsymbol{X})^{-1}\boldsymbol{X^T W}(u_n, v_n)y \end{bmatrix}$$

v $= tr(R_0 - R_1)$

δ_1 $= tr(R_1)$

R_0 $= (1 - \boldsymbol{H})^T(1 - \boldsymbol{H})$

R_1 $= (1 - \boldsymbol{L})^T(1 - \boldsymbol{L})$

Dengan kriteria keputusan tolak H_0 jika $F_{hitung} > F_{\alpha,df_1,df_2}$ atau jika *p-value* < 0,05. Derajat bebas yang digunakan yaitu $df_1 = \frac{v^2}{v^*}$ dan $df_2 = \frac{\delta_1^2}{\delta_2}$ dengan $v^* = tr[(R_0 - R_1)^2]$ dan $\delta_2 = tr[(R_1)^2]$.

2.5.3.2 Uji Signifikansi Parameter Model GWR

Uji signifikansi model GWR bertujuan untuk melihat parameter mana saja yang memiliki pengaruh signifikan terhadap variabel respon (Putra, *et al.*, 2022). Uji signifikansi parameter atau uji parsial parameter model GWR dituliskan dalam hipotesis berikut:

H_0: Tidak ada pengaruh variabel prediktor terhadap variabel respon

H_1: Terdapat pengaruh variabel prediktor terhadap variabel respon

Statistik Uji:

$$t_{hitung} = \frac{\hat{\beta}_k(u_i, v_i)}{SE\left[\hat{\beta}_k(u_i, v_i)\right]} \tag{2.23}$$

Dengan SE adalah *Standard Error* dari $\hat{\beta}_k(u_i, v_i)$ dan kriteria keputusannya adalah tolak H_0 jika $|t_{hitung}| > t_{\alpha/_{2}(n-p-1)}$ atau jika *p-value* $< 0{,}05$.

2.6 Regresi Ridge

Regresi Ridge diajukan sebagai solusi untuk mengatasi masalah multikolinearitas dan menjaga kestabilan koefisien regresi. Regresi Ridge mengurangi dampak multikolinieritas dengan memilih estimasi yang memiliki bias namun memiliki varians yang lebih rendah daripada estimasi dalam regresi linear berganda. Metode Regresi Ridge diperoleh dengan cara yang sama seperti metode kuadrat terkecil, yaitu dengan meminimumkan jumlah kuadrat sisaan. Dalam regresi ridge, kendala (tetapan bias) ditambahkan pada kuadrat terkecil, sehingga koefisien menurun dan mendekati nol (Ali & Nugraha, 2019). Pada Regresi Ridge konstanta λ pada diagonal matriks $\boldsymbol{X^T X}$ sehingga koefisien penduga Ridge dipengaruhi oleh besarnya λ. Estimasi parameter pada Regresi Ridge dilakukan dengan menstandarisasi variabel prediktor dan variabel respon (Fadliana, *et al.*, 2019). Standardisasi variabel pada Regresi Ridge adalah dengan menggunakan rumus sebagai berikut:

$$Y_i^* = \frac{1}{\sqrt{n-1}}\left(\frac{Y_i - \bar{Y}}{s_Y}\right) \text{ dan } X_{ik}^* = \frac{1}{\sqrt{n-1}}\left(\frac{X_{ik} - \bar{X}_k}{s_X}\right) \tag{2.24}$$

Menurut Kutner, *et al.*, (2005), Regresi Ridge dituliskan dalam bentuk persamaan di bawah ini:

$$Y_i^* = \beta_1^* X_{1i}^* + \beta_2^* X_{2i}^* + \cdots + \beta_k^* X_{ki}^* + \varepsilon_i^* \tag{2.25}$$

Penduga koefisien pada Regresi Ridge diperoleh dengan cara meminimumkan persaamaan berikut:

$$\hat{\beta}^R = \underset{\beta}{\arg min} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{k=1}^{p} X_{ik}\beta_k \right)^2 + \lambda \sum_{k=1}^{p} \beta_k^2 \right\} \qquad (2.26)$$

Dengan kendala $\sum_{k=1}^{p} \beta_k^2 \geq \lambda$, dimana λ merupakan besaran yang mengendalikan besarnya penyusutan dan nilai $\lambda \geq 0$. Ketika nilai $\lambda = 0$, maka Regresi Ridge akan menghasilkan estimasi yang sama dengan MKT. Ketika $\lambda \rightarrow \infty$ maka estmasi koefisien mendekati nol (Yulita, 2016). Dengan meminimumkan jumlah kuadrat galat maka diperoleh penduga koefisien regresi Ridge adalah sebagai berikut:

$$\hat{\beta}^R = (\boldsymbol{X^T X} + \lambda \boldsymbol{I})^{-1} \boldsymbol{X^T Y} \qquad (2.27)$$

Dengan ***I*** merupakan matriks identitas dan pemilihan nilai λ dapat diperoleh dengan menggunakan metode *Cross Validation* (CV).

2.7 *Locally Compensated Ridge Geographically Weighted Regression* (LCR-GWR)

Metode *Locally Compensated Ridge Geographically Weighted Regression* (LCR-GWR) merupakan pengembangan dari regresi Ridge untuk mengatasi multikolinieritas pada analisis data spasial yang pertama kali diperkenalkan oleh Gollini, *et al*., (2014). Metode ini menggunakan satu koefisien bias untuk suatu lokasi pengamatan tertentu, jika terdapat *n* lokasi pengamatan maka akan terdapat *m* koefisien bias ridge yang berbeda. Metode ini akan menghasilkan koefisien bias Ridge secara lokal. Parameter ridge dibiarkan bervariasi di setiap wilayah menyesuaikan dengan efek multikolinieritas antarvariabel prediktor di setiap wilayah pengamatan sehingga koefisien parameter pada model diharapkan lebih akurat (Fadliana, *et al*., 2019). Model LCR-GWR ditulis dalam bentuk sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, \lambda_i)x_{ik} + \varepsilon_i \tag{2.28}$$

dengan:

y_i = Nilai variabel respon pada lokasi pengamatan ke- i

x_{ik} = Nilai variabel prediktor k pada lokasi pengamatan ke- i

$\beta_0(u_i, v_i)$ = Intersep pada lokasi pengamatan ke- i

$\beta_k(u_i, v_i, \lambda_i)$ = Nilai koefisien regresi lokal ke- k dengan koefisien bias regresi Ridge pada lokasi ke- i

ε_{ik} = Galat pengamatan ke- i

(u_i, v_i) = Titik koordinat dari lokasi pengamatan ke- i

i = $1, 2, 3, \ldots, n$

k = $1, 2, 3, \ldots, p$

2.7.1 Pendugaan Parameter Model LCR-GWR

Pada model LCR-GWR, di mana parameter juga diestimasi dengan prosedur *Weighted Least Square* (WLS), dengan membuat sistem pembobotan bergantung pada lokasi dalam ruang geografis. Estimasi pada moldel LCR-GWR dituliskan dalam bentuk sebagai berikut:

$$\hat{\boldsymbol{\beta}}^{\boldsymbol{LCR}}((u_i, v_i) = \left(\boldsymbol{X}^{*\boldsymbol{T}}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^{*} + \boldsymbol{\lambda I}(u_i, v_i)\right)^{-1}\boldsymbol{X}^{*\boldsymbol{T}}\boldsymbol{W}(u_i, v_i)\boldsymbol{Y}^{*} \tag{2.29}$$

$\lambda\boldsymbol{I}(u_i, v_i)$ yang merupakan nilai *Locally Compensated* (LC) dari λ di wilayah pengamatan (u_i, v_i). Nilai parameter regresi Ridge diperoleh dengan menghubungkan nilai eigen dan *conditional number* (c) yang diperoleh dari perkalian matriks $\boldsymbol{X^T W}(u_i, v_i)\boldsymbol{X}$ yaitu diperoleh dari $\left((\epsilon_1 - \epsilon_p)/(c-1)\right) - \epsilon_p$ Dimana ϵ_1 merupakan nilai eigen terbesar dan ϵ_p adalah nilai eigen terkecil. Sedangkan *Conditional number* (c) diperoleh dari pembagian nilai eigen terbesar dengan nilai eigen terkecil. Menurut Hocking (2003), multikolinieritas dapat dideteksi jika nilai dari *conditional number* (c) lebih besar dari 30. Ketika nilai dari *conditional number* (c) yang diperoleh lebih besar dari 30 memaksa nilai *Locally Compensated* (LC) untuk tidak melebihi ambang batas yang sama, sehingga mengikuti konvensi dan menentukan ambang batas 30 (Gollini, *et al.*, 2014).

2.7.2 Pengujian Hipotesis Model LCR-GWR

Pengujian hipotesis pada model LCR-GWR bertujuan untuk melihat parameter mana saja yang memiliki pengaruh signifikan terhadap variabel respon (Fadliana, *et al.*, 2019). Uji signifikansi parameter atau uji parsial parameter model LCR-GWR dituliskan dalam hipotesis berikut:

H_0: Tidak ada pengaruh variabel prediktor terhadap variabel respon

H_1: Terdapat pengaruh variabel prediktor terhadap variabel respon

Statistik Uji:

$$t_{hitung} = \frac{\hat{\beta}_k(u_i, v_i, \lambda_i)}{SE\left[\hat{\beta}_k(u_i, v_i, \lambda_i)\right]} \tag{2.30}$$

Dengan SE adalah Standard Error dari $\hat{\beta}_k(u_i, v_i,)$ dan kriteria keputusannya adalah tolak H_0 jika $|t_{hitung}| > t_{\alpha/2(n-p-1)}$ atau jika *p-value* < 0,05.

2.8 *Least Absolute Shrinkage and Selection Operator* (LASSO)

Metode *Least Absolute Shrinkage and Selection Operator* (LASSO) dikemukakan pertama kali oleh Tibshirani pada 1996. Menurut Tibshirani (1996), estimasi OLS yang biasa digunakan sering memiliki bias yang kecil namun dengan ragam yang tinggi yang mengakibatkan berkurangnya akurasi prediksi. Akurasi prediksi dapat ditingkatkan dengan menyusutkan beberapa koefisien sampai ke nol. Metode mengecilkan dan menyusutkan koefisien yang berkorelasi tinggi dan tidak signifikan sampai ke nol disebut dengan Metode LASSO. Seperti regresi Ridge yang menyebabkan koefisien menyusut ke arah nol. Namun, pada LASSO beberapa koefisien menyusut sampai ke nol (Herawati, *et al.,* 2018). Estimasi dalam metode LASSO dirumuskan sebagai berikut:

$$\hat{\beta}^L = \underset{\beta}{\arg min} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{k=1}^{p} X_{ik}\beta_k \right)^2 \tag{2.31}$$

Dengan kendala $\sum_{k=1}^{p} |\beta_k| \leq t$. Dimana nilai t merupakan penyusutan parameter yang nilainya ditentukan oleh CV. Dengan kendala tersebut, mengakibatkan adanya penambahan $\lambda \sum_{k=1}^{p} |\beta_k|$ sehingga persamaan (2.21) equivalen dengan:

$$\hat{\beta}^L = \underset{\beta}{\arg min} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{k=1}^{p} X_{ik}\beta_k \right)^2 + \lambda \sum_{k=1}^{p} |\beta_k| \right\} \tag{2.32}$$

Karena nilai absolut dari koefisien regresi dibatasi menyebabkan persamaan menjadi nonlinier sehingga metode OLS tidak dapat menyelesaikannya. Untuk menyelesaikannya diperlukan pemrograman kuadratik dengan algoritma LARS (Setyorini, *et al.*, 2017). Sebelum solusi akhir dari LASSO diperoleh, nilai parameter dari t harus diestimasi terlebih dahulu dengan persamaan sebagai berikut:

$$t = \frac{\sum_{k=1}^{p} |\hat{\beta}_k|}{\sum_{k=1}^{p} \left|\hat{\beta}_k^{OLS}\right|} \tag{2.33}$$

Nilai t dari 0 sampai 1. Jika $\hat{\beta}_k^{OLS}$ estimasi parameter dari metode kuadrat terkecil, maka nilai dari $\sum_{k=1}^{p} |\hat{\beta}_k| < \sum_{k=1}^{p} \left|\hat{\beta}_k^{OLS}\right|$ menyebabkan parameter menjadi nol (Yuliana, 2018).

2.9 *Geographically Weighted Lasso* (GWL)

Metode *Geographically Weighted Lasso* (GWL) adalah penerapan metode LASSO untuk menangani masalah multikolinieritas lokal dan heterogenitas spasial pada model GWR. Estimasi pada model GWL dirumuskan sebagai berikut:

$$\hat{\beta}^L = \underset{\beta}{\arg\min}\left\{\sum_{i=1}^{n}\left(y_i - \beta_0(u_i, v_i) - \sum_{k=1}^{p} X_{ik}\beta_k\,(u_i, v_i)\right)^2 + \lambda\sum_{k=1}^{p}|\beta_k\,(u_i, v_i)|\right\} \tag{2.34}$$

Dengan kendala $\sum_{k=1}^{p}|\beta_k\,(u_i, v_i)| \leq t$ yang harus diboboti pada setiap lokasi sehingga parameter penyusutan akan berbeda setiap lokasinya (Setyorini*, et al*., 2017).

Langkah-langkah dalam melakukan estimasi pada model GWL adalah sebagai berikut:

1. Menghitung nilai *bandwidth* optimum dengan metode CV
2. Dengan menggunakan *bandwidth* optimum yang diperoleh selanjutnya menghitung matriks pembobot $\boldsymbol{W}$ yang berukuran $n \times n$ dengan fungsi eksponensial.
3. Pada setiap lokasi $i = 1, 2, 3, \dots, n$
   a. Menghitung $\boldsymbol{W}_{(\boldsymbol{i})}^{\frac{1}{2}} = \sqrt{\left(diag\left(\boldsymbol{W}_{(\boldsymbol{i})}\right)\right)}$
   b. Menghitung $\boldsymbol{X}_W = \boldsymbol{W}_{(\boldsymbol{i})}^{\frac{1}{2}}\boldsymbol{X}$ dan $\boldsymbol{y}_W = \boldsymbol{W}_{(\boldsymbol{i})}^{\frac{1}{2}}\boldsymbol{y}$ menggunakan akar dari pembobot $\boldsymbol{W}_{(\boldsymbol{i})}$ di setiap lokasi ke- i.
   c. Menentukan solusi dari LASSO yang sesuai dengan CV berdasarkan nilai t pada setiap lokasi ke- i dengan algoritma LARS.

Nilai parameter t pada setiap lokasi ke- i (t_i) dihitung dengan rumus sebagai berikut:

$$t_i = \frac{\sum_{k=1}^{p} |\hat{\beta}_k\,(u_i, v_i)|}{\sum_{k=1}^{p} \left|\hat{\beta}_k^{OLS}\right|} \tag{2.35}$$

2.10 Seleksi Model Terbaik

Menurut Wang*, et al.* (2022), model GWR menggunakan kriteria evaluasi berfokus pada pengukuran kesalahan kuadrat global, termasuk *Root Mean Squared Error* (RMSE) dan *Akaike Information Criterion Score* (AIC).
AIC adalah metode yang paling umum untuk mengevaluasi kinerja kecocokan pada model GWR yang merupakan estimasi yang tidak bias. Dalam AIC nilai yang lebih rendah menunjukkan model yang memiliki kinerja kecocokan yang lebih baik (Herawati, *et al*., 2018). AIC dirumuskan sebagai berikut:

$$AIC = 2k + n \ln\left(\frac{JKG}{n}\right) \tag{2.36}$$

Dimana k merupakan jumlah parameter yang diestimasi, n merupakan jumlah pengamatan, dan $JKG = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ yang merupakan Jumlah Kuadrat Galat.

RMSE merupakan akar kuadrat dari JKG dibagi dengan jumlah pengamatan atau akar dari selisih antara nilai prediksi dan aktual. Nilai RMSE yang kecil menunjukkan prediksi yang akurat (Setiyorini, *et al*., 2017). RMSE dirumuskan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{2.37}$$

2.11 *Gross Domestic Product* (GDP)

Gross Domestic Product (GDP) merupakan total nilai semua jenis barang dan jasa yang dihasilkan oleh suatu negara dalam suatu periode tertentu (biasanya per tahun), sedangkan *Gross Regional Domestic Product* (GRDP) merupakan GDP di suatu wilayah (Sanusi, *et al.,* 2021). Urrutia & Tampis (2017), mendefinisikan GDP sebagai nilai pasar semua barang dan jasa yang diproduksi di dalam batas-batas negara yang menjadi faktor pertumbuhan ekonomi suatu negara. GDP dianggap sebagai salah satu indikator utama dalam menentukan kesehatan ekonomi. GDP paling umum diukur dengan menggunakan metode pengeluaran, dimana GDP dihitung dengan menambahkan pengeluaran untuk barang konsumsi baru, pengeluaran investasi baru, pengeluaran pemerintah, dan nilai ekspor dikurangi dengan nilai impor.

Menurut U.S Bureau of Economic Analysis, perekonomian Amerika Serikat merupakan perkenomian yang terbesar di dunia berdasarkan pengukuran GDP nominal. Sektor jasa ekonomi, yang meliputi keuangan, real estat, asuransi, layanan profesional dan bisnis, serta perawatan kesehatan menjadi contributor terbesar GDP. Amerika Serikat merupakan kekuatan geopolitik dominan di dunia yang mampu mempertahankan hutang luar negeri sebagai produsen mata uang cadangan utama dunia. Menurut Bank Dunia, GDP Nominal Amerika Serikat dalam Dolar AS pada tahun 2022 adalah sebesar $25,46 triliun. GDP Amerika Serikat naik 2,1% pada tahun 2022 (*Economic Comission for Latin America and the Carribean* (ECLAC), 2022). Menurut Hasan, *et al*., (2016), salah satu variabel yang memengaruhi GDP adalah *Labor force*. Menurut Stanic & Racic (2019), inflasi merupakan salah satu faktor yang memengaruhi GDP. Selain *Labor force* dan inflasi, pendidikan termasuk salah satu faktor yang berpengaruh terhadap pertumbuhan GDP (Syed & Shaikh, 2013)

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun ajaran 2023/2024 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang akan digunakan pada penelitian ini adalah data sekunder, yaitu data *Gross Domestic Product* (GDP) Amerika Serikat Tahun 2022 beserta faktor-faktornya yang diperoleh dari Bureau of Economic Analysis U.S Department of Analysis (https://www.bea.gov/). Data berjumlah 51 observasi dengan 11 variabel prediktor dan 1 (satu) variabel respon.

Tabel.1 Variabel Data

Variabel	Keterangan
Y	*Gross Domestic Product* (US Dollar)
X_1	*Income* (US Dollar)
X_2	*Personal Consumption Expenditures* (US Dollar)
X_3	*Unemployment Rate* (Persen)
X_4	*Labor Force Participation Not Seasonally* (Persen)
X_5	*Population Growth Rate* (Persen)
X_6	*Population Density* (Jiwa)

Tabel.1 Lanjutan

X_7	*High School Graduate Rate* (Persen)
X_8	*College Graduate Rate* (Persen)
X_9	*Assosiate, Bachelor, Professional Degree Rate* (Persen)
X_{10}	*Poverty Rate* (Persen)
X_{11}	*Inflation Rate* (Persen)

3.3 Metode Penelitian

Adapun langkah-langkah yang dilakukan pada penelitian ini adalah sebagai barikut:

1. Melakukan analisis dan penyajian data dengan statistika deskriptif.
2. Melakukan identifikasi multikolinieritas dengan nilai *Variance Inflation Factor* (VIF).
3. Melakukan penskalaan dan pemusatan data.
4. Menguji dependensi spasial dengan Uji Z berdasarkan nilai Indeks Moran.
5. Menguji heterogenitas spasial dengan uji Breusch-Pagan.
6. Melakukan pemodelan GWR dengan langkah:
   a. Mencari nilai *bandwidth* (h) optimum dengan menggunakan metode CV pada persamaan (2.18).
   b. Membentuk matriks pembobot $\boldsymbol{W}(u_i, v_i)$ untuk setiap lokasi pengamatan dengan menggunakan nilai *bandwidth* (h) yang telah diperoleh.
   c. Melakukan pendugaan parameter regresi untuk setiap lokasi pengamatan berdasarkan matriks pembobot dengan persamaan (2.15).
7. Melakukan Uji kesesuaian model dengan uji F dan uji signifikansi parameter dengan uji t pada model GWR
8. Melakukan identifikasi multikolinieritas lokal dengan nilai *Variance Inflation Factor* (VIF).
9. Melakukan pemodelan LCR-GWR dengan langkah:
   a. Mencari nilai *bandwidth* (h) optimum dengan menggunakan metode CV

b. Membentuk matriks pembobot $\boldsymbol{W}(u_i, v_i)$ untuk setiap lokasi pengamatan dengan menggunakan nilai *bandwidth* (h) yang telah diperoleh.
c. Menghitung nilai eigen untuk mendapatkan nilai conditional number.
d. Melakukan pendugaan parameter regresi untuk setiap lokasi pengamatan berdasarkan matriks pembobot.

10. Melakukan pemodelan GWL dengan langkah:
    a. Menghitung nilai *bandwidth* (h) optimum dengan metode CV
    b. Dengan menggunakan *bandwidth* (h) optimum yang diperoleh selanjutnya menghitung matriks pembobot $\boldsymbol{W}$ yang berukuran $n \times n$ dengan fungsi eksponensial.
    c. Pada setiap lokasi $i = 1, 2, 3, \dots, n$
       - Menghitung $\boldsymbol{W}_{(\boldsymbol{i})}^{\frac{1}{2}} = \sqrt{\left(diag\left(\boldsymbol{W}_{(i)}\right)\right)}$
       - Menghitung $\boldsymbol{X}_W = \boldsymbol{W}_{(\boldsymbol{i})}^{\frac{1}{2}} \boldsymbol{X}$ dan $\boldsymbol{y}_W = \boldsymbol{W}_{(i)}^{\frac{1}{2}} \boldsymbol{y}$ menggunakan akar dari pembobot $\boldsymbol{W}_{(i)}$ di setiap lokasi ke- i.
       - Melakukan pendugaan parameter dengan algoritma LARS dari $\boldsymbol{X}_W$ dan $\boldsymbol{y}_W$ yang telah diperoleh.
11. Membandingkan nilai RMSE dan AIC yang diperoleh pada model GWR dan GWL untuk mengetahui metode yang terbaik dalam menduga nilai GDP.
12. Kesimpulan

V. KESIMPULAN

Berdasarkan hasil penelitian dalam memodelkan regresi pada data spasial terdapat perbedaan yang signifikan antara model GWR dan model OLS. Sedangkan dalam menangani multikolinieritas lokal pada model GWR, metode GWL mampu menangani multikolinieritas dengan mendapatkan nilai AIC dan RMSE yang lebih kecil dibandingkan model LCR-GWR. Sehingga pada contoh kasus di lokasi pengamatan Maryland diperoleh model regresi yaitu:

$$\hat{y}^L(u_5, v_5) = -0{,}10890X_3 + 0{,}04997X_8 + 0{,}04570X_{10} + 0{,}45762X_{11}$$

Berdasarkan model yang diperoleh untuk lokasi pengamatan ke-5 yaitu Maryland, GDP dipengaruhi oleh *Unemployment Rate* (X_3), *College Graduate Rate* (X_8), *Poverty Rate* (X_{10}), dan *Inflation* (X_{11}).

DAFTAR PUSTAKA

Ali, R. G. & Nugraha, J. 2019. Penerapan Metode Regresi Ridge dalam Mengatasi Masalah Multikolinearitas pada Kasus Indeks Pembangunan Manusia di Indonesia Tahun 2017, hlm. 239-248. Prosiding Sendika. Universitas Muhammadiyah Purworejo, Purworejo.

Amin, C., Sari, D. N., Priyono, K. D., & Hidayah, B. 2023. The Spatial Pattern of COVID-19 Incidence in Relation to Poverty Across Central Java Province, hlm. 450-463. Proceedings of the International Conference of Geography and Disaster Management (ICGDM 2022), Universitas Muhammadiyah Surakarta, Surakarta.

Anselin, L. 1998. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Netherlands.

Boot, B. 2001. *Spatial Pattern*. Elseiver ltd., Amsterdam.

Draper, N. R. & Smith, H. 1998. *Applied Regression Analysis*. John Wiley & Sons, New York.

Fadhilah, N. 2015. Geographically Weighted Regression dan Spatial Pattern Analysis untuk Pemodelan Kejadian Penyakit Malaria dan Faktor yang Mempengaruhi di Provinsi Papua (Skripsi). Jurusan Statistika FMIPA, Institut Teknologi Sepuluh Nopember, Surabaya.

Fadliana, A., Pramoedyo, H., & Fitriani, R. 2020. Implementation of Locally Compensated Ridge-Geographically Weighted Regression Model in Spatial Data with Multicollinearity Problems (Case Study: Stunting among Children Aged under Five Years in East Nusa Tenggara Province). *Media Statistika*. **13**(2): 125-135.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*. John Wiley & Sons, Hoboken.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. 2014. GWmodel: an R Package for Exploring Spatial Heterogeneity using Geographically Weighted Models. *Journal of Statistical Software*. **63**(17): 1-50.

Harlan, J. 2018. *Analisis Regresi Linear*. Gunadarma Pondokcina, Depok.

Hasan, M. N., Rana, S., Malek, M. B., Das, K. R., & Sultana, N. 2016. Modeling Bangladesh's Gross Domestic Product Using Regression Approach. *Malaysian Journal of Mathematical Sciences*. **10**(2): 233–246.

Herawati, N., Nisa, K., Setiawan, E., Nusyirwan, & Tiryono. 2018. Regularized Multiple Regression Methods to Deal with Severe Multicollinearity. *International Journal of Statistics and Applications*. **8**(4): 167-172.

Hidayah, N. R. & Indrasetianingsih, A. 2019. Analisis Regresi Spatial Durbin Model (SDM) untuk Pemodelan Kemiskinan Provinsi Jawa Timur Tahun 2017. *Jurnal Statistika*. **11**(2): 40-46.

Hocking, R. R. 2003. *Methods and Applications of Linear Models*. John Wiley & Sons, New York.

Klar, R. G. 2021. *Geographically Weighted Regression based Investigation of Transport Policies for Increased Public Transport Ridership A Case Study of Stockholm*. Stocholm, Sweden.

Kutner, M. H., Natchtsheim, C. J., Neter, J., & Li, W. 2005. *Applied Linear Statistical Models*. McGraw-Hill, Boston.

Leung, Y., Mei, C. L., & Zhang, W. X. 2000. Statistical Tests for Spatial Nonstationarity Based on The Geographically Weighted Regression Model. *Journal of Environment and Planning A*. 32: 9-32.

Lutfiani, N., Sugiman, & Mariani, S. 2017. Pemodelan Geographically Weighted Regression (GWR) dengan Fungsi Pembobot Kernel Gaussian dan Bi-Square. *UNNES Journal of Mathematics*. **8**(1): 82-91.

Montgomery, D. G., Peck, E. A., & Vining, G. G. 2012. *Introduction to Linear Regression Analysis*. John Wiley & Sons, Inc., New Jersey.

Nadya, M. 2017. Analisis Geographically Weighted Regression (GWR) pada Kasus Pneumonia Balita di Provinsi Jawa Barat (Skripsi). Jurusan Matermatika FMIPA, Universitas Negeri Jakarta, Jakarta.

Putra, R., Tyas, S. W., & Fadhlurrahman, M. G. 2022. Geographically Weighted Regression with The Best Kernel Function on Open Unemployment Rate Data in East Java Province. *Enthusiastic International Journal of Statistics and Data Science*. **2**(1): 26-36.

Sanusi, R. N. M., Saputro, D. R. S., & Setiyowati, R. 2021. Application of Robust Principal Component Analysis for Gross Regional Domestic Product of Provinces in Indonesia. *Journal Physics*. 1776: 1-6.

Setiyorini, A., Suprijadi, J., & Handoko, B. 2017. Implementations of Geographically Weighted Lasso in Spatial Data with Multicollinearity (Case Study: Poverty modeling of Java Island, hlm. 1-14. AIP Conference Proceedings, Universitas Padjajaran, Bandung.

Stanic, S. & Racic, Z. V. 2019. Analysis of Macroeconomic Factors Effect to Gross Domestic Product of Bosnia and Herzegovina Using the Multiple Linear Regression Model. *Economics*. **7**(2): 91-97.

Tibshirani, R. 1996. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. **58**(1): 267-288.

Syed, A. A. S. G. & Shaikh, F. M. 2013. Effects of Macroeconomic Variables on Gross Domestic Product (GDP) in Pakistan, hlm. 703-711. *International Conference on Applied Economics* (ICOAE), Elsevier ltd, Amsterdam.

Urrutia, J. D. & Tampis, R. L. 2017. Regression Analysis of The Economic Factors of The Gross Domestic Product in The Philippines. *Journal of Fundamental and Applied Sciences*. **9**(75): 190-201.

Wang, H., Huang, Z., Yin, G., Bao, Y., Zhou, X., & Gao, Y. 2022. *GWRBoost: A Geographically Weighted Gradient Boosting Method for Explainable Quantification of Spatially-Varying Relationships*. Cornell University, New York.

Wang, J. & Zuo, R. 2020. Assessing Geochemical Anomalies Using Geographically Weighted Lasso. *Applied Geochemistry*. 119: 104668-104668.

Wheeler, D. C. 2009. Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression: The Geographically Weighted Lasso. *Journal of Environment and Planning A*. **41** (3): 722-742.

Yang, H., Xu, T., Chen, D., Yang, H., & Pu, L. 2020. Direct Modeling of Subway Ridership at The Station Level: A Study Based on Mixed Geographically Weighted Regression. *Canadian Journal of Civil Engineering*. **47**(5): 534-545.

Yan, X. & Su, X. G. 2009. *Linear Regression Analysis Theory and Computing*. World Scientific Publishing Co. Pte. Ltd., Singapore.

Yuliana. 2018. Model Geographically Weighted Least Absolute Shrinkage and Selection Operator (GW-LASSO) pada Data Penderita Diare di Indonesia (Skripsi). Jurusan Matematika FMIPA, Universitas Sebelas Maret, Surakarta.

Yulita, T. 2016. Pemodelan Geographically Weighted Ridge Regression dan Geographically Weighted Lasso pada Data Spasial dengan Multikolinieritas (Thesis). Sekolah Pascasarjana, Institut Pertanian Bogor, Bogor.