

ARTIKEL PENELITIAN

Pemodelan *Geographically Weighted Ridge Regression* Pada Tingkat Pengangguran Terbuka di Kalimantan Barat

Ferry Adrian¹, Yundari^{2,*}, Siti Aprizkiyandari³

¹Program Studi Statistika, Universitas Tanjungpura, Pontianak-KalBar, Indonesia

²Program Studi Matematika, Universitas Tanjungpura, Pontianak-KalBar, Indonesia

³Program Studi Agroteknologi, Universitas Tanjungpura, Pontianak-KalBar, Indonesia

*Penulis korespondensi: yundari@math.untan.ac.id

Diterima: 16 Juli 2023; Direvisi: 14 Agustus 2023; Disetujui: 15 Agustus 2023; Dipublikasi: 18 September 2023.

Abstrak: *Geographically Weighted Ridge Regression* (GWRR) adalah pengembangan dari regresi *ridge* yaitu dengan menambahkan unsur pembobot sebagai informasi tambahan. GWRR digunakan untuk mengatasi data spasial yang mengandung multikolinearitas lokal dengan memodifikasi regresi *ridge* dan GWR. Tujuan penelitian ini adalah untuk membentuk model GWRR dan menganalisis keefektifan model GWRR dalam mengatasi masalah multikolinearitas pada data spasial. Data penelitian yang digunakan terdiri dari variabel respon yaitu nilai tingkat pengangguran terbuka (Y) di kabupaten/kota Provinsi Kalimantan Barat tahun 2021 dan variabel penjelas yang digunakan adalah upah minimum kabupaten/kota (X_1), jumlah penduduk (X_2), persentase penduduk miskin (X_3), angkatan kerja (X_4), dan rata-rata lama sekolah (X_5). Proses analisis terlebih dahulu dilakukan dengan pemodelan regresi linear berganda dan pemodelan GWR. Setelah itu dilakukan transformasi pemusatan dan penskalaan pada data, dilanjutkan dengan pemodelan GWRR untuk mengatasi masalah multikolinearitas pada data spasial. Setelah didapatkan semua model dilakukan perbandingan model dilihat dari nilai MSE. Berdasarkan hasil penelitian diperoleh 14 model GWRR dengan pembagian tiga kelompok wilayah berdasarkan faktor-faktor yang mempengaruhi tingkat pengangguran terbuka untuk setiap kabupaten/kota di Provinsi Kalimantan Barat. Pada perbandingan model yang digunakan, jika dilihat dari nilai MSE, model GWRR memiliki nilai MSE paling besar dibandingkan model regresi linear dan model GWR. Meskipun nilai MSE pada model GWRR paling besar diantara ketiga model, namun pada model GWRR masalah multikolinearitas sudah dapat teratasi.

Kata Kunci: Data Spasial, Multikolinearitas, Matriks Pembobot, Regresi Spasial

Abstract: *Geographically Weighted Ridge Regression* (GWRR) is a development of ridge regression by adding a weighting element as additional information. GWRR is used to overcome spatial data containing local multicollinearity by modifying the ridge regression and GWR. The purpose of this study was to construct the GWRR model and analyze the effectiveness of the GWRR model in overcoming multicollinearity problems in spatial data. The research data used consists of response variables, namely the value of the open unemployment rate (Y) in districts/cities of West Kalimantan Province in 2021 and the explanatory variables used are district/city minimum wage (X_1), population (X_2), percentage of poor population (X_3), labor force (X_4), and average years

of schooling (X_5). The analysis process was first carried out by using multiple linear regression modeling and GWR modeling. After that, centering and scaling transformations were carried out on the data, followed by GWRR modeling to overcome the problem of multicollinearity in spatial data. After all the models were obtained, a model comparison was made in terms of the MSE value. Based on the results of the study, 14 GWRR models were obtained by dividing into three regional groups based on the factors that affect the open unemployment rate for each district/city in West Kalimantan Province. In the comparison of the models used, when viewed from the MSE value, the GWRR model has the highest MSE value compared to the linear regression model and the GWR model. Even though the MSE value in the GWRR model is the largest among the three models, in the GWRR model the problem of multicollinearity can be resolved.

Keywords: Spatial Data, Multicollinearity, Weighting Matrix, Spatial Regression

1. Pendahuluan

Tingkat pengangguran merupakan isu ekonomi yang dihadapi oleh banyak negara, termasuk Indonesia. Hampir semua negara menghadapi tantangan pengangguran. Salah satu faktor yang menyebabkan pengangguran adalah pertumbuhan populasi yang cepat, yang menyebabkan peningkatan jumlah angkatan kerja setiap tahunnya, sementara lapangan pekerjaan yang tersedia tidak bertambah sebanding [1]. Pengangguran merujuk pada kondisi di mana individu yang termasuk dalam angkatan kerja berkeinginan untuk bekerja tetapi belum berhasil menemukan pekerjaan. Orang yang tidak bekerja dan tidak aktif mencari pekerjaan tidak dianggap sebagai pengangguran [2].

Berdasarkan informasi yang disajikan oleh Badan Pusat Statistik Kalimantan Barat, terdapat tren peningkatan dalam tingkat pengangguran terbuka (TPT) di Kalimantan Barat dari tahun 2019 hingga 2021, dengan variasi peningkatan yang berbeda-beda antar kabupaten/kota [3]. TPT di suatu daerah diyakini dipengaruhi oleh TPT di daerah sekitarnya. Fenomena ini terjadi karena adanya faktor-faktor yang terkait dengan kedekatan geografis atau hubungan antar daerah tersebut. Keberadaan pengaruh spasial pada variabel respon menyebabkan perkiraan menjadi tidak akurat karena asumsi keacakan galat tidak terpenuhi, yang mengakibatkan adanya ketidakkonstanan dalam variansi galat yang dihasilkan (heteroskedastisitas). Oleh karena itu, diperlukan penerapan model regresi spasial untuk mengatasi masalah ini [4].

Geographically Weighted Regression (GWR) merupakan suatu bentuk model regresi spasial yang digunakan untuk memodelkan hubungan antara variabel respon dan variabel penjelas pada data yang menunjukkan adanya ketergantungan spasial. Dalam model GWR, matriks pembobot yang diterapkan memiliki besaran yang ditentukan oleh kedekatan geografis antar lokasi pengamatan yang terlibat. Pada penerapan model GWR, terkadang muncul situasi di mana terdapat korelasi antar variabel penjelas, situasi ini dikenal sebagai multikolinearitas lokal. Jika terdapat multikolinearitas lokal dalam data spasial, pendekatan *Geographically Weighted Ridge Regression* (GWRR) dapat diterapkan untuk mengatasi masalah ini. Metode GWRR merupakan perluasan dari metode regresi *ridge* dan GWR, yang melibatkan unsur pembobot sebagai komponen informasi tambahan dalam analisis [5].

Dalam penelitian ini, penulis mengaplikasikan metode GWRR dalam memodelkan tingkat pengangguran terbuka di kabupaten/kota Provinsi Kalimantan Barat tahun 2021. Selain itu, metode GWRR digunakan untuk mengatasi masalah multikolinearitas pada data TPT di kabupaten/kota Provinsi Kalimantan Barat tahun 2021.

2. Metode Penelitian

Penelitian yang dilakukan adalah penelitian kuantitatif karena data yang digunakan dalam penelitian ini berbentuk angka. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistika (BPS) Provinsi Kalimantan Barat berupa data tingkat pengangguran terbuka (TPT) dan faktor-faktor yang mempengaruhi tingkat pengangguran terbuka yang meliputi seluruh kabupaten/kota di Kalimantan Barat pada tahun 2021. Variabel penjelas yang digunakan dalam penelitian ini adalah upah minimum kabupaten/kota, jumlah penduduk, persentase penduduk miskin, angkatan kerja, dan rata-rata lama sekolah. Adapun langkah-langkah penelitian yang dilakukan pada penelitian ini adalah:

1. Mengambil data sekunder dari Badan Pusat Statistik Kalimantan Barat tahun 2021.
2. Menentukan variabel penjelas (X) dan variabel respon (Y)
 - X_1 : Upah minimum kabupaten/kota (Rupiah).
 - X_2 : Jumlah penduduk (Jiwa).
 - X_3 : Persentase penduduk miskin (Persen).
 - X_4 : Angkatan kerja (Jiwa).
 - X_5 : Rata-rata lama sekolah (Tahun).
 - Y : Tingkat pengangguran terbuka (Persen).
3. Menentukan model regresi linear [6]

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k X_{ik} + \varepsilon_i; i = 1, 2, \dots, n; k = 1, 2, \dots, p$$

dengan y_i merupakan variabel respon pengamatan ke- i , x_{ik} merupakan variabel penjelas ke- k pengamatan ke- i , β_k merupakan koefisien regresi dari variabel x_k , β_0 merupakan *intercept*, dan ε_i merupakan *error* random ke- i

4. Melakukan uji secara simultan [6]

$$F_{hitung} = \frac{MSR}{MSE}$$

Pengambilan keputusan paling tidak ada satu parameter yang signifikan dalam model jika $F_{hitung} > F_{\alpha, p, n-p-1}$, dengan n adalah banyaknya observasi dan p adalah banyaknya variabel penjelas.

5. Melakukan uji parsial untuk model regresi linear [6]

$$t_{hitung} = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)}$$

dengan $se(\hat{\beta}_k) = \sqrt{var\hat{\beta}_k}$ dan $var\hat{\beta}_k = C_{kk}\sigma^2$. Untuk C_{kk} = elemen diagonal dari matriks $(\mathbf{X}^T\mathbf{X})^{-1}$. Parameter model dikatakan signifikan jika $|t_{hitung}| > t_{\alpha/2, n-p-1}$, dengan n adalah banyaknya observasi dan p adalah banyaknya variabel penjelas.

6. Menguji asumsi klasik pada model regresi linear [7]
7. Menguji Heterogenitas spasial [8]

$$BP = \left(\frac{1}{2}\right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f}$$

dengan elemen vektor f dirumuskan $f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$ dimana e_i merupakan residual untuk pengamatan ke- i , matriks f merupakan matriks berukuran $n \times (p+1)$ yang sudah distandarisasi untuk setiap pengamatan, dan σ^2 merupakan ragam dari e_i . Terjadi heterogenitas spasial jika $BP > \chi_{(a;p)}^2$ atau nilai $p - value < 0,05$ dengan p adalah banyaknya variabel penjelas.

8. Menghitung jarak *Euclidean* dengan rumus $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ [9].
9. Menentukan *bandwidth* optimum (h) menggunakan metode *Cross Validation* (CV) berdasarkan nilai CV minimum [6]

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

dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir dimana pengamatan di lokasi- i dihilangkan dari proses penaksiran.

10. Menghitung matriks pembobot untuk masing-masing wilayah menggunakan fungsi pembobot *kernel fixed Gaussian* dengan rumus [9]:

$$w_{ij} = \exp\left(-\frac{1}{2} \left(\frac{d_{ij}}{h}\right)^2\right)$$

dengan h adalah nilai *bandwidth* dan d_{ij} merupakan jarak *Euclidean*

11. Menentukan model *Geographically Weighted Regression* (GWR) [6]

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

dengan y_i merupakan nilai observasi variabel respon ke- i , x_{ik} merupakan nilai observasi variabel penjelas ke- k di lokasi pengamatan ke- i , $\beta_0(u_i, v_i)$ adalah konstanta/*intercept* pada pengamatan ke- i , (u_i, v_i) merupakan titik koordinat letak geografis lokasi pengamatan ke- i , $\beta_k(u_i, v_i)$ adalah koefisien regresi variabel penjelas ke- k di lokasi pengamatan ke- i , dan ε_i merupakan *error* pengamatan ke- i

12. Melakukan uji parsial model GWR [6]

$$t_{hit} = \frac{\beta_k(\hat{u}_i, \hat{v}_i)}{\hat{\sigma} \sqrt{C_{kk}}}$$

dengan C_{kk} adalah elemen diagonal ke- k dari matriks $\mathbf{C}\mathbf{C}^T$. Untuk $\mathbf{C} = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i)$ dan $\hat{\sigma} = \frac{SSE(H_1)}{\partial_1}$. Jika diberikan tingkat signifikansi α , maka parameter model diindikasikan signifikan jika $|t_{hit}| \geq t_{\alpha/2; df}$ dengan derajat bebas $df = \frac{\partial_1^2}{\partial_2}$.

13. Mendeteksi multikolinearitas pada model GWR [10]

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

dengan $R_k^2(u_i, v_i)$ adalah koefisien determinasi variabel ke- k pada lokasi ke- i . Terjadi multikolinearitas jika nilai $VIF > 10$.

14. Menentukan koefisien *ridge* menggunakan metode *generalized cross validation* (GCV) [11]

$$GCV = \frac{\sum_{i=1}^n \varepsilon_{i,c}^2}{[n - (1 + \text{tr}(\mathbf{H}_c))]^2}$$

dengan $\varepsilon_{i,c}^2$ merupakan sisaan kuadrat ke- i untuk nilai c tertentu dan \mathbf{H}_c merupakan matriks $\mathbf{X}^* (\mathbf{X}^{*T} \mathbf{X}^* + c\mathbf{I})^{-1} \mathbf{X}^{*T}$.

15. Melakukan centering dan scaling pada variabel penjelas dan variabel respon [10]

$$\mathbf{X}^* = \begin{bmatrix} \left(\frac{x_{11}-\bar{x}_1}{S_1\sqrt{n-1}} \right) & \left(\frac{x_{21}-\bar{x}_2}{S_2\sqrt{n-1}} \right) & \cdots & \left(\frac{x_{p1}-\bar{x}_p}{S_p\sqrt{n-1}} \right) \\ \left(\frac{x_{12}-\bar{x}_1}{S_1\sqrt{n-1}} \right) & \left(\frac{x_{22}-\bar{x}_2}{S_2\sqrt{n-1}} \right) & \cdots & \left(\frac{x_{p2}-\bar{x}_p}{S_p\sqrt{n-1}} \right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{x_{1n}-\bar{x}_1}{S_1\sqrt{n-1}} \right) & \left(\frac{x_{2n}-\bar{x}_2}{S_2\sqrt{n-1}} \right) & \cdots & \left(\frac{x_{pn}-\bar{x}_p}{S_p\sqrt{n-1}} \right) \end{bmatrix}; \mathbf{y}^* = \begin{bmatrix} \left(\frac{y_1-\bar{y}}{S_y\sqrt{n-1}} \right) \\ \left(\frac{y_2-\bar{y}}{S_y\sqrt{n-1}} \right) \\ \vdots \\ \left(\frac{y_n-\bar{y}}{S_y\sqrt{n-1}} \right) \end{bmatrix}$$

dengan \mathbf{X}^* adalah matriks variabel penjelas hasil transformasi berukuran $n \times p$, \mathbf{y}^* adalah vektor variabel respon hasil transformasi berukuran $n \times 1$, dimana $\bar{x}_k = \frac{\sum_{i=1}^n x_{ki}}{n}$ yang merupakan rata-rata dari masing-masing variabel penjelas ke- k , dengan $k = 1, 2, \dots, p$ dan $i = 1, 2, \dots, n$; $S_k = \sqrt{\sum_{i=1}^n \frac{(x_{ki}-\bar{x}_k)^2}{n-1}}$; $S_y = \sqrt{\sum_{i=1}^n \frac{(y_i-\bar{y})^2}{n-1}}$; dan $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ yang merupakan rata-rata dari masing-masing variabel respon ke- i , dengan $i = 1, 2, \dots, n$.

16. Menentukan model *Geographically Weighted Ridge Regression* (GWRR)
 17. Melakukan uji parsial model GWRR [10]

$$t_{hit} = \frac{\hat{\beta}_k(u_i, v_i, k)}{\hat{\sigma} \sqrt{C_{kk}}}$$

dengan C_{kk} adalah elemen diagonal ke- k dari matrik $\mathbf{C}\mathbf{C}^T$. Untuk $\mathbf{C} = (\mathbf{X}^{*T}\mathbf{W}(u_i, v_i)\mathbf{X}^* + c\mathbf{I})^{-1}\mathbf{X}^T\mathbf{W}(u_i, v_i)$ dan $\hat{\sigma}$ adalah akar dari $\hat{\sigma}^2 = \frac{SSE}{n-2tr(\mathbf{G})+tr(\mathbf{G}^T\mathbf{G})}$. Jika tingkat signifikansi α , maka parameter model diindikasikan signifikan jika $|t_{hit}| \geq t_{\alpha/2;df}$ dimana $df = \left(\frac{u_1^2}{u_2} \right)$.

18. Mendeteksi multikolinearitas pada model GWRR
 Nilai VIF model GWRR merupakan nilai dari elemen diagonal utama *inverse* dari $(\mathbf{X}^{*T}\mathbf{W}\mathbf{X}^* + c\mathbf{I})$ dengan c merupakan koefisien *ridge* optimum [10]. Terjadi multikolinearitas jika nilai $VIF > 10$.
 19. Membandingkan model yang digunakan dengan metode *Mean Square Error* (MSE) [12]

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n}$$

dengan y_i adalah data aktual pada pengamatan ke- i , sedangkan \hat{y}_i adalah hasil prediksi pada pengamatan ke- i , dan n adalah jumlah observasi.

3. Hasil dan Pembahasan

3.1. Analisis Data

Data yang digunakan merupakan data sekunder yang didapat dari Badan Pusat Statistik Kalimantan Barat tahun 2021. Data ini terdiri dari 14 kabupaten/kota di Provinsi Kalimantan Barat. Statistik deskriptif pada data dapat dilihat pada Tabel 3.1.

Tabel 3.1: Statistik Deskriptif

Variabel	Minimum	Rata-rata	Maksimum
Y	2,66	5,41	12,38
X ₁	2.422.594	2.555.311	2.860.323
X ₂	128.550	390.771	663.713
X ₃	4,34	7,47	12,01
X ₄	55.171	188.269	317.161
X ₅	6,02	7,30	10,43

Pada Tabel 3.1 memberikan informasi bahwa wilayah dengan nilai TPT tertinggi berada di Kota Pontianak sebesar 12,38%, sedangkan Kabupaten Melawi merupakan wilayah dengan nilai

TPT paling rendah yaitu sebesar 2,66%. Rata-rata nilai TPT dari 14 kabupaten/kota di Provinsi Kalimantan Barat pada tahun 2021 sebesar 5,41%. Sementara itu Kabupaten Ketapang merupakan wilayah dengan nilai UMK tertinggi yaitu Rp. 2.860.323, sedangkan Kabupaten Mempawah wilayah dengan nilai UMK terendah yaitu Rp. 2.422.594. Penduduk dengan jumlah tertinggi berada di wilayah Kota Pontianak sebanyak 663.713 jiwa, sedangkan Kabupaten Kayong Utara wilayah dengan jumlah penduduk terendah sebesar 128.550 jiwa. Sedangkan untuk wilayah dengan persentase penduduk miskin tertinggi berada di Kabupaten Melawi yaitu sebanyak 12,01% dan wilayah dengan persentase penduduk miskin terendah berada di Kabupaten Kubu Raya yaitu sebanyak 4,34%. Jumlah angkatan kerja Kota Pontianak memiliki jumlah paling tinggi yaitu 317.161 jiwa dan Kabupaten Kayong Utara memiliki jumlah angkatan kerja paling rendah yaitu 55.171 jiwa. Sementara itu, wilayah dengan rata-rata lama sekolah tertinggi berada di Kota Pontianak dengan nilai 10,43 tahun, sedangkan Kabupaten Kayong Utara wilayah dengan rata-rata lama sekolah terendah dengan nilai 6,02 tahun. Berdasarkan deskripsi data diketahui bahwa TPT yang berada di wilayah Kota Pontianak cenderung memiliki nilai TPT yang tinggi, dibandingkan nilai TPT pada Kabupaten Melawi yang memiliki nilai terendah. Hal ini terjadi karena Kota Pontianak merupakan ibu kota Provinsi Kalimantan Barat yang mempunyai penduduk wilayah yang padat namun belum dapat menyediakan lapangan pekerjaan yang memadai dibandingkan Kabupaten Melawi yang jumlah penduduknya masih belum padat, sehingga dalam hal ini dapat disimpulkan bahwa diduga terdapat pengaruh wilayah yang mempengaruhi tinggi rendahnya nilai TPT di Kalimantan Barat.

3.2. *Pemodelan Regresi Linear*

Pemodelan regresi linear ini bertujuan untuk melakukan analisis awal pada data TPT untuk mengetahui adanya pengaruh antara variabel penjelas dan variabel respon. Pendugaan parameter pada regresi linear menggunakan metode *Ordinary Least Square* (OLS). Berdasarkan dugaan awal parameter diperoleh model regresi linear sebagai berikut:

$$\hat{y} = -15,21 + 0,000003x_1 + 0,000023x_2 - 0,3405x_3 - 0,000051x_4 + 2,089x_5$$

Setelah didapatkan model regresi linear dilakukan uji secara simultan. Berdasarkan uji F diperoleh nilai $F_{hitung} = 8,661$ dan $p - value = 0,004$ dengan $\alpha = 0,05$ sehingga dapat disimpulkan bahwa secara simultan paling tidak ada satu parameter regresi yang signifikan dalam model regresi. Selanjutnya dilakukan uji parsial dengan menggunakan uji t yang dapat dilihat pada Tabel 3.2.

Tabel 3.2: Uji Parsial Model Regresi Linear

Variabel	$ t_{hitung} $	Keputusan
X_1	0,791	tidak signifikan
X_2	1,985	tidak signifikan
X_3	1,774	tidak signifikan
X_4	2,111	tidak signifikan
X_5	4,392	signifikan

Berdasarkan Tabel 3.2, bahwa hanya variabel X_5 yang signifikan dengan nilai $t_{hitung} > t_{(0,025;8)} = 2,751$. Hal tersebut mengindikasikan hanya variabel rata-rata lama sekolah yang mempengaruhi TPT di Provinsi Kalimantan Barat. Selanjutnya dilakukan pengujian asumsi klasik pada model regresi linear untuk mengetahui kelayakan model. Uji normalitas dengan uji *Shapiro-Wilk* menunjukkan bahwa galat berdistribusi normal dengan nilai $p - value = 0,8959$ lebih dari $\alpha = 0,05$. Pada uji autokorelasi dapat dilakukan dengan pengujian *Durbin Watson* (DW), diperoleh nilai *Durbin-Watson* (DW) sebesar 2,003 yang berada di rentang nilai 1,65~2,35 yang menunjukkan tidak adanya autokorelasi.

Selanjutnya dilakukan uji multikolinearitas dengan melihat *variance inflation factor* (VIF). Uji multikolinearitas dipenuhi jika nilai VIF kurang dari 10.

Tabel 3.3: Nilai VIF Model Regresi Linear

Variabel	X_1	X_2	X_3	X_4	X_5
VIF	1,526	27,725	1,632	26,811	1,449

Berdasarkan Tabel 3.3, hasil nilai VIF menunjukkan bahwa terjadi multikolinearitas pada variabel X_2 dan X_4 dengan nilai VIF lebih besar dari 10. Pada uji heteroskedastisitas dilakukan dengan uji *Breusch-Pagan*. Berdasarkan uji *Breusch-Pagan* diperoleh nilai $p - value = 0,048 < \alpha = 0,05$ maka dapat disimpulkan bahwa terjadi heteroskedastisitas pada data. Pada pengujian asumsi klasik model regresi linear, didapatkan kesimpulan bahwa uji normalitas, uji multikolinearitas dan uji heteroskedastisitas terpenuhi, sedangkan uji autokorelasi tidak terpenuhi, sehingga diperlukan metode lain untuk menangani masalah asumsi klasik ini.

Dikarenakan data yang digunakan memenuhi asumsi heteroskedastisitas, maka diduga terdapat heterogenitas spasial sehingga dilakukan pengujian dengan menggunakan uji *Breusch Pagan* dan diperoleh hasil nilai $BP = 11,19 > \chi^2_{(0,05;5)} = 11,07$, sehingga dapat disimpulkan bahwa terdapat pengaruh heterogenitas spasial di setiap lokasi pengamatan pada taraf nyata $\alpha = 0,05$. Hal tersebut mengakibatkan variansi sisaan tidak homogen dan terdapat indikasi pengelompokan wilayah. Oleh karena itu, model regresi tidak tepat digunakan untuk pemodelan TPT di Kalimantan Barat, sehingga digunakan pendekatan model lain menggunakan regresi spasial yaitu model GWR.

3.3. *Geographically Weighted Regression (GWR)*

Langkah awal untuk memperoleh model GWR pada setiap lokasi diperlukan nilai *bandwidth* optimum yang diperoleh dengan metode *Cross Validation (CV)* yang selanjutnya digunakan untuk memperoleh matriks pembobot pada proses pendugaan parameternya. Fungsi pembobot yang digunakan adalah fungsi *kernel fixed gaussian*. Diperoleh nilai *bandwidth* optimum sebesar 1,8231 dengan CV minimum sebesar 44,407. Selanjutnya nilai *bandwidth* yang diperoleh digunakan untuk menentukan matriks pembobot dalam melakukan estimasi parameter dengan menggunakan metode *Weighted Least Square (WLS)*. Fungsi pembobotnya adalah sebagai berikut:

$$w_{ij} = \exp \left(-\frac{1}{2} \left(\frac{d_{ij}}{1,8231} \right)^2 \right)$$

Metode WLS digunakan untuk estimasi parameter, selanjutnya dilakukan uji secara parsial untuk melihat variabel yang signifikan pada model GWR. Variabel yang signifikan dari model GWR berbeda untuk setiap wilayahnya. Hal ini dapat dilihat dari nilai statistik uji t jika variabel penjelas memiliki $|t_{hitung}| > t_{(0,025;8)} = 2,751$. Variabel penjelas yang berpengaruh signifikan terhadap variabel respon dikelompokkan menjadi tiga kelompok. Pengelompokkan variabel signifikan pada model GWR dapat dilihat pada Tabel 3.4.

Tabel 3.4: Variabel yang Signifikan Dalam Model GWR

Kelompok	Kabupaten/Kota	Variabel yang Signifikan
1	Sintang, Kapuas Hulu, Sekadau, Melawi	X_5
2	Sambas, Bengkayang, Ketapang, Kota Singkawang	X_4 dan X_5
3	Landak, Mempawah, Sanggau, Kayong Utara, Kubu Raya, Kota Pontianak	X_2, X_4 dan X_5

Tabel 3.5: Model GWR Tiap Kabupaten/Kota di Provinsi Kalimantan Barat

Kab/Kota	Model GWR
Sambas	$\hat{y}_1 = -10,836 - 0,000080x_{1,4} + 2,249x_{1,5}$
Bengkayang	$\hat{y}_2 = -10,382 - 0,000081x_{2,4} + 2,180x_{2,5}$
Landak	$\hat{y}_3 = -11,759 + 0,000035x_{3,2} - 0,000076x_{3,4} + 2,148x_{3,5}$
Mempawah	$\hat{y}_4 = -10,614 + 0,000035x_{4,2} - 0,000076x_{4,4} + 2,148x_{4,5}$
Sanggau	$\hat{y}_5 = -13,700 + 0,000030x_{5,2} - 0,000066x_{5,4} + 2,102x_{5,5}$
Ketapang	$\hat{y}_6 = -12,137 - 0,000063x_{6,4} + 1,907x_{6,5}$
Sintang	$\hat{y}_7 = -21,574 + 2,010x_{7,5}$
Kapuas Hulu	$\hat{y}_8 = -23,732 + 2,108x_{8,5}$
Sekadau	$\hat{y}_9 = -15,176 + 2,066x_{9,5}$
Melawi	$\hat{y}_{10} = -16,950 + 1,973x_{10,5}$
Kayong Utara	$\hat{y}_{11} = -11,485 + 0,000034x_{11,2} - 0,000072x_{11,4} + 1,990x_{11,5}$
Kubu Raya	$\hat{y}_{12} = -10,823 + 0,000037x_{12,2} - 0,000079x_{12,4} + 2,063x_{12,5}$
Pontianak	$\hat{y}_{13} = -10,795 + 0,000037x_{13,2} - 0,000080x_{13,4} + 2,101x_{13,5}$
Singkawang	$\hat{y}_{14} = -10,380 + 0,000081x_{14,4} + 2,202x_{14,5}$

Berdasarkan hasil pengelompokkan variabel yang signifikan, maka dapat dibentuk permodelan GWR untuk masing-masing kabupaten/kota. Model GWR untuk masing-masing kabupaten/kota dapat dilihat pada Tabel 3.5.

Karena terdapat beberapa faktor yang mempengaruhi TPT tiap wilayah sehingga memungkinkan terjadinya masalah multikolinearitas, sehingga perlu dilakukan uji multikolinearitas terhadap model GWR. Adanya multikolinearitas ditunjukkan dengan nilai VIF yang lebih besar dari 10. Seperti terlihat pada Tabel 3.6.

Tabel 3.6: Nilai VIF Model GWR

Kab/Kota	Nilai VIF				
	X_1	X_2	X_3	X_4	X_5
Sambas	1,341	44,458	1,600	45,278	1,475
Bengkayang	1,431	47,922	1,679	48,801	1,528
Landak	1,461	40,943	1,638	41,086	1,512
Mempawah	1,478	46,883	1,696	47,550	1,540
Sanggau	1,493	34,171	1,592	33,490	1,514
Ketapang	1,926	41,188	1,712	39,509	1,703
Sintang	1,615	17,841	1,418	15,520	1,457
Kapuas Hulu	1,524	16,444	1,417	14,681	1,330
Sekadau	1,520	30,127	1,561	28,909	1,526
Melawi	1,647	25,832	1,517	23,672	1,591
Kayong Utara	1,753	42,953	1,739	42,261	1,630
Kubu Raya	1,624	46,028	1,750	46,180	1,586
Pontianak	1,554	46,102	1,725	46,478	1,563
Singkawang	1,401	47,585	1,658	48,510	1,515

Berdasarkan Tabel 3.6, diketahui pada variabel X_2 dan X_4 untuk semua kabupaten/kota memiliki nilai VIF lebih besar dari 10 sehingga dapat disimpulkan terjadi masalah multikolinieritas pada model GWR. Dikarenakan terdapat multikolinieritas pada model GWR, sehingga diperlukan metode lain untuk mengatasi masalah multikolinieritas ini yaitu menggunakan model GWRR untuk analisis selanjutnya.

3.4. *Geographically Weighted Ridge Regression (GWRR)*

Model GWRR digunakan untuk mengatasi multikolinieritas pada model regresi lokal yaitu model GWR. Pendugaan parameter model GWRR menggunakan *Weighted Least Square (WLS)* yang telah ditambah dengan nilai parameter *ridge* sehingga melibatkan matriks pembobot. Matriks pembobot yang digunakan sama seperti matriks pembobot untuk model GWR dan nilai parameter *ridge* optimum diperoleh sebesar 0,085 dengan nilai GCV minimum 0,196. Dalam model ini digunakan variabel hasil yang sudah dilakukan pemusatan dan penskalaan. Pemusatan dan penskalaan adalah bagian dari membakukan (*standardized*) variabel. Pemusatan artinya perbedaan antara masing-masing pengamatan dengan rata-rata dari seluruh pengamatan untuk variabel. Sedangkan penskalaan meliputi gambaran pengamatan pada kesatuan (unit) standar deviasi dari pengamatan untuk variabel [13]. Variabel hasil pemusatan dan penskalaan dapat dilihat pada Tabel 3.7.

Tabel 3.7: Variabel Hasil Pemusatan dan Penskalaan

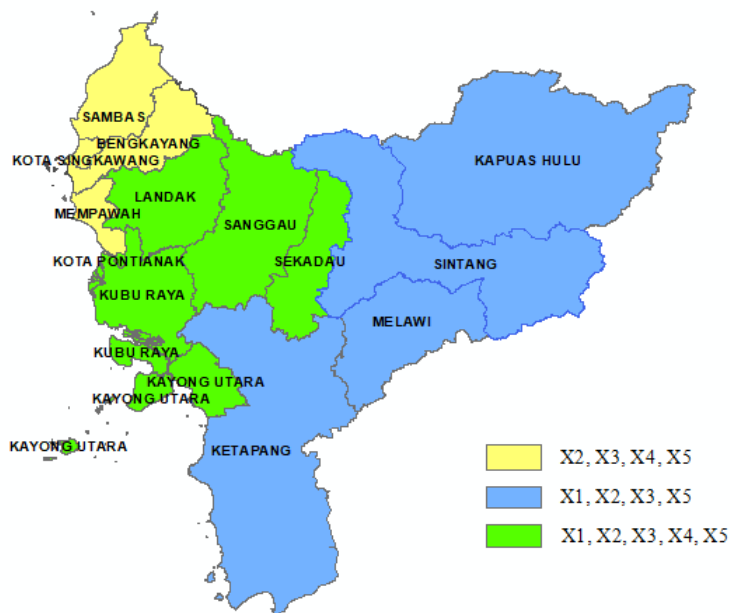
No	Y^*	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*
1	-0,141	0,106	0,382	0,020	0,332	-0,161
2	-0,097	0,025	-0,154	-0,104	-0,181	-0,139
3	-0,214	-0,013	0,016	0,371	0,033	-0,051
4	0,225	-0,315	-0,132	-0,241	-0,220	-0,073
5	-0,192	-0,095	0,151	-0,308	0,204	0,024
6	0,149	0,723	0,293	0,280	0,209	0,043
7	-0,143	0,098	0,051	0,191	0,181	-0,062
8	-0,120	-0,171	-0,212	0,154	-0,135	0,063
9	-0,244	-0,224	-0,275	-0,127	-0,258	-0,126
10	-0,269	-0,171	-0,247	0,478	-0,239	-0,109
11	-0,159	0,378	-0,406	0,196	-0,433	-0,355
12	0,157	-0,290	0,347	-0,329	0,344	-0,084
13	0,681	-0,096	0,422	-0,304	0,420	0,865
14	0,366	0,044	-0,237	-0,278	-0,254	0,165

Selanjutnya nilai *bandwidth* dan nilai parameter *ridge* yang diperoleh digunakan dalam menentukan estimasi parameter pada model GWRR. Setelah itu dilakukan uji secara parsial untuk melihat variabel apa saja yang signifikan pada model GWRR. Variabel yang signifikan pada model GWRR berbeda untuk setiap wilayahnya. Hal ini dapat dilihat dari nilai statistik uji t jika variabel penjelas memiliki $|t_{hitung}| > t_{0,025;8} = 2,751$ maka variabel penjelas berpengaruh signifikan terhadap variabel respon. Variabel penjelas yang berpengaruh signifikan terhadap variabel respon dikelompokkan menjadi tiga kelompok. Pengelompokan variabel signifikan pada model GWRR dapat dilihat pada Tabel 3.8.

Tabel 3.8: Variabel yang Signifikan Dalam Model GWRR

Kelompok	Kabupaten/Kota	Variabel yang Signifikan
1	Sambas, Bengkayang, Mempawah, dan Kota Singkawang	$X_2, X_3, X_4,$ dan X_5
2	Ketapang, Sintang, Kapuas Hulu, dan Melawi	$X_1, X_2, X_3,$ dan X_5
3	Landak, Sanggau, Sekadau, Kayong Utara, Kubu Raya, dan Kota Pontianak	$X_1, X_2, X_3, X_4,$ dan X_5

Untuk melihat pengelompokan variabel yang signifikan pada model GWRR setiap wilayah dapat dilihat pada Gambar 3.1. Pemodelan GWRR memiliki perbedaan pada setiap kabupaten/kota di Kalimantan Barat, sehingga kabupaten/kota yang memiliki kesamaan signifikan terhadap model dikelompokkan. Pada Gambar 3.1 menunjukkan beberapa wilayah yang berdekatan memiliki kesamaan variabel yang signifikan yang menjelaskan adanya pengaruh spasial dalam pemodelan GWRR pada TPT di Kalimantan Barat.



Gambar 3.1: Pembagian Wilayah variabel yang Signifikan Pada Model GWRR

Berdasarkan hasil pengelompokan variabel yang signifikan, maka dapat dibentuk permodelan GWRR untuk masing-masing kabupaten/kota. Proses pembentukan X^* ke X dapat dilakukan berdasarkan nilai penduga parameter dengan menggunakan hubungan linear [14].

$$\beta_1 = \left(\frac{S_Y}{S_{xk}} \right) \beta_1^*, k = 1, 2, \dots, \beta_0 = \bar{Y} - \beta_1 \bar{X}_1 - \beta_2 \bar{X}_2 - \dots - \beta_k \bar{X}_k$$

Parameter penduga yang diperoleh dengan metode GWRR dalam bentuk baku, kemudian akan dikembalikan ke dalam bentuk awal. Model awal GWRR untuk masing-masing kabupaten/kota dapat dilihat pada Tabel 3.9.

Tabel 3.9: Model Awal GWRR Tiap Kabupaten/Kota di Provinsi Kalimantan Barat

Kab/Kota	Model GWR
Sambas	$\hat{y}_1 = -9,191 + 0,000002x_{1,2} - 0,418x_{1,3} - 0,000007x_{1,4} + 1,798x_{1,5}$
Bengkayang	$\hat{y}_2 = -10,035 + 0,000002x_{2,2} - 0,432x_{2,3} - 0,000008x_{2,4} + 1,791x_{2,5}$
Landak	$\hat{y}_3 = -13,667 + 0,000004x_{3,1} + 0,000003x_{3,2} - 0,397x_{3,3} - 0,000009x_{3,4} + 1,806x_{3,5}$
Mempawah	$\hat{y}_4 = -11,560 + 0,000003x_{4,2} - 0,425x_{4,3} - 0,000008x_{4,4} + 1,789x_{4,5}$
Sanggau	$\hat{y}_5 = -16,185 + 0,000004x_{5,1} + 0,000004x_{5,2} - 0,355x_{5,3} - 0,000008x_{5,4} + 1,777x_{5,5}$
Ketapang	$\hat{y}_6 = -16,119 - 0,000005x_{6,1} + 0,000003x_{6,2} - 0,318x_{6,3} + 1,559x_{6,5}$
Sintang	$\hat{y}_7 = -15,397 + 0,000005x_{7,1} + 0,000003x_{7,2} - 0,272x_{7,3} + 1,194x_{7,5}$
Kapuas Hulu	$\hat{y}_8 = -14,056 + 0,000005x_{8,1} + 0,000003x_{8,2} - 0,272x_{8,3} + 1,194x_{8,5}$
Sekadau	$\hat{y}_9 = -17,154 + 0,000005x_{9,1} + 0,000004x_{9,2} - 0,328x_{9,3} - 0,000007x_{9,4} + 1,729x_{9,5}$
Melawi	$\hat{y}_{10} = -17,217 + 0,000005x_{10,1} + 0,000003x_{10,2} - 0,292x_{10,3} + 1,570x_{10,5}$
Kayong Utara	$\hat{y}_{11} = -15,907 + 0,000005x_{11,1} - 0,000004x_{11,2} - 0,361x_{11,3} - 0,000007x_{11,4} + 1,686x_{11,5}$
Kubu Raya	$\hat{y}_{12} = -14,172 + 0,000004x_{12,1} - 0,000003x_{12,2} - 0,403x_{12,3} - 0,000008x_{12,4} + 1,749x_{12,5}$
Pontianak	$\hat{y}_{13} = -13,222 + 0,000004x_{13,1} - 0,000003x_{13,2} - 0,413x_{13,3} - 0,000008x_{13,4} + 1,773x_{13,5}$
Singkawang	$\hat{y}_{14} = -9,432 + 0,000002x_{14,2} - 0,431x_{14,3} - 0,000008x_{14,4} + 1,793x_{14,5}$

Selanjutnya dilakukan pengujian asumsi multikolinearitas pada model GWRR untuk melihat apakah pada model GWRR masih terjadi multikolinearitas atau tidak. Nilai VIF pada model GWRR dapat dilihat pada Tabel 3.10.

Tabel 3.10: Nilai VIF Model GWRR

Kab/Kota	Nilai VIF				
	X_1	X_2	X_3	X_4	X_5
Sambas	2,620	5,869	2,149	5,826	1,661
Bengkayang	2,224	5,825	2,023	5,738	1,479
Landak	1,751	5,695	1,712	5,590	1,417
Mempawah	1,961	5,776	1,899	5,677	1,415
Sanggau	1,566	5,684	1,615	5,531	1,597
Ketapang	1,481	6,046	2,109	5,937	2,242
Sintang	2,829	5,549	2,475	5,933	4,424
Kapuas Hulu	3,879	6,755	2,989	6,253	5,070
Sekadau	1,561	5,745	1,644	5,534	1,862
Melawi	1,678	6,013	1,863	5,667	2,644
Kayong Utara	1,423	5,822	1,828	5,712	1,663
Kubu Raya	1,592	5,760	1,811	5,652	1,432
Pontianak	1,713	5,749	1,814	5,641	1,396
Singkawang	2,366	5,843	2,074	5,679	1,526

Pada Tabel 3.10 terlihat bahwa nilai $VIF < 10$ untuk semua variabel pada setiap kabupaten/kota. Sehingga dapat disimpulkan sudah tidak terjadi multikolinearitas pada model GWRR.

3.5. Perbandingan Model Yang Digunakan

Dalam membandingkan model yang digunakan dilihat dari nilai MSE yang terkecil antara model regresi linier berganda, GWR dan GWRR, diperoleh hasil sebagai berikut:

Tabel 3.11: Hasil Perbandingan Model Regresi Linear, GWR, dan GWRR

Model	Regresi Linear	GWR	GWRR
MSE	1,165	0,700	1,421

Berdasarkan Tabel 3.11, model GWR memiliki nilai MSE (0,700) paling kecil jika dibandingkan dengan model regresi linier dan GWRR [15]. Meskipun nilai MSE pada model GWRR paling besar diantara ketiga model, namun pada model GWRR masalah multikolinearitas sudah berhasil teratasi dengan tidak adanya nilai VIF yang lebih dari 10 sehingga model GWRR terbukti mampu mengatasi masalah multikolinearitas pada data spasial.

4. Kesimpulan

Berdasarkan hasil dan pembahasan diperoleh kesimpulan bahwa terdapat 14 model GWRR untuk masing-masing kabupaten/kota di Provinsi Kalimantan Barat. Dari 14 model GWRR diperoleh tiga kelompok wilayah berdasarkan faktor-faktor yang mempengaruhi TPT untuk setiap kabupaten/kota di Provinsi Kalimantan Barat. TPT untuk Kabupaten Sambas, Bengkayang, Mempawah, dan Kota Singkawang dipengaruhi oleh faktor jumlah penduduk, persentase penduduk miskin, angkatan kerja, dan rata-rata lama sekolah. Sementara, untuk Kabupaten Ketapang, Sintang, Kapuas Hulu, dan Melawi dipengaruhi faktor UMK, jumlah penduduk, persentase penduduk miskin, dan rata-rata lama sekolah. Sedangkan untuk Kabupaten Landak, Sanggau, Sekadau, Kayong Utara, Kubu Raya dan Kota Pontianak dipengaruhi oleh faktor UMK, jumlah penduduk, persentase penduduk miskin, angkatan kerja, dan rata-rata lama sekolah. Pada perbandingan model yang digunakan, jika dilihat dari nilai MSE, model GWRR memiliki nilai MSE paling besar dibandingkan model regresi linear dan model GWR. Meskipun nilai MSE pada model GWRR paling besar diantara ketiga model, namun pada model GWRR masalah multikolinearitas sudah dapat teratasi.

Referensi

- [1] E. Amalia and L. K. Sari, "Analisis spasial untuk mengidentifikasi tingkat pengangguran terbuka berdasarkan kabupaten/kota di pulau jawa tahun 2017," *Indonesian Journal of Statistics and Its Applications*, vol. 3, no. 3, pp. 202–215, 2019. [View online](#).
- [2] R. Helvira and E. P. Rizki, "Pengaruh investasi, upah minimum dan ipm terhadap tingkat pengangguran terbuka di provinsi kalimantan barat," *JIsEB*, vol. 1, no. 1, pp. 53–62, 2020. [View online](#).
- [3] B. P. K. Barat, *Kalimantan Barat Dalam Angka 2022*. BPS Provinsi Kalimantan Barat, 2022. [View online](#).
- [4] M. Agustina, B. Abapihi, G. N. A. Wibawa, I. Yahya, *et al.*, "Pemodelan faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di indonesia dengan pendekatan regresi spasial," in *Seminar Nasional Sains dan Terapan VI*, vol. 6, pp. 56–70, 2022. [View online](#).
- [5] Y. Arthayanti, I. G. A. M. Srinadi, and G. Gandhiadi, "Geographically weighted ridge regression dalam kasus multikolinearitas pada indeks pembangunan manusia di kabupaten/kota provinsi jawa timur," *Jurnal Matematika Vol*, vol. 7, no. 2, pp. 124–131, 2017. [View online](#).
- [6] R. A. Pamungkas, H. Yasin, and R. Rahmawati, "Perbandingan model gwr dengan fixed dan adaptive bandwidth untuk persentase penduduk miskin di jawa tengah," *Jurnal Gaussian*, vol. 5, no. 3, pp. 535–544, 2016. [View online](#).
- [7] A. R. I. M. Sari and F. G. Putra, "Pengaruh sisa hasil usaha pada koperasi menggunakan regresi linear berganda," *MAJU: Jurnal Ilmiah Pendidikan Matematika*, vol. 7, no. 2, pp. 110–120, 2020. [View online](#).
- [8] M. Nadya, *Analisis Geographically Weighted Regression (Gwr) Pada Kasus Pneumonia Balita Di Provinsi Jawa Barat*. PhD thesis, UNIVERSITAS NEGERI JAKARTA, 2017. [View online](#).
- [9] P. Astuti, N. N. Debataraja, and E. Sulistianingsih, "Analisis kemiskinan dengan pemodelan geographically weighted regression (gwr) di provinsi nusa tenggara timur," *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 7, no. 3. [View online](#).

- [10] D. Veronica, H. Yasin, and T. Widiharih, "Pemodelan pendapatan asli daerah (pad) di kabupaten dan kota di jawa tengah menggunakan geographically weighted ridge regression," *Jurnal Gaussian*, vol. 5, no. 3, pp. 383–393, 2016. [View online](#).
- [11] D. R. S. Saputro, R. D. Hastutik, and P. Widyaningsih, "The modeling of human development index (hdi) in papua—indonesia using geographically weighted ridge regression (gwrr)," in *AIP Conference Proceedings*, vol. 2326, AIP Publishing, 2021. [View online](#).
- [12] D. Kurniasih, S. Mariani, and S. Sugiman, "Efisiensi relatif estimator fungsi kernel gaussian terhadap estimator polinomial dalam peramalan usd terhadap jpy," *UNNES Journal of Mathematics*, vol. 2, no. 2, 2013. [View online](#).
- [13] N. Anggraini, D. Kusnandar, and N. N. Debatara, "Metode generalized ridge regression dalam mengatasi multikolinearitas," *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 8, no. 4, 2019. [View online](#).
- [14] N. A. Abdullah, Amran, and Saleh, "Pemodelan geographically weighted regression multikolinearitas menggunakan regresi ridge," *Jurnal Universitas Hasanuddin*, 2017. [View online](#).
- [15] D. Wirwana, M. N. Aidi, and A. Fitrianto, "analysis of factors affecting district/city grdp in kalimantan island," *International Journal of Sciences*, vol. 64, no. 1, 2022. [View online](#).

Format Sitasi IEEE:

F. Adrian, Yundari, S. Aprizkiyandari, "Pemodelan *Geographically Weighted Ridge Regression* Pada Tingkat Pengangguran Terbuka Di Kalimantan Barat", *Jurnal Diferensial*, vol. 5(2), pp. 83–95, 2023.

This work is licensed under a [Creative Commons "Attribution-ShareAlike 4.0 International"](#) license.

