Estimasi, Vol. 4, No. 1, Januari, 2023, Hal. 100-113
DOI: 10.20956/ejsa.vi.25426

Pemodelan *Mixed Geographically Weighted Regression* yang Mengandung Multikolinearitas dengan Regresi Ridge

Suritman[1*], Raupong[2], Anisa Kalondeng[3]
[123]Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin, Kota Makassar, 90245, Indonesia
*** Corresponding author, email: suritmancs@gmail.com**

Abstract

In the Mixed Geographically Weigted Regression (MGWR) model, some variables are local and some are global. In MGWR modeling, it is often found that the data has multicollinearity. To overcome this problem, MGWR models with ridge regression are used. The MGWR model can be applied to poverty cases because it can experience spatial heterogeneity due to differences in geographical, cultural, and economic policies that vary in each region. In this study, the estimation of MGWR model parameters with ridge regression is then applied to data on the poor population of South Sulawesi in 2016. Data on the poor population of South Sulawesi experience multicollinearity, so it is solved using the MGWR model with ridge regression. Variables that have a significant effect globally are x_3 and x_6. while the variables that have a significant local effect are $x_2, x_4, x_5, x_7, x_8, x_9$ and x_{10}. The AIC value of the MGWR model with ridge regression of 63.64473 is smaller than the MGWR model, meaning that the addition of ridge regression to the MGWR model makes the model better in overcoming multicollinearity problems.

Keywords: *Mixed Geographically Weigted Regression, Ridge Regression, Multicollinearity, Poor Population.*

Abstrak

Pada model Mixed Geographically Weigted Regression (MGWR) sebagian variabel bersifat lokal dan sebagian lainnya bersifat global. Pada pemodelan MGWR sering didapati data mengalami multikolinearitas, untuk mengatasi hal tersebut digunakan model MGWR dengan regresi ridge. Model MGWR dapat diterapkan pada kasus kemiskinan karena dapat mengalami heterogenitas spasial akibat perbedaan geografis, budaya, dan kebijakan ekonomi yang bervariasi di setiap wilayah. Pada penelitian ini estimasi parameter model MGWR dengan regresi ridge kemudian diterapkan pada data penduduk miskin Sulawesis Selatan Tahun 2016. Data penduduk miskin Sulawesi Selatan mengalami multikolinearitas, sehingga diselesaikan menggunakan model MGWR dengan regresi ridge. Variabel yang berpengaruh signifikan secara global adalah x_3 dan x_6. sementara variabel yang berpengaruh signifikan secara lokal adalah x_2, x_4, x_5, x_7, x_8, x_9 dan x_{10}. Nilai AIC model MGWR dengan regresi ridge sebesar 63.64473 lebih kecil dari model MGWR, artinya penambahan regresi ridge pada model MGWR menjadikan model lebih baik dalam mengatasi masalah multikolinearitas.

Kata Kunci: Mixed Geographically Weigted Regression, Regresi Ridge, Multikolinearitas, Penduduk Miskin.

1. Pendahuluan

Dalam model *Geographically Weighted Regression* (GWR) sering didapati tidak semua variabel berpengaruh secara lokal, akan tetapi ada variabel yang berpengaruh secara global. Untuk mengatasi pengaruh lokal dan global variabel prediktor, digunakan model *Mixed Geographically Weigted Regression* (MGWR) [1]. Kekurangan pada model GWR dan MGWR adalah tidak dapat mengatasi multikolinieritas. Multikolinieritas dapat diatasi dengan metode regresi *ridge*. Regresi *ridge* adalah metode untuk mengendalikan kestabilan penduga kuadrat terkecil yang mengalami multikoliniearitas dengan cara menambahkan koefisien tetapan bias positif pada proses pendugaan parameter, sehingga hasil yang diperoleh memiliki ragam yang lebih kecil daripada hasil dengan menggunakan metode kuadrat terkecil [2].

Berdasarkan hasil Susenas Maret 2017, jumlah penduduk miskin di Sulawesi Selatan keadaan Maret 2017 berjumlah 813,07 ribu jiwa atau 9,38 persen dari total penduduk. Jumlah penduduk miskin di Provinsi Sulawesi Selatan mengalami mengalami kenaikan sebesar 6,04 ribu jiwa jika dibandingkan kondisi Maret 2016 yang besarnya 9,40 persen atau 807,03 ribu jiwa. Penduduk miskin di Provinsi Sulawesi Selatan sebagian besar berada di daerah perdesaan. Pada Maret 2017 penduduk miskin di perdesaan mencapai 12,59 persen sedangkan di perkotaan sebesar 4,48 persen.

Berdasarkan penjelasan yang telah diuraikan, peneliti tertarik mengambil atau melakukan penelitian dengan judu "Pemodelan Mixed Geographically Weighted Regression yang Mengandung Multikolinearitas dengan Regresi Ridge", untuk menentukan model MGWR dengan regresi *ridge* data penduduk miskin kabupaten/kota di Provinsi Sulawesi Selatan pada tahun 2016.

2. Material dan Metode

Data yang digunakan adalah data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Sulawesi Selatan dan Profil Dinas Kesehatan Sulawesi Selatan. Penelitian ini juga menggunakan data letak astronomi yang meliputi letak lintang dan letak bujur tiap kabupaten/kota di Sulawesi Selatan. Lokasi penelitian terdiri dari 24 kabupaten/kota di Provinsi Sulawesi Selatan. Variabel respon yang digunakan adalah persentase penduduk miskin di kabupaten/kota Provinsi Sulawesi Selatan pada tahun 2016 (y). Adapun variabel prediktor yang digunakan yaitu rumah tangga memiliki sumber penerangan bukan listrik (x_1), rumah tangga menurut luas lantai rumah < 8m^2 (x_2), rumah tangga memiliki dinding terluas bukan tembok dan kayu (x_3), penduduk usia 15 tahun ke atas (<SD) (x_4), penduduk usia 15 tahun ke atas yang tidak bekerja (x_5), angka harapan hidup (x_6), pengeluaran perkapita untuk makanan (x_7), jumlah fasilitas kesehatan (x_8), rata-rata lama sekolah (x_9), dan kepadatan penduduk (x_{10}).

2.1 Model GWR

Model GWR adalah pengembangan dari model regresi dimana setiap parameter dihitung pada setiap titik [5]. Model GWR dapat ditulis sebagai berikut:

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

$$\boldsymbol{Y} = \boldsymbol{X\beta}(u_i, v_i) + \boldsymbol{\varepsilon} \tag{2}$$

y_i : nilai pengamatan variabel respon untuk lokasi ke-i

$\beta_0(u_i, v_i)$: intersep pada lokasi pengamatan ke-i

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

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

ε_i : galat pengamatan ke- i

Estimasi parameter model GWR adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X'W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X'W}(u_i, v_i)\boldsymbol{Y} \tag{3}$$

2.2 Uji Variabilitas Model GWR

Uji variabilitas adalah uji yang dilakukan untuk megetahui variabel yang bersifat lokal dan variabel yang bersifat global. Pengujian ini dilakukan dengan Hipotesis sebagai berikut:

H_0 : $\beta_{1k}(u_1, v_i) = \beta_{2k}(u_2, v_2) = \cdots = \beta_{nk}(u_n, v_n)$ untuk setiap $k = 1,2, \dots, p$

H_1 : minimal ada satu $\beta_{ik}(u_i, v_i) \neq 0$ untuk setiap $i = 1,2, \dots, n$

Statistik uji:

$$F = \frac{\dfrac{V_k^2}{tr\left(\frac{1}{n}\boldsymbol{B}_k^t\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\right)\boldsymbol{B}_k}}{\dfrac{SSE(H_1)}{b_1}} \sim F(a, p) \tag{4}$$

dengan $\boldsymbol{B}_k = \boldsymbol{e}_k^t[\boldsymbol{X'W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X'W}(u_i, v_i), \boldsymbol{\beta}_k(u_i, v_i) = \left[\hat{\beta}_1(u_i, v_i) \dots \hat{\beta}_k(u_i, v_i)\right]'$, $V_k^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_k(u_i, v_i) - \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_k(u_i, v_i)\right)^2 = \frac{1}{n}\boldsymbol{\beta}'_k\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\boldsymbol{\beta}_k$,

$b_t = tr([(\boldsymbol{I} - \boldsymbol{L})'(\boldsymbol{I} - \boldsymbol{L})]^t), t = 1,2$ serta $c_t = tr\left(\frac{1}{n}\boldsymbol{B}'_k\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\boldsymbol{B}'_k\right)^t, t = 1,2.$ $\boldsymbol{J}$ adalah matriks berukuran ($n \times n$) yang semua elemennya adalah 1 dan $\boldsymbol{e}_k$ adalah vektor kolom berukuran (p + 1) yang bernilai satu untuk elemen ke-k dan nol untuk lainnya. Maka H_0 ditolak apabila nilai $F \geq F_{\alpha,df1,df2}$, artinya variabel prediktor bersifat lokal. Sebaliknya apabila nilai $F < F_{\alpha,df1,df2}$, artinya variabel prediktor bersifat global. Dimana $df_1 = \frac{c_1^2}{c_2}$ dan $df_2 = \frac{b_1^2}{b_2}$.

2.3 Regresi *Ridge*

Metode regresi *ridge* didasarkan pada modifikasi metode kuadrat terkecil, yakni dengan menambahkan suku $\lambda\boldsymbol{I}$ pada $(\boldsymbol{X}'\boldsymbol{X})$ sebelum diinverskan sehingga menyebabkan multikolinearitas melemah [3]. Estimasi parameter regresi *ridge* didefinisikan sebagai berikut:

$$(\boldsymbol{X}'\boldsymbol{X} + \lambda\boldsymbol{I})\widehat{\boldsymbol{\beta}}_R = \boldsymbol{X}'\boldsymbol{Y} \text{ atau } \widehat{\boldsymbol{\beta}}_R = (\boldsymbol{X}'\boldsymbol{X} + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{Y} \tag{5}$$

Untuk pemilihan nilai konstan λ dapat digunakan metode Estimator Lawless & Wang (LW) yang diperoleh dengan rumus sebagai berikut:
untuk tetapan bias global:

$$\lambda = \frac{p\hat{\sigma}^2}{\widehat{\boldsymbol{\beta}}'_{OLS}(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}}_{OLS}} \tag{6}$$

untuk tetapan bias lokal:

$$\lambda(u_i, v_i) = \frac{p\hat{\sigma}^2}{\widehat{\boldsymbol{\beta}}'_{WLS}(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}}_{WLS}} \tag{7}$$

dengan p merupakan jumlah parameter $\hat{\beta}$ tanpa $\hat{\beta}_0$, $\hat{\sigma}^2$ merupakan *Mean square error (MSE)*, $\widehat{\boldsymbol{\beta}}_{OLS}$ merupakan estimasi yang diperoleh dengan *OLS*, dan $\hat{\beta}_{WLS}$ merupakan estimasi yang diperoleh dengan WLS.

2.4 Model MGWR

Model MGWR dengan variabel prediktor sebanyak p, variabel prediktor yang bersifat global sebanyak q, dan variabel prediktor yang bersifat lokal sebanyak $(p - q)$, serta lokasi pengamatan sebanyak i dan mengasumsikan bahwa intersep model bersifat lokal dapat dituliskan sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{q} \beta_k x_{ik} + \sum_{k=q+1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \tag{8}$$

$$\boldsymbol{Y} = \boldsymbol{X}_g\boldsymbol{\beta}_g + \boldsymbol{X}_l\boldsymbol{\beta}_l(u_i, v_i) + \boldsymbol{\varepsilon} \tag{9}$$

dengan $\boldsymbol{X}_g$ adalah matriks variabel prediktor bersifat global, $\boldsymbol{X}_l$ adalah matriks variabel prediktor bersifat lokal, $\boldsymbol{\beta}_g$ adalah vektor penduga parameter bersifat global, dan $\boldsymbol{\beta}_l(u_i, v_i)$ adalah vektor penduga parameter bersifat lokal [6].
Estimasi parameter untuk model MGWR adalah:

$$\widehat{\boldsymbol{\beta}}_l(u_i, v_i) = (\boldsymbol{X}'_l\boldsymbol{W}(u_i, v_i)\boldsymbol{X}_l)^{-1}\boldsymbol{X}'_l\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} \tag{10}$$

$$\hat{\beta}_g = \left[X_g'(I - S_l)'(I - S_l)X_g\right]^{-1} X_g'(\boldsymbol{I} - \boldsymbol{S}_l)'(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{Y} \tag{11}$$

dengan

$$\boldsymbol{S}_l = \begin{bmatrix} \left[x_{i1}'\boldsymbol{W}(u_1, v_1)\boldsymbol{X}_g\right]^{-1} \boldsymbol{X}_l'\boldsymbol{W}(u_1, v_1) \\ \left[x_{i2}'\boldsymbol{W}(u_2, v_2)\boldsymbol{X}_g\right]^{-1} \boldsymbol{X}_l'\boldsymbol{W}(u_2, v_2) \\ \vdots \\ \left[x_{in}'\boldsymbol{W}(u_n, v_n)\boldsymbol{X}_g\right]^{-1} \boldsymbol{X}_l'\boldsymbol{W}(u_n, v_n) \end{bmatrix}$$

2.4.1 Uji Kesesuaian Model MGWR

Uji hipotesis ini membandingkan model MGWR dengan model regresi linear. Pengujian ini dilakukan dengan hipotesis berikut:

H_0 : $\beta_k(u_i, v_i) = \beta_k$ untuk setiap $k = q+1, q+2, \dots, p$ dan $i = 1,2, \dots, n$

H_1 : minimal ada satu $\beta_{q+r}(u_i, v_i) \neq \beta_k$

Statistik uji:

$$F_1 = \frac{\dfrac{\boldsymbol{Y}'[(\boldsymbol{I}-\boldsymbol{H})-(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}]}{v_1}}{\dfrac{\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}}{u_i}} \sim F_{(a,p)} \tag{12}$$

dengan $\boldsymbol{S} = \boldsymbol{S}_l + (\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\left[\boldsymbol{X}_g'(\boldsymbol{I} - \boldsymbol{S}_l)'(\boldsymbol{I} - \boldsymbol{S}_l)\right]^{-1} \boldsymbol{X}'_g(\boldsymbol{I} - \boldsymbol{S}_l)'(\boldsymbol{I} - \boldsymbol{S})$ dan nilai $u_t = tr([(\boldsymbol{I} - \boldsymbol{S})'(\boldsymbol{I} - \boldsymbol{S})^t]), t = 1,2$ serta $v_t = tr([(\boldsymbol{I} - \boldsymbol{H})(\boldsymbol{I} - \boldsymbol{S})'(\boldsymbol{I} - \boldsymbol{S})]^t), t = 1,2$. Maka H_0 ditolak apabila nilai $F_1 < F_{a,df_1,df_2}$ dimana $df_1 = \frac{v_1^2}{v_2}$ dan $df_2 = \frac{u_1^2}{u_2}$.

2.4.2 Uji Serentak Parameter Model MGWR

Uji ini digunakan untuk menguji secara serentak bagaimana signifikansi dari variabel-variabel model MGWR. Ada dua pengujian yang dilakukan, pertama adalah pengujian hipotesis serentak pada parameter variabel prediktor global [7].

H_0 : $\beta_1 = \beta_2 = \dots = \beta_q = 0$

H_1 : minimal ada satu $\beta_k \neq 0$ untuk setiap $k = 1,2, \dots, q$

Statistik uji:

$$F_2 = \frac{\dfrac{\boldsymbol{Y}'[(\boldsymbol{I}-\boldsymbol{S}_l)'(\boldsymbol{I}-\boldsymbol{S}_l)-(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}]}{r_1}}{\dfrac{\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}}{u_1}} \sim F_{(a,p)} \tag{12}$$

Dengan $r_t = tr([(I - S)'(I - S) - (I - S)]^t), t = 1,2$. H_0 ditolak apabila nilai $F_2 \geq F_{a,df_1,df_2}$. Dimana $df_1 = \frac{r_1^2}{r_2}$ dan $df_2 = \frac{u_1^2}{u_2}$.

Selanjutnya uji hipotesis serentak yang kedua dilakukan pada parameter variabel prediktor lokal dengan hipotesis sebagai berikut:

H_0 : $\beta_{q+1}(u_i, v_i) = \beta_{q+2}(u_i, v_i) = \dots = \beta_p(u_i, v_i) = 0$ untuk setiap $i = 1,2, \dots n$

H_1 : minimal ada satu $\beta_k(u_i, v_i) \neq 0$ untuk setiap $k = q+1, q+2, \dots, p$

Statistik uji:

$$F_3 = \frac{\frac{\boldsymbol{Y}'\left[(\boldsymbol{I}-\boldsymbol{S}_g)'(\boldsymbol{I}-\boldsymbol{S}_g)-(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\right]\boldsymbol{Y}}{t_1}}{\frac{\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}}{u_i}} \sim F_{(a,p)} \tag{13}$$

dengan $t_t = tr([(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S}) - (\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})]^t), t = 1,2$. Maka H_0 ditolak apabila nilai $F_3 \geq F_{\alpha,df_1,df_2}$. Dimana $df_1 = \frac{t_1^2}{t_2}$ dan $\mathrm{df}_2 = \frac{u_1^2}{u_2}$.

2.4.3 Uji Parsial Parameter Model MGWR

Uji ini digunakan untuk mengetahui variabel global dan lokal yang berpengaruh signifikan terhadap respon pada model MGWR. Pada pengujian ini akan di lakukan dua kali, yang pertama pengujian signifikansi variabel global dan yang kedua pengujian signifikansi pada variabel lokal [7]. Untuk pengujian signifikansi pada variabel global digunakan hipotesis yaitu:

H_0 : $\beta_k = 0$ untuk setiap $k = 1,2, \dots, q$

H_1 : $\beta_k \neq 0$

Statistik uji:

$$F_3 = \frac{\frac{\boldsymbol{Y}'\left[(\boldsymbol{I}-\boldsymbol{S}_g)'(\boldsymbol{I}-\boldsymbol{S}_g)-(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\right]\boldsymbol{Y}}{t_1}}{\frac{\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}}{u_1}} \sim F_{(a,p)} \tag{14}$$

dengan g_{kk} adalah elemen diagonal ke k dari hasil perkalian matriks $\boldsymbol{GG}'$. Matriks $\boldsymbol{G} = \left[X_g'(I-S)'(I-S)X_g\right]^{-1}X'_g(I-S_l)'(I-S)$ dan $\hat{\sigma}^2 = \frac{\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{Y}}{tr((\boldsymbol{I}-\boldsymbol{S})'(\boldsymbol{I}-\boldsymbol{S}))}$. Maka tolak H_0 jika $|T_{g\_hit}| > T_{a/2,df}$. Dimana $df = \frac{u_1^2}{u_2}$.

Uji hipotesis selanjutnya ditunjukkan untuk mengetahuai variabel lokal yang berpengaruh signifikan terhadap respon pada model *MGWR*. Untuk menguji signifikansi suatu variabel lokal digunakan hipotesis yaitu:

H_0 : $\beta_k(u_i, v_i) = 0$ untuk setiap $k = q+1, q+2, \dots p$ dan $i = 1,2, \dots n$

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

Statistik uji:

$$T_{l\_hit} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma}\sqrt{m_{kk}}} \sim T_{(a/2,df)} \tag{15}$$

dengan m_{kk} adalah elemen diagonal ke k dari hasil perkalian matriks $\boldsymbol{MM}'$. Matriks $\boldsymbol{M} = [\boldsymbol{X}_l'\boldsymbol{W}(u_i, v_i)\boldsymbol{X}_l]^{-1}\boldsymbol{X}_l'\boldsymbol{W}(u_i, v_i)(\boldsymbol{I} - \boldsymbol{X}_g\boldsymbol{G})$ Maka tolak H_0 jika $|T_{t\_hit}| > T_{a/2,df}$.

2.5 Uji Serentak Parameter Model MGWR

Ada beberapa metode yang dapat digunakan untuk memilih model terbaik, salah satunya adalah *Akaike Information Criterion* (AIC) yang didefinisikan sebagai berikut:

$$AIC = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n + tr(\boldsymbol{S}) \tag{16}$$

dengan $\hat{\sigma}$ merupakan nilai estimator standar deviasi dari bentuk residual dan $\boldsymbol{S}$ merupakan matriks proyeksi dari model. Pemilihan model terbaik dilakukan dengan menentukan nilai AIC terkecil [6].

3. Hasil dan Diskusi

3.1 Estimasi Parameter Model MGWR dengan Regresi *Ridge*

Model MGWR diasumsikan mengandung multikoliniearitas sehingga penyelesaiannya dapat menggunakan model MGWR dengan regresi *ridge*. Untuk mendapatkan estimasi parameter model MGWR dengan regresi *ridge*, langkah pertama yaitu menuliskan model MGWR dalam bentuk GWR:

$$\widetilde{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}_g\boldsymbol{\beta}_g = \boldsymbol{X}_l\boldsymbol{\beta}_l(u_i, v_i) + \boldsymbol{\varepsilon} \tag{17}$$

$$\boldsymbol{\varepsilon} = \widetilde{\boldsymbol{Y}} - \boldsymbol{X}_l\boldsymbol{\beta}_l(u_i, v_i) \tag{18}$$

Mengestimasi model MGWR untuk variabel lokal dilakukan dengan menggunakan metode WLS yaitu dengan mengalikan $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$ dengan matriks pembobot $\boldsymbol{W}(u_i, v_i)$.

$$\begin{aligned}\boldsymbol{\varepsilon}'\boldsymbol{W}(u_i, v_i)\boldsymbol{\varepsilon} = \widetilde{\boldsymbol{Y}}'\boldsymbol{W}(u_i, v_i)\widetilde{\boldsymbol{Y}} - \boldsymbol{2}\boldsymbol{\beta}_l'(u_i, v_i)\boldsymbol{X}_l'\boldsymbol{W}(u_i, v_i)\widetilde{\boldsymbol{Y}} \\ +\boldsymbol{\beta}_l'(u_i, v_i)\boldsymbol{X}_l'\boldsymbol{W}(u_i, v_i)\boldsymbol{X}_l\boldsymbol{\beta}_l(u_i, v_i)\end{aligned} \tag{19}$$

Sebelum diturunkan terhadap $\boldsymbol{\beta}_l{}'(ui, vi)$, persamaan 19 ditambahkan fungsi kendala $\lambda(u_i, v_i)(\boldsymbol{\beta}_l^2(u_i, v_i) - c)$ yang berguna untuk menurunkan nilai VIF pada data yang mengakibatkan multikolinearitas. Sehingga estimasi parameter model lokal adalah:

$$\widehat{\boldsymbol{\beta}}_l(u_i, v_i) = [\boldsymbol{X}_l{}'\boldsymbol{W}(u_i, v_i)\boldsymbol{X}_l + \lambda(u_i, v_i)\boldsymbol{I}]^{-1}\boldsymbol{X}_l'\boldsymbol{W}(u_i, v_i)\widetilde{\boldsymbol{Y}} \tag{20}$$

Misalkan $x_{li}' = \left(1, x_{1i}, x_{2i}, \dots, x_{qi}\right)$ adalah elemen baris ke-i dari matriks $\boldsymbol{X}_i$, maka nilai prediksi untuk $\tilde{\boldsymbol{Y}}$ pada (u_i, v_i) dapat diperoleh dengan persamaan:

$$\hat{\widetilde{\boldsymbol{Y}}} = x_{li}'\widehat{\boldsymbol{\beta}}_l(u_i, v_i) = x_{li}'[\boldsymbol{X}_l{}'\boldsymbol{W}(u_i, v_i)\boldsymbol{X}_l + \lambda(u_i, v_i)\boldsymbol{I}]^{-1}\boldsymbol{X}_l'\boldsymbol{W}(u_i, v_i)\widetilde{\boldsymbol{Y}} \tag{21}$$

Sehingga untuk seluruh pengamatan dapat ditulis:

$$\widehat{\widetilde{\boldsymbol{Y}}} = \left(\widehat{\widetilde{\boldsymbol{Y}}}_1, \widehat{\widetilde{\boldsymbol{Y}}}_2, \ldots \widehat{\widetilde{\boldsymbol{Y}}}_n\right)' = \boldsymbol{S}_l \widetilde{\boldsymbol{Y}} \tag{22}$$

dengan

$$S_l = \begin{bmatrix} [x_{i1}' \boldsymbol{W}(u_1, v_1) \boldsymbol{X}_l + \lambda(u_1, v_1) I]^{-1} \boldsymbol{X}_l' \boldsymbol{W}(u_1, v_1) \\ [x_{i2}' \boldsymbol{W}(u_2, v_2) \boldsymbol{X}_l + \lambda(u_2, v_2) I]^{-1} \boldsymbol{X}_l' \boldsymbol{W}(u_2, v_2) \\ \vdots \\ [x_{in}' \boldsymbol{W}(u_n, v_n) \boldsymbol{X}_l + \lambda(u_n, v_n) I]^{-1} \boldsymbol{X}_l' \boldsymbol{W}(u_n, v_n) \end{bmatrix}$$

Selanjutnya mensubstitusikan elemen dari $\widehat{\boldsymbol{\beta}}_l(u_i, v_i)$ ke dalam model MGWR dengan regresi *Ridge* sehingga persamaan 9 menjadi:

$$(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{Y} = (\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g \boldsymbol{\beta}_g + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} = (\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{Y} - (\boldsymbol{I} - \boldsymbol{S})\boldsymbol{X}_g \boldsymbol{\beta}_g$$

Dengan menggunakan metode OLS dan penambahan fungsi kendala $\lambda(\beta_g' \beta_g - c)$ yang berguna untuk mengatasi multikolinearitas pada model global, maka dapat diperoleh estimasi untuk koefisien global sebagai berikut:

$$\widehat{\boldsymbol{\beta}}_g = \left[\boldsymbol{X}_g'(\boldsymbol{I} - \boldsymbol{S}_l)'(\boldsymbol{I} - \boldsymbol{S})\boldsymbol{X}_g + \lambda \boldsymbol{I}\right]^{-1} \boldsymbol{X}_g'(\boldsymbol{I} - \boldsymbol{S})'(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{Y} \tag{23}$$

Dengan mensubstitusikan nilai $\widetilde{\boldsymbol{Y}}$ pada persamaan 17 ke persamaan 20 maka diperoleh estimasi untuk koefisien lokal yaitu:

$$\widehat{\boldsymbol{\beta}}_l(u_i, v_i) = [\boldsymbol{X}_l' \boldsymbol{W}(u_i, v_i) \boldsymbol{X}_l + \lambda(u_i, v_i) \boldsymbol{I}]^{-1} \boldsymbol{X}_l \boldsymbol{W}(u_i, v_i)\left(\boldsymbol{Y} - \boldsymbol{X}_g \widehat{\boldsymbol{\beta}}_g\right) \tag{24}$$

3.2 Pengujian Multikolinearitas

Berdasarkan nilai VIF global diketahui variabel jumlah fasilitas kesehatan (x_8), rata-rata lama sekolah (x_9) dan variabel kepadatan penduduk (x_{10}) memiliki nilai VIF lebih besar dari 10 yaitu masing-masing sebesar 11.226956, 16.550084, dan 15.483114. Sehingga disimpulkan terdapat multikolinieritas pada model global.

Setelah diketahui variabel yang mengalami multikolinearitas secara global, berikutnya adalah uji multikolinearitas secara lokal. Sebagai contoh, uji multikolinearitas lokal untuk wilayah Kab. Barru. Variabel rata-rata lama sekolah (x_9) dan variabel kepadatan penduduk (x_{10}) memiliki nilai VIF lebih besar dari 10 yaitu masing-masing sebesar 11.257458 dan 14.102280. Sehingga disimpulkan terdapat multikolinieritas pada model lokal.

3.3 Pengujian Pengaruh Spasial

Pengujian pengaruh spasial dilakukan untuk mengetahui keragaman antar lokasi. Diperoleh nilai BP yaitu $19.134 > 18.307$ (χ^2) sehingga kriteria keputusan adalah tolak H_0 artinya terdapat heterogenitas spasial pada data kemiskinan.

3.4 Model GWR

Langkah pertama dalam memodelkan GWR adalah mencari *bandwidth* optimum untuk tiap lokasi dengan metode CV. Nilai bandwith optimum yang didapat dengan fungsi pembobot *adaptive gaussian kernel* dapat dilihat pada Tabel 1 berikut:

Tabel 1. Bandwith untuk setiap lokasi

Lokasi	*Bandwith*	Lokasi	*Bandwith*	Lokasi	*Bandwith*
1	2.31406	9	0.59154	17	1.80656
2	0.91096	10	0.60859	18	1.13052
3	0.70506	11	0.84949	19	2.08973
4	0.77075	12	0.54323	20	1.93342
5	0.73361	13	0.84604	21	1.50227
6	0.69186	14	0.57822	22	1.16536
7	1.44820	15	0.81976	23	0.62055
8	0.57010	16	0.99491	24	1.09637

Setelah mendapatkan nilai *bandwith* optimum, langkah selanjutnya adalah menentukan matriks pembobot pada masing-masing lokasi dengan terlebih dahulu menentukan jarak *euclidian* pada setiap lokasi pengamatan.

Tabel 2. Matriks Pembobot $\boldsymbol{w}_{ij}$

Wilayah	Selayar	Bulukumba	Bantaeng	...	Palopo
Selayar	1	0.0948	0.0088	...	0.0007
Bulukumba	0.6941	1	0.6522	..	0.1311
Bantaeng	0.6447	0.7741	1		0.1002
⋮	⋮	⋮	⋮	⋱	⋮
Palopo	0.1984	0.0527	0.0038	...	1

Sebagai contoh, berikut estimasi parameter model GWR untuk wilayah Kab. Barru

Tabel 3. Estimasi Parameter Model GWR Kab. Barru

Variabel	$\boldsymbol{\beta}$	Thit	Keterangan
intersep	120.258		
x_1	0.218	0.196	Tidak Signifikan
x_2	0.160	0.061	Tidak Signifikan
x_3	0.118	0.024	Tidak Signifikan
x_4	-0.266	0.139	Tidak Signifikan
x_5	-0.229	0.051	Tidak Signifikan
x_6	-1.241	2.422	Signifikan

x_7	-0.078	0.107	Tidak Signifikan
x_8	-0.004	$4.952e^{-5}$	Tidak Signifikan
x_9	$0.17e^{-3}$	$1.68e^{-7}$	Tidak Signifikan
x_{10}	-0.955	4.690	Signifikan

Berdasarkan Tabel 3, dengan menggunakan α = 5%, niliai T tabel adalah 2,101 diperoleh variabel signifikan untuk wilayah Kab. Barru adalah angka harapan hidup (x_6) dan Kepadatan penduduk (x_{10}).

3.5 Uji Variabilitas Model GWR

Uji variabilitas dilakukan untuk mengetahui variabel lokal dan variabel global. Berdasarkan Tabel 4, terdapat 2 variabel yang berpengaruh secara global yaitu x_3 dan x_6. Sedangkan 8 variabel yang berpengaruh secara lokal adalah x_1, x_2, x_4, x_5, x_7, x_8, x_9, dan x_{10}.

Tabel 4. Uji Variabilitas Model GWR

Variabel	$F\ hitung$	$F\ tabel$
x_1	7.382	2.340
x_2	3.490	2.340
x_3	1.348	2.377
x_4	3.101	2.340
x_5	2.483	2.422
x_6	1.964	2.340
x_7	4.539	2.307
x_8	3.355	2.543
x_9	13.143	2.543
x_{10}	11.470	2.255

3.6 Standarisasi Data

Standarisasi ini dilakukan untuk membakukan variabel sehingga menghasilkan variabel baru. Sebagai contoh, berikut hasil standarisasi data untuk wilayah Kab. Barru:

Tabel 5. Standarisasi Data Kab. Barru

Keterangan	**y**	**x1**	**x2**	**x3**	**...**	**x10**
Data Awal	9.45	4.37	24.17	41.01	...	7.61
Hasil Standarisasi	-0.24	0.51	-0.04	1.05	...	0.01

3.7 Penentuan Tetapan Bias

Dalam model MGWR dengan regresi *ridge*, penentuan tetapan bias nantinya akan ada dua yaitu tetapan bias global dan tetapan bias lokal. Pertama adalah menentukan tetapan bias global, diperoleh tetapan bias global sebesar 0.2336796. Kedua adalah penentuan tetapan bias untuk setiap lokasi.

Tabel 6. Nilai Tetapan Bias Lokal

Lokasi	λ	Lokasi	λ	Lokasi	λ
1	0.21914	9	0.18067	17	0.20861
2	0.17051	10	0.19939	18	0.18443
3	0.08408	11	0.23242	19	0.20358
4	0.11219	12	0.17950	20	0.20794
5	0.10199	13	0.21959	21	0.20068
6	0.10490	14	0.14593	22	0.16947
7	0.22740	15	0.15999	23	0.19772
8	0.14972	16	0.17200	24	0.17493

3.8 Model MGWR dengan Regresi *Ridge*

Hasil penghitungan untuk estimasi parameter global $\hat{\beta}_g$ diperoleh $x_3 = 0.6062921$ dan $x_3 = -0.5939490$. Sementara untuk estimasi parameter lokal, berikut ini contoh untuk wilayah Kab. Barru:

Tabel 7. Estimasi Parameter Variabel Lokal Kab. Barru

Wilayah	Variabel	$\hat{\beta}_l$
Kab. Barru	intersep	-0.145571
	x_1	0.132471
	x_2	0.392285
	x_4	-0.683726
	x_5	-0.650583
	x_7	-0.114886
	x_8	-0.090477
	x_9	-0.077568
	x_{10}	-0.382985

3.9 Uji Kesesuaian Model

Pengujian ini dilakukan untuk melihat apakah terdapat perbedaan yang signifikan antara model regresi linear dengan model MGWR. Diperoleh nilai statistik uji F sebesar

1.01408 yang lebih kecil dari nilai F tabel sebesar 1.80660 dengan α = 5% maka terima H_0 artinya model MGWR tidak berbeda secara signifikan dengan model regresi global.

3.10 Uji Serentak

Pada pengujian ini akan ada dua, yaitu pengujian hipotesis pada parameter variabel prediktor global dan lokal. Pertama adalah pengujian hipotesis serentak pada parameter variabel global. Diperoleh nilai statistik uji F sebesar 5.88728 dengan menggunakan $\alpha =$ 5%, nilai F tabel adalah 2.89510 maka menolak H_0 sehingga dimpulkan bahwa variabel prediktor global secara serentak berpengaruh terhadap model. Kedua adalah pengujian hipotesis serentak pada variabel prediktor lokal. Diperoleh nilai statistik uji F sebesar 2.82994 dengan menggunakan $\alpha = 1\%$ nilai F tabel adalah 2.12326 maka menolak H_0 sehingga dapat disimpulkan bahwa variabel prediktor lokal secara serentak berpengaruh terhadap model.

3.11 Uji Parsial

Uji ini digunakan untuk mengetahui variabel global dan lokal yang berpengaruh signifikan terhadap respon pada model MGWR. Diperoleh variabel prediktor global yang berpengaruh signifikan adalah rumah tangga memiliki dinding terluas bukan tembok dan kayu (x_3) dan angka harapan hidup (x_6).

Tabel 8. Uji Parsial Parameter Global Model MGWR

Variabel	T Hitung	T Tabel	Keterangan
x_3	3.5321	2.093024	Signifikan
x_6	-4.63152		Signifikan

Selanjutnya adalah menguji signifikansi parameter variabel lokal. Berikut contohnya pengujian parameter secara parsial pada Kab. Barru:

Tabel 9. Uji Parsial Parameter Lokal Model MGWR dengan regresi ridge

Variabel	T hitung	T Tabel	Keterangan
x_1	0.98575	2.43344	Tidak Signifikan
x_2	4.61461		Signifikan
x_4	-9.96984		Signifikan
x_5	-7.07199		Signifikan
x_7	-1.09842		Tidak Signifikan
x_8	-1.56298		Tidak Signifikan
x_9	-1.74963		Tidak Signifikan
x_{10}	-11.0399		Signifikan

3.12 Interpretasi Model dan Pemilihan Model Terbaik

Model MGWR dengan regresi *ridge* yang dihasilkan akan berbeda-beda pada setiap lokasi pengamatan bergantung pada nilai parameter model dan variabel yang signifikan. Misalkan pada lokasi pengamatan Kab. Barru, model MGWR dengan regresi *ridge* yang dihasilkan setelah diubah ke bentuk awal:

$$y = 72.795 + 0.177x_2 + 0.046x_3 - 0.058x_4 - 0.161x_5 - 0.578x_6 - 0.217x_{10}$$

artinya bahwa jika nilai x_2 dinaikkan satu satuan sedangkan variabel lainnya tetap maka taksiran y akan meningkat sebesar 0.177, begitupun untuk x_3, x_4, x_5, x_6, dan x_{10}. Membandingkan model dilakukan untuk mengetahui apakah setelah penambahan regresi *ridge*, model yang diterapkan untuk menggambarkan presentase penduduk miskin di Kabupaten/Kota Sulawesi Selatan tahun 2016 bisa lebih baik.

Tabel 10. Perbandingan AIC model MGWR dan MGWR dengan regresi ridge

Model	**AIC**
MGWR	64.56255
MGWR *ridge*	63.64473

Berdasarkan Tabel 10 diperoleh model MGWR dengan regresi *ridge* mempunyai nilai AIC terkecil yaitu 63.64473. Sehingga dapat disimpulkan bahwa model MGWR dengan regresi *ridge* lebih baik digunakan untuk memodelkan tingkat penduduk miskin di Kabupaten/Kota Sulawesi Selatan tahun 2016 yang mengalami multikolinearitas.

4. Kesimpulan

Berdasarkan hasil analisis data penelitian yang dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Estimasi parameter model *Mixed Geographically Weighted Regression* (MGWR) dengan regresi *ridge* adalah:

$$\widehat{\boldsymbol{\beta}}_g = \left[\boldsymbol{X}_g'(\boldsymbol{I}-\boldsymbol{S}_l)'(\boldsymbol{I}-\boldsymbol{S}_l)\boldsymbol{X}_g + \lambda\boldsymbol{I}\right]^{-1}\boldsymbol{X}_g'(\boldsymbol{I}-\boldsymbol{S}_l)'(\boldsymbol{I}-\boldsymbol{S}_l)\boldsymbol{Y}$$

$$\widehat{\boldsymbol{\beta}}_l(u_i,v_i) = [\boldsymbol{X}_l'\boldsymbol{W}(u_i,v_i)\boldsymbol{X}_l + \lambda(u_i,v_i)\boldsymbol{I}]^{-1}\boldsymbol{X}_l'\boldsymbol{W}(u_i,v_i)(\boldsymbol{Y}-\boldsymbol{X}_g\widehat{\boldsymbol{\beta}}_g)$$

2. Berdasarkan model MGWR dengan regresi *ridge*, faktor-faktor yang berpengaruh signifikan terhadap persentase penduduk miskin di Kabupaten/Kota Sulawesi Selatan tahun 2016 secara lokal adalah rumah tangga menurut luas lantai rumah < 8 m^2 (x_2), Penduduk usia 15 tahun ke atas < SD (x_4), penduduk usia 15 tahun ke atas yang tidak bekerja (x_5), pengeluaran perkapita untuk makanan (x_7), jumlah fasilitas kesehatan (x_8), rata-rata lama sekolah (x_9) dan kepadatan penduduk (x_{10}). Sedangkan variabel yang berpengaruh signifikan secara global adalah rumah tangga memiliki dinding terluas bukan tembok dan kayu (x_3) dan angka harapan hidup (x_6).

Daftar Pustaka

[1] Mahmuda dan Harini, S. *Statistik Uji Parsial Pada Model Mixed Geographically Weighted Regression*. Malang: UIN Maulana Malik Ibrahim. 2014.

[2] Yulita, T. *Pemodelan Geographically Weighted Ridge Regression dan Geographically Weighted Lasso pada Data Spasial dengan Multikolinearitas*. Institut Pertanian Bogor. 2016.

[3] Ifadah, A. *Analisis Metode Principal Component Analysis (Komponen Utama) dan Regresi Ridge dalam Mengatasi Dampak Multikolinearitas dalam Analisis Regresi Linear Berganda.* Semarang: Fakultas MIPA UNS. 2011.

[4] Anselin, L. *Spatial Econometrics*. Dallas: School of Social Science, 2009.

[5] Fotheringham, A.S., Brundson, C. dan Charlthon, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Ltd. UK. 2002.

[6] Wuryanti, I.F., Purnami, S.W., dan Purhadi. *Pemodelan Mixed Geographically Weighted Regression (MGWR) pada Angka Kematian Balita Kabupaten Bojonegoro Tahun 2011*. Surabaya: Fakultas MIPA ITS. 2013.

[7] Purhadi dan Yasin, H. Mixed Geographically Weighted Regression Model Case Study: The Percentage Of Poor Households In Mojokerto 2008. *European Journal of Scientific Research*. 2012.