

Model *Robust Geographically Weighted Regression* pada Data Kemiskinan di Sulawesi Selatan Tahun 2019

Aqilah Salsabila Rahman^{1*}, Georgina Maria Tinungki², Erna Tri Herdiani³

¹²³Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Hadanuddin, Makassar, 60294, Indonesia

* Corresponding author, email: aqilahsr123@gmail.com

Abstract

Geographically Weighted Regression (GWR) is a method of spatial analysis that can be used to perform analysis by assigning weights based on the geographical distance of each observation location and the assumption of having spatial heterogeneity. The result of this analysis is an equation model whose parameter values apply only to each observation location and are different from other observation locations. However, when there are outliers at the observation location, a more robust estimation method is needed. One of the robust methods that can be applied to the GWR model is the Least Absolute Deviation method. In this study, model estimation was carried out on the factors that affect poverty in South Sulawesi in 2019 using Robust Geographically Weighted Regression (RGWR) with the Least Absolute Deviation (LAD) method. Determination of weighting is done by using the adaptive kernel bisquare weighting function. The results obtained are RGWR models which are different and apply only to each district/city in South Sulawesi. In addition, it was also found that the RGWR model with the LAD method was the best model for data that experienced spatial heterogeneity and contained outliers.

Keywords: *Adaptive Kernel Bisquare, Geographically Weighted Regression, Least Absolute Deviation, Poverty, Robust.*

Abstrak

*Geographically Weighted Regression (GWR) adalah salah satu metode analisis spasial yang dapat digunakan untuk melakukan analisis dengan pemberian pembobot berdasarkan jarak setiap lokasi pengamatan secara geografis serta asumsi memiliki keragaman spasial. Hasil dari analisis ini adalah model persamaan yang nilai-nilai parameternya berlaku hanya pada masing-masing lokasi pengamatan dan berbeda dengan lokasi pengamatan lainnya. Namun, saat terdapat pencilan pada lokasi pengamatan, diperlukan sebuah metode estimasi yang lebih kekar (robust). Salah satu metode robust yang dapat diterapkan pada model GWR adalah dengan metode Least Absolute Deviation. Pada penelitian ini estimasi model dilakukan terhadap faktor-faktor yang mempengaruhi kemiskinan di Sulawesi Selatan Tahun 2019 menggunakan model *Robust Geographically Weighted Regression* (RGWR) dengan metode *Least Absolute Deviation* (LAD). Penentuan pembobot dilakukan dengan menggunakan fungsi pembobot adaptive kernel bisquare. Hasil yang didapatkan adalah model RGWR yang berbeda-beda dan berlaku hanya pada masing-masing kabupaten/kota di Sulawesi Selatan. Selain itu, didapatkan pula hasil bahwa model RGWR dengan metode LAD adalah model yang paling baik pada data yang mengalami keragaman secara spasial dan mengandung pencilan.*

Kata Kunci: *Adaptive Kernel Bisquare, Adaptive Kernel Bisquare, Kemiskinan, Least Absolute Deviation, Robust.*

1. Pendahuluan

Kemiskinan merupakan salah satu indikator kesejahteraan rakyat. Menurut Badan Pusat Statistik (BPS), kemiskinan merupakan permasalahan yang hampir dialami oleh semua negara di dunia termasuk negara yang sudah maju sekalipun. Menurut Bank Dunia, pada tahun 2015 sepuluh persen dari populasi penduduk dunia, yaitu sebanyak 731 juta jiwa hidup di bawah garis kemiskinan internasional. Hal tersebut mengindikasikan bahwa kemiskinan masih menjadi isu global yang perlu diselesaikan. Berdasarkan data BPS, persentase jumlah penduduk miskin di Indonesia pada Maret 2020 sebesar 9.78 persen, meningkat 0.56 persen dari September 2019 dan meningkat 0.37 persen dari Maret 2019. Lebih khususnya lagi, persentase penduduk miskin di Sulawesi Selatan pada Maret 2020 sebesar 8.72 persen, meningkat sebesar 0.16 persen dari September 2019 dan meningkat sebesar 0.03 persen dari Maret 2019. Peningkatan jumlah penduduk miskin tersebut mengindikasikan bahwa pengambilan kebijakan masih perlu dikaji mengenai faktor-faktor penyebabnya.

Salah satu metode yang dapat digunakan untuk menganalisis faktor-faktor yang menyebabkan kemiskinan adalah regresi linier berganda. Regresi linier berganda merupakan metode yang memodelkan hubungan antara variabel respon dan variabel prediktor [1]. Metode estimasi yang digunakan untuk menganalisis regresi adalah metode kuadrat terkecil [2]. Dalam menggunakan suatu metode, terdapat asumsi yang harus dipenuhi agar metode tersebut dapat digunakan. Metode kuadrat terkecil mempunyai asumsi-asumsi tertentu [3]. Salah satu asumsi yang harus dipenuhi ialah kehomogenan ragam. Kehomogenan ragam sulit didapat karena adanya perbedaan karakteristik pada suatu wilayah yang mengakibatkan terjadinya keragaman spasial [4]. Keragaman spasial merupakan suatu keadaan saat pengukuran hubungan antar variabel berbeda-beda antara satu lokasi pengamatan dengan lokasi pengamatan yang lainnya [5]. Sehingga, diperlukan pendekatan analisis yang memperhatikan keadaan geografis yang dalam hal ini disebut analisis spasial.

Geographically Weighted Regression (GWR) adalah salah satu metode analisis spasial yang dapat digunakan untuk melakukan analisis dengan pemberian pembobot berdasarkan jarak masing-masing lokasi pengamatan secara geografis serta asumsi memiliki keragaman spasial. Hasil dari analisis ini adalah model persamaan yang nilai-nilai parameternya berlaku hanya pada masing-masing lokasi pengamatan dan berbeda dengan lokasi pengamatan lainnya.

Seluruh penelitian yang baru saja disebutkan menggunakan metode estimasi parameter kuadrat terkecil terboboti. Menurut Metode tersebut tidak kekar terhadap keberadaan pencilan. Sehingga, apabila terdapat pencilan pada data, maka akan mengakibatkan terciptanya estimasi parameter yang bias dan mengakibatkan kekeliruan dalam menyimpulkan hubungan regresi [6]. Alternatif yang dapat dilakukan pada analisis regresi untuk mengatasi adanya pencilan pada data, yaitu menggunakan metode regresi kekar [7]. Oleh karena itu, metode analisis *Robust Geographically Weighted*

Regression (RGWR) sesuai untuk menganalisis data yang mencakup beberapa lokasi pengamatan dengan asumsi memiliki keragaman spasial serta mengandung pencilan.

Salah satu metode estimasi parameter diperkenalkan oleh Roger Joseph Boscovich pada tahun 1957, yaitu metode Least Absolute Deviation. Metode ini digunakan dengan meminimumkan jumlah mutlak galat untuk memperoleh hasil estimasi parameter regresi. Salah satu penelitian telah membandingkan metode LAD dan estimasi M, sehingga diperoleh hasil kesimpulan bahwa metode LAD merupakan metode yang paling baik [8].

Berdasarkan uraian di atas, maka peneliti tertarik untuk memodelkan faktor-faktor yang mempengaruhi kemiskinan di Sulawesi Selatan Tahun 2019 menggunakan Robust Geographically Weighted Regression dengan metode Least Absolute Deviation.

2. Material dan Metode

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) melalui Survei Sosial Ekonomi Nasional (SUSENAS), dan Survei Angkatan Kerja Nasional (SAKERNAS). Cakupan data penelitian adalah 24 kabupaten/kota di Provinsi Sulawesi Selatan tahun 2019. Variabel yang digunakan dalam penelitian ini yaitu Persentase Penduduk Miskin (Y), Indeks Pembangunan Manusia (X_1), Tingkat Pengangguran Terbuka (X_2), Angka Melek Huruf Penduduk Miskin (X_3), Persentase Penduduk Miskin yang Belum Tamat SD (X_4), Persentase Penduduk Miskin Bekerja di Sektor Informal (X_5), Persentase Pengeluaran Perkapita Penduduk Miskin untuk Makanan (X_6), Persentase Rumah Tangga Miskin yang Menggunakan Air Layak (X_7), Persentase Rumah Tangga Miskin yang Menggunakan Jamban Sendiri/Bersama (X_8), Persentase Rumah Tangga Miskin Penerima Rastra atau BPNT (X_9).

Data dianalisis dengan menggunakan model RGWR dengan menggunakan metode LAD. Diberikan koordinat lintang dan bujur (u_i, v_i) pada lokasi pengamatan ke- i , maka model GWR pada lokasi pengamatan tersebut dinyatakan sebagai berikut [9].

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

Apabila d_{ij} adalah jarak (u_i, v_i) antar lokasi pengamatan ke- i , maka koefisien regresi pada persamaan (1) diestimasi dengan metode LAD berdasarkan lokasi (u_i, v_i) yang dinyatakan sebagai:

$$\min \sum_{i=1}^n W_{ij} |\varepsilon_i| \quad (2)$$

dengan fungsi pembobot *adaptive kernel bisquare* W_{ij} sebagai berikut:

$$W_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h_i}\right)^2\right)^2 & ; d_{ij} \leq h_i \\ 0 & ; d_{ij} > h_i \end{cases} \quad (3)$$

Apabila dilakukan dekomposisi pada galat ε_i , dan dimisalkan $\varepsilon_i = y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}}_i$ [10], maka estimasi parameter model RGWR dengan metode *Least Absolute Deviation* pada lokasi pengamatan ke- i dapat dinyatakan seperti pada Persamaan (4) berikut:

$$\hat{\boldsymbol{\beta}}_{i,LAD} = \arg \min \left\{ \sum_{i=1; \varepsilon_i \geq 0}^n W_{ij} (y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}}_i) + \sum_{i=1; \varepsilon_i < 0}^n W_{ij} (y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}}_i) \right\} \quad (4)$$

Persamaan (4) selanjutnya diselesaikan dengan menggunakan metode simpleks dengan masalah optimasi linier sebagai berikut:

1. Meminimumkan $\sum_{i=1}^n W_{ij} |\varepsilon_i| = \sum_{i=1, \varepsilon_i \geq 0}^n W_{ij} (\varepsilon_i^+) + \sum_{i=1, \varepsilon_i < 0}^n W_{ij} (\varepsilon_i^-)$
2. Fungsi kendala $\beta_{i0} + \sum_{k=1}^p \beta_{ik} x_{ik} + \varepsilon_i^+ - \varepsilon_i^- = y_i, \varepsilon_i^+, \varepsilon_i^- \geq 0; i = 1, 2, \dots, n.$

Dari estimasi parameter model RGWR dengan menggunakan metode LAD yang diperoleh pada Persamaan (3) selanjutnya diterapkan pada data kemiskinan di Sulawesi Selatan tahun 2019 dengan estimasi model RGWR sebagai berikut:

$$\begin{aligned} \hat{y}_i &= \hat{\beta}_{i,0,LAD} + \hat{\beta}_{i,1,LAD} x_{i1} + \hat{\beta}_{i,2,LAD} x_{i2} + \dots + \hat{\beta}_{i,n,LAD} x_{in} \\ &= \mathbf{x}_i \hat{\boldsymbol{\beta}}_{i,LAD} \end{aligned}$$

Langkah awal pada penelitian ini adalah menentukan model regresi linier berganda menggunakan metode *Ordinary Least Squares* (OLS). Selanjutnya, melakukan pengujian pengaruh spasial yang dalam hal ini pengaruh keragaman spasial (*spatial heterogeneity*) dengan menggunakan uji *Breusch-Pagan*. Selanjutnya, menentukan model GWR menggunakan metode *Weighted Least Squares* (WLS). Kemudian, melakukan pendeteksian pencilan. Setelah itu, menentukan model RGWR dengan menggunakan metode LAD).

3. Hasil dan Diskusi

3.1 Pemodelan Regresi Linier Berganda dengan Metode *Ordinary Least Squares*

Estimasi parameter model regresi linier berganda diperoleh dengan menggunakan metode OLS, sehingga diperoleh $\hat{\boldsymbol{\beta}}_{OLS}$ yang berlaku pada seluruh kabupaten/kota di Sulawesi Selatan. Selanjutnya, dari $\hat{\boldsymbol{\beta}}_{OLS}$ didapatkan estimasi model regresi linier berganda sebagai berikut:

$$\begin{aligned} y_i &= 26.236 - 0.377x_{i1} + 0.259x_{i2} - 0.008x_{i3} - 0.158x_{i4} + 0.048x_{i5} + 0.209x_{i6} \\ &\quad - 0.016x_{i7} - 0.042x_{i8} + 0.099x_{i9} \end{aligned}$$

Salah satu asumsi yang harus dipenuhi saat menggunakan metode OLS ialah asumsi kehomogenan ragam. Sehingga, pengujian keragaman perlu dilakukan untuk mengetahui adanya keragaman pada data.

3.2 Keragaman Spasial

Keragaman spasial ditunjukkan oleh perbedaan karakteristik antara satu lokasi pengamatan dengan lokasi pengamatan lainnya. Pengujian keragaman spasial ini dapat dilakukan dengan menggunakan statistik uji Breusch-Pagan (BP) dengan prosedur pengujian hipotesis sebagai berikut:

a. Hipotesis:

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{24}^2 = \sigma^2$ (Tidak terdapat keragaman spasial)

$H_1 : \text{Terdapat } \sigma_i^2 \neq \sigma^2$ (Terdapat keragaman spasial)

b. Statistik uji:

$$BP = \left(\frac{1}{2}\right) \mathbf{f}^T \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} = 17.9327$$

c. Kesimpulan:

Karena $BP = 17.9327 > \chi_{(0,05,9)}^2 = 16.9190$, maka tolak H_0 dan terima H_1 .

Dapat dikatakan bahwa terdapat keragaman spasial atau keragaman antar lokasi setiap variabel pada data kemiskinan. Oleh karena itu, pemodelan regresi dapat dilanjutkan dengan memperhatikan lokasi yang dalam hal ini digunakan model GWR.

3.3 Pemodelan GWR dengan Metode WLS

Estimasi parameter model GWR diperoleh menggunakan metode kuadrat terkecil terboboti (WLS) dengan fungsi pembobot *adaptive kernel bisquare*, sehingga diperoleh $\hat{\beta}_{i,WLS}$ yang berlaku pada setiap 24 kabupaten/kota di Sulawesi Selatan pada tabel 1 berikut:

Tabel 1. Estimasi Parameter Model GWR

Variabel	Maksimum	Minimum
$\hat{\beta}_{i,0,WLS}$	39.602	28.590
$\hat{\beta}_{i,1,WLS}$	-0.358	-0.511
$\hat{\beta}_{i,2,WLS}$	0.541	-0.034
$\hat{\beta}_{i,3,WLS}$	-0.051	-0.114
$\hat{\beta}_{i,4,WLS}$	-0.171	-0.276
$\hat{\beta}_{i,5,WLS}$	0.056	0.042
$\hat{\beta}_{i,6,WLS}$	0.321	0.143
$\hat{\beta}_{i,7,WLS}$	0.038	-0.018
$\hat{\beta}_{i,8,WLS}$	-0.046	-0.067
$\hat{\beta}_{i,9,WLS}$	0.113	0.082

Selanjutnya, dari estimasi parameter $\hat{\beta}_{i,WLS}$ didapatkan estimasi model GWR untuk persentase penduduk miskin di Sulawesi Selatan tahun 2019 yang berlaku pada setiap kabupaten/kota di Sulawesi Selatan.

3.4 Deteksi Pencilan

Pendeteksian pencilan pada galat model GWR dilakukan dengan metode *Boxplot* dan diperoleh bahwa terdapat pencilan pada galat model GWR, yaitu pada Kabupaten Wajo. Keberadaan pencilan ini mengindikasikan bahwa model GWR yang diperoleh tidak kekar terhadap keberadaan pencilan. Sehingga, pemodelan GWR perlu dilanjutkan menggunakan metode yang lebih kekar (*robust*) yang dalam hal ini selanjutnya disebut dengan model RGWR.

3.4.1 Pemodelan Robust Geographically Weighted Regression dengan Metode Least Absolute Deviation

Estimasi parameter model RGWR diperoleh menggunakan metode LAD) dengan fungsi pembobot *adaptive kernel bisquare*, sehingga diperoleh $\hat{\beta}_{i,LAD}$ yang berlaku pada setiap 24 kabupaten/kota di Sulawesi Selatan sebagai berikut.

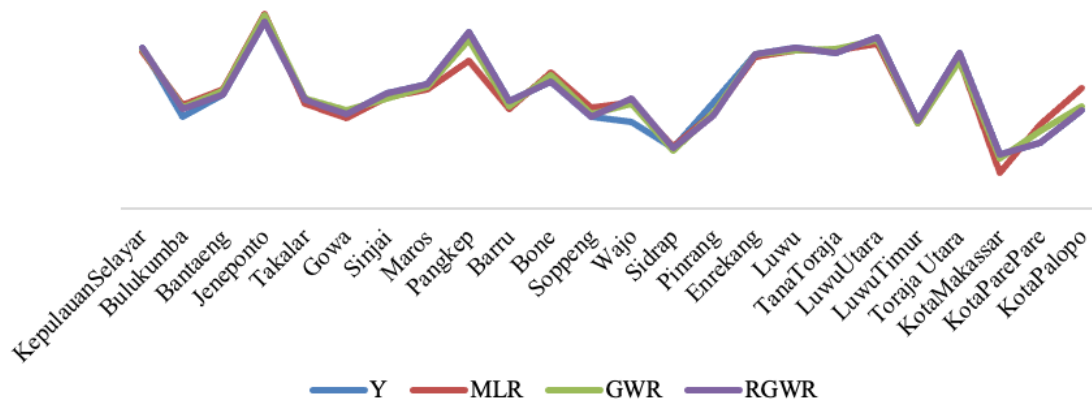
Tabel 2. Estimasi Parameter Model RGWR

Variabel	Maksimum	Minimum
$\hat{\beta}_{i,0,LAD}$	79.822	31.267
$\hat{\beta}_{i,1,LAD}$	-0.265	-0.601
$\hat{\beta}_{i,2,LAD}$	0.734	-0.475
$\hat{\beta}_{i,3,LAD}$	0.014	-0.214
$\hat{\beta}_{i,4,LAD}$	-0.143	-0.324
$\hat{\beta}_{i,5,LAD}$	0.106	-0.038
$\hat{\beta}_{i,6,LAD}$	0.331	-0.107
$\hat{\beta}_{i,7,LAD}$	0.074	-0.013
$\hat{\beta}_{i,8,LAD}$	-0.036	-0.110
$\hat{\beta}_{i,9,LAD}$	0.122	-0.004

Selanjutnya, dari estimasi parameter $\hat{\beta}_{i,LAD}$ didapatkan estimasi model RGWR untuk persentase penduduk miskin di Sulawesi Selatan tahun 2019 yang berlaku pada setiap kabupaten/kota di Sulawesi Selatan.

3.4.2 Pemilihan Model Terbaik

Perbandingan nilai estimasi persentase penduduk miskin di Sulawesi Selatan tahun 2019 untuk setiap model Regresi Lininer Berganda, GWR, dan RWGR disajikan pada Gambar 1 berikut:



Gambar 1. Perbandingan Model MLR, GWR, dan RGWR

Berdasarkan Gambar 1 di atas dapat dilihat bahwa nilai estimasi persentase penduduk miskin di Sulawesi Selatan tahun 2019 ($\hat{Y}$) pada model RGWR lebih mendekati nilai aktual persentase penduduk miskin (Y). Selanjutnya, perlu dilakukan pengukuran model terbaik, yaitu kriteria *Mean Squares of Error* (MSE) dan Koefisien Determinasi (R^2). Nilai MSE dengan derajat kebebasan galat sebesar 14 dan nilai R^2 yang diperoleh disajikan pada Tabel 3 berikut:

Tabel 3. Perbandingan Nilai MSE dan R^2

Metode	MSE	R^2 (%)
MLR	1.397	90.848
GWR	0.437	93.863
RGWR	0.367	97.517

Berdasarkan Tabel 3 dapat dilihat bahwa model regresi yang menghasilkan nilai MSE paling kecil dan nilai R^2 paling besar adalah model RGWR. Oleh karena itu, dapat dikatakan bahwa model RGWR merupakan model yang paling baik untuk digunakan dalam mengestimasi persentase penduduk miskin di Sulawesi Selatan tahun 2019 yang mengalami keragaman secara spasial dan mengandung pencilan.

4 Kesimpulan

Berdasarkan hasil yang diperoleh, dapat disimpulkan bahwa model RGWR menghasilkan estimasi model yang berbeda-beda untuk setiap kabupaten/kota di Sulawesi Selatan. Selain itu, model RGWR dengan metode LAD memberikan hasil estimasi yang paling baik pada data kemiskinan di Sulawesi Selatan tahun 2019 yang mengalami keragaman secara spasial dan mengandung pencilan dengan nilai MSE sebesar 0.367 dan R^2 sebesar 97.51%.

Daftar Pustaka

- [1] Dewi, P. L. A. *Pemodelan Faktor-Faktor Penyebab Kecelakaan Lalu Lintas Berdasarkan Metode Geographically Weighted Regression di Jawa Timur*. Skripsi, Institut Teknologi Sepuluh Nopember, 2016.
- [2] Montgomery, D. C. & Peck, E. A. *Introduction to Linear Regression Analysis, 2nd edition*. New York: John Wiley & Sons, Inc, 1992.
- [3] Gujarati, D. *Dasar-dasar Ekonometrika, Edisi ketiga, jilid 1*. Julius A. Mulyadi, S.E, penerjemah. Jakarta: Erlangga. Terjemahan dari: Essentials of Econometrics, 2007.
- [4] Erda, G. *Pendugaan Model Regresi Terboboti Geografis dan Temporal Kekar Menggunakan Penduga-M*. Tesis. Institut Pertanian Bogor, 2018.
- [5] Fotheringham, dkk. *Geographically Weighted Regression : The Analysis of Spatially Varying Relationships*. John Wiley and Sons, 2002.
- [6] Djuraidah, dkk. *Pemodelan Regresi Terboboti Kekar dengan Simpangan Mutlak Terkecil dan Penduga-M (Studi Kasus: PDRB Kabupaten/Kota Pulau Jawa Tahun 2015)*. Tesis. Institut Pertanian Bogor, 2019.
- [7] Sari, R. A. *Perbandingan Beberapa Metode Kekar pada Pendugaan Parameter Regresi Linier Sederhana untuk Data yang Mengandung Pencilan*. Skripsi. Institut Pertanian Bogor, 2016.
- [8] Wulandari, dkk. Robust Geographically Weighted Regression Modeling using Least Absolute Deviation and M-Estimator. *International Journal of Scientific Research in Science, Engineering and Technology*, 6(1), 238-245, 2019.
- [9] Huang, dkk. Geographically and Temporally Weighted Regression For Modeling Spatio-Temporal Variation In House Prices. *International Journal of Geographical Information Science*, 24(3), 383-401, 2010.
- [10] Wheeler, D. C., Fischer, M., Nijkamp, P. *Geographically Weighted Regression. Handbook of Regional Science*, Berlin, Heidelberg: Springer, 2014.