



***ROBUST GEOGRAPHICALLY WEIGHTED REGRESSION DENGAN
METODE LEAST ABSOLUTE DEVIATION PADA KASUS
PENYEBARAN COVID 19 DI INDONESIA***

JURNAL ILMIAH

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Statistika

Oleh

Dilla Permata Sari

B2A20004

PROGRAM STUDI STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS MUHAMMADIYAH SEMARANG

2021

ROBUST GEOGRAPHICALLY WEIGHTED REGRESSION DENGAN METODE LEAST ABSOLUTE DEVIATION PADA KASUS PENYEBARAN COVID 19 DI INDONESIA

Dilla Permata sari¹, Tiani Wahyu Utami², Indah Manfaati Nur³

^{1,2,3}Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

Alamat e-mail : dillapermatasari98@gmail.com

ABSTRAK

Corona Virus Disease-19 (Covid-19) adalah virus yang bisa menyerang sistem pernapasan manusia dan hewan. Covid-19 pertama kali ditemukan di Indonesia pada tanggal 2 Maret 2020 dimana dua warga Depok dinyatakan positif Covid-19. Pandemi Covid-19 telah merebak di seluruh Indonesia. Penelitian ini bertujuan untuk mengetahui model dan faktor-faktor yang mempengaruhi penyebaran kasus covid-19 di Indonesia menggunakan *Robust Geographically Weighted Regression* dengan Metode *Least Absolute Deviation*. Metode ini digunakan karena adanya heterogenitas spasial dan pencilaan sehingga pendekatan dengan analisis regresi sederhana tidak bisa dilakukan. Model yang dihasilkan bersifat lokal untuk setiap provinsi di Indonesia. Data yang digunakan adalah data penyebaran kasus positif Covid-19 di Indonesia dari tanggal 2 maret 2020 hingga 24 mei 2021. Hasil penelitian menunjukkan bahwa rata-rata kasus positif Covid-19 di Indonesia adalah 52.199,26 dengan kasus positif tertinggi dan terendah terjadi Provinsi DKI Jakarta dan Maluku Utara. Berdasarkan nilai MAPE terkecil model RGWR mampu memberikan hasil estimasi yang lebih baik dibandingkan dengan model GWR, sehingga estimasi kasus positif Covid-19 dengan RGWR lebih mendekati keadaan sebenarnya pada 34 Provinsi di Indonesia.

Kata Kunci: Covid-19, *Least Absolute Deviation*, *Robust Geographically Weighted Regression*,

ABSTRACT

Corona Virus Disease-19 (Covid-19) is a virus that can attack the respiratory system of humans and animals. Covid-19 was first discovered in Indonesia on March 2, 2020 where two Depok residents tested positive for Covid-19. The Covid-19 pandemic has spread throughout Indonesia. This study aims to determine the models and factors that influence the spread of COVID-19 cases in Indonesia using *Robust Geographically Weighted Regression* with the *Least Absolute Deviation Method*. The method used is due to spatial and pencil heterogeneity so that an approach with simple regression analysis cannot be carried out. The resulting model is local for each province in Indonesia. The data used is data on the spread of positive Covid-19 cases in Indonesia from March 2, 2020 to May 24, 2021. The results show that the average positive case of Covid-19 in Indonesia is 52,199.26 with the highest and lowest positive cases occurring in DKI Jakarta and North Maluku Provinces. Based on the MAPE value compared to the RGWR model, it is able to provide better estimation results with the

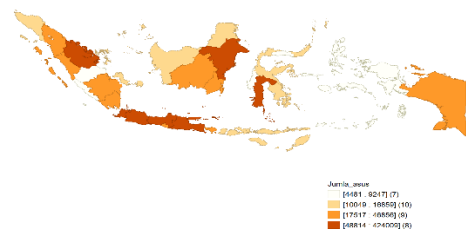
GWR model, so that the estimation of positive cases of Covid-19 with the RGWR is closer to the actual situation in 34 provinces in Indonesia.

Keyword: Covid-19, Least Absolute Deviation, Robust Geographically Weighted Regression,

PENDAHULUAN

Corona Virus Disease-19 (Covid-19) adalah virus yang bisa menyerang sistem pernapasan manusia dan hewan. Pada manusia biasanya menyebabkan penyakit infeksi saluran pernapasan mulai dari flu biasa hingga penyakit yang serius seperti *Middle East Respiratory Syndrome* (MERS) dan *Sindrom Pernafasan Akut Berat/Severe Acute Respiratory Syndrome* (SARS). Kasus Covid-19 pertama kali dilaporkan di Kota Wuhan, Provinsi Hubei, Tiongkok pada Desember 2019 yang diberi nama *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-COV2).

Covid-19 pertama kali ditemukan di Indonesia pada tanggal 2 Maret 2020 dimana 2 warga Depok dinyatakan positif. Hingga pada tanggal 31 Maret 2020, presiden Indonesia menandatangani peraturan pemerintah nomor 21 Tahun 2020 yang mengatur Pembatasan Sosial Berskala Besar (PSBB) serta menandatangani keputusan presiden Nomor 11 Tahun 2020 yang menyatakan pandemi virus corona sebagai bencana nasional. Berikut adalah peta sebaran Covid-19 di Indonesia dari tanggal 2 Maret 2020 hingga 24 Mei 2021.



Gambar 1. 1 Penyebaran Kasus Covid-19 di Indonesia

Berdasarkan peta diatas terdapat beberapa tingkatan penyebaran kasus Covid-19 di Indonesia yang terbagi atas empat zona yaitu zona merah, zona orange, zona kuning dan zona putih. Zona merah menandakan penyebaran kasus Covid-19 nya sangat tinggi yang terdiri dari delapan provinsi. Zona orange menandakan penyebaran Covid-19 nya tinggi yang terdiri dari sembilan provinsi. Zona kuning menandakan penyebaran Covid-19 dikategori sedang yang terdiri dari sepuluh provinsi. Sedangkan zona putih menandakan kasus penyebaran Covid-19 dikategori rendah yang terdiri dari tujuh provinsi. Jadi, berdasarkan peta di atas dapat disimpulkan bahwa Indonesia masih berada dalam kondisi rawan Covid-19.

Menurut Hardianto (2020), kepadatan penduduk memiliki andil dalam penyebaran Covid-19 di Indonesia, hal ini merujuk pada kenyataan bahwa kawasan perkotaan yang memiliki tingkat kepadatan penduduk tinggi akan menyebabkan transmisi penyakit lebih cepat. Kementerian Kesehatan Republik Indonesia (2020) menjelaskan penyebaran Covid-19 di Indonesia dipengaruhi oleh persentase penduduk lanjut usia hal ini dikarenakan menurunnya imunitas seiring dengan bertambahnya usia. Angka kemiskinan memiliki pengaruh signifikan terhadap total kasus Covid-19 hal ini terjadi karena individu dan populasi miskin tidak memiliki akses ke layanan Kesehatan dan mungkin mendapat informasi yang salah karena kurangnya akses ke saluran informasi sehingga mereka lebih cenderung mengabaikan peringatan Kesehatan

masyarakat (Ahmed, Pissarides & Stiglitz, 2020). Selain itu, sumber air bersih sangat mempengaruhi penyebaran Covid-19 di Indonesia hal ini dikarenakan kebijakan dari WHO yang menganjurkan mencuci tangan dengan air bersih dan mandi apabila pulang dari bepergian sebagai wujud perilaku hidup bersih dan sehat.

Penyebaran Covid-19 yang cepat dari suatu lokasi menuju lokasi lainnya mengindikasikan adanya ketergantungan spasial dan ketidakstasioneran spasial (Anselin: 1988). Ketidakstasioneran spasial merujuk pada nilai varians yang tidak sama untuk setiap wilayah atau dikenal dengan heterogenitas spasial. Oleh sebab itu, model regresi dikembangkan dengan membiarkan varians dari error yang dihasilkan model berbeda untuk setiap wilayah yaitu dengan cara menjadikan koefisien regresi bersifat lokal, artinya setiap wilayah akan memiliki koefisien regresi masing-masing (Fotheringham, dkk: 2002). Salah satu pengembangan dari model regresi yang mengakomodasi parameter yang bersifat lokal adalah *Geographically Weighted Regression* (GWR) (Fotheringham, dkk: 2002). GWR dibangun dengan penaksir *Weighted Least Squares* (WLS) yaitu dengan menambahkan pembobot spasial pada model regresi. Penaksiran parameter pada GWR dengan pendekatan *Least Squares* (LS), sangat sensitif terhadap keberadaan pencilan pada data (Afifah: 2017). Keberadaan pencilan pada data akan mengakibatkan hasil estimasi parameter menjadi tidak efisien dikarenakan residual yang dihasilkan cenderung besar. Oleh sebab itu, diperlukan metode yang lebih *robust* guna mengakomodasi adanya pencilan data pada model GWR. Salah satu metode yang digunakan adalah dengan menggunakan *Least Absolute Deviation* (LAD). Pada kasus Covid-19, model RGWR dengan metode LAD akan menggambarkan hubungan antara

jumlah kasus Covid-19 dengan faktor-faktor penyebabnya untuk setiap wilayah. Hal ini akan lebih efektif karena penanganan penyebaran Covid-19 disesuaikan dengan kondisi wilayah masing-masing. Oleh sebab itu, pada penelitian ini akan dilakukan kajian tentang model RGWR dengan menggunakan metode LAD dalam penaksiran parameternya pada kasus penyebaran Covid-19 pada 34 provinsi di Indonesia.

TINJAUAN PUSTAKA

1. *Geographically Weighted Regression*

Geographically Weighted Regression (GWR) adalah salah satu solusi yang dapat digunakan jika efek spasial, yaitu heterogenitas spasial terdeteksi. GWR adalah metode statistika yang menghasilkan model bersifat lokal, yaitu hanya berlaku untuk masing-masing lokasi sehingga perbedaan varians tidak berpengaruh pada data.

Model GWR dapat dirumuskan sebagai berikut (Fotheringham, dkk: 2002)

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^K \beta_k(u_i, v_i)x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n; k = 1, 2, \dots, p \quad (2.1)$$

Dengan

y_i : variabel respon pada lokasi ke- i
 x_{ik} : variabel prediktor ke- k pada lokasi ke- i
 (u_i, v_i) : koordinat pada titik i , dimana u_i menunjukkan *longitude* dan v_i menunjukkan *latitude*, yang selanjutnya akan dilambangkan dengan i

$\beta_0, \beta_k(u_i, v_i)$: intersep dan koefisien regresi ke- k pada lokasi ke- i ;

ε_i : residual pada observasi ke- i

bentuk matriks dapat dituliskan sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1n} \\ 1 & x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix} \begin{bmatrix} \beta_0(i) \\ \beta_1(i) \\ \vdots \\ \beta_p(i) \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Berdasarkan persamaan (2.1), diasumsikan bahwa data observasi yang dekat dengan lokasi ke- i memiliki pengaruh yang

lebih besar dalam mengestimasi $\beta_k(i)$ daripada data observasi yang jauh dari lokasi ke- i , sehingga dalam menaksir parameter GWR, biasanya dilakukan dengan menggunakan metode kuadrat terkecil terboboti atau yang lebih dikenal dengan nama *Weighted Least Square* (WLS). Data dengan observasi yang lebih dekat dengan lokasi ke- i memiliki bobot yang lebih tinggi dari pada data yang jauh dari lokasi ke- i . Persamaan WLS dalam mengestimasi parameter dapat dituliskan sebagai berikut:

$$\sum_{j=1}^n w_j(i) \varepsilon_j^2 = \sum_{j=1}^n w_j(i) (y_j - \beta_0(i) - \sum_{k=1}^K \beta_k(i) x_{jk})^2 \quad (2.2)$$

dalam bentuk matriks:

$$\varepsilon^T W(i) \varepsilon = Y^T W(i) Y - 2 \beta^T(i) X^T W(i) Y + \beta^T(i) X^T W(i) X \beta(i)$$

Jumlah kuadrat error akan minum jika persamaan didiferensiasi sehingga parameter β dapat diestimasi, dengan formula sebagai berikut:

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

2. Pedeteksian *Outlier*

Salah satu kelemahan dari metode GWR adalah sensitif terhadap *outlier*., sehingga mengakibatkan residual semakin besar dan estimasi parameter dengan menggunakan *Weighted Least Square* tidak akan menghasilkan parameter bersifat BLUE.

Keberadaan *outlier* dapat dideteksi dengan berbagai cara, salah satunya adalah dengan menggunakan *Boxplot*. Metode *Boxplot* menggunakan jangkauan interkuartil sebagai batas untuk mendekteksi *outlier*. Nilai jangkauan antarkuartil dapat dihitung dengan menggunakan rumus:

$$IQR = Q_3 - Q_1$$

data dikatakan *outlier* jika:

1. nilainya lebih kecil dari $\{Q_1 - (1.5 \times IQR)\}$
2. nilainya lebih besar dari $\{Q_3 + (1.5 \times IQR)\}$

Pengecekan *outlier* juga dapat dilakukan melalui peta area, yaitu dengan menggunakan metode *boxmap*. Dalam pengecekan *outlier*, metode ini menggunakan

cara yang sama dengan *boxplot* namun, *outlier* ditampilkan dalam bentuk peta area.

3. Pembobot Nilai RGWR

Metode yang digunakan untuk menghitung nilai pembobot adalah Fungsi kernel. Fungsi Kernel memberikan pembobot sesuai dengan parameter penghalus (bandwidth). Beberapa jenis fungsi kernel adaptif yaitu:

1. Fungsi Kernel adaptif *Gaussian*

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

2. Fungsi kernel *exponential*

$$w_{ij} = \exp \left[-\left(\frac{d_{ij}}{h_i} \right) \right]$$

4. Bandwidth Optimum pada RGWR

Pemilihan bandwidth memiliki pengaruh yang besar terhadap hasil yang diperoleh maka dibutuhkan metode dapat digunakan untuk mencari nilai h (bandwidth) yang optimum adalah metode *Cross Validation* (CV) dengan rumus sebagai berikut:

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(h))^2 \quad (2.3)$$

Dengan $\hat{y}_{\neq i}(h)$ adalah penaksir dari y_i , dengan pengamatan ke- i dihilangkan dari proses estimasi. Namun, metode CV yang digunakan untuk pemilihan *bandwidth* optimum sangat sensitif terhadap *outlier*, karena residual dari *outlier* akan mendominasi nilai keseluruhan dari CV. Alternatif lain yang dapat digunakan adalah dengan menggunakan metode *Absolute Cross Validation* (ACV). ACV menggunakan nilai mutlak untuk mengestimasi $\hat{y}_{\neq i}(h)$ dengan rumus sebagai berikut:

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

Kriteria pemilihan *bandwidth* optimum adalah dengan melihat nilai ACV yang paling kecil.

5. Pengujian Efek Spasial

Efek heterogenitas spasial diuji dengan menggunakan uji *Breusch-Pagan* (Anselin, 1998). Hipotesisnya yang digunakan adalah $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (tidak ada efek heterogenitas spasial)

H_1 : minimal ada satu $\sigma_1^2 \neq \sigma^2$ (ada efek heterogenitas spasial)

Statistik uji:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \sim \chi^2_{(K)}$$

dimana

$$\mathbf{f} = (f_1, f_2, \dots, f_n)^T$$

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right)$$

$$e_i = y_i - \hat{y}_i$$

dengan

σ^2 : varians error

\mathbf{Z} : matriks variabel prediktor berukuran $n \times (K + 1)$ hipotesis H_0 ditolak jika $BP > \chi^2_{(K)}$ dengan K adalah banyak variabel.

6. *Robust Geographically Weighted Regression*

Seperti yang telah dibahas sebelumnya, model GWR sensitif terhadap outlier karena estimasi parameter menggunakan Weighted Least Square (WLS), dengan meminimumkan jumlah kuadrat residual, sehingga diperlukan teknis yang robust untuk mengatasi adanya outlier. Salah satu metode yang bisa digunakan adalah metode Least Absolute Deviation (LAD). Metode LAD menaksir parameter dengan meminimumkan harga mutlak residual, sehingga efek dari outlier pada penaksir LAD lebih kecil dibandingkan penaksir WLS. Estimasi parameter dengan menggunakan metode LAD juga tidak membutuhkan asumsi residual harus berdistribusi normal, sehingga salah satu asumsi regresi linear dapat dilanggar, yaitu normalitas. Metode LAD baik digunakan karena mengestimasi menggunakan pendekatan median sehingga robust terhadap outlier. Penelitian mengenai penggunaan LAD untuk mengatasi outlier telah banyak

dilakukan, seperti penelitian Zhang dan Mei (2011) menggabungkan model Local Linear GWR dan Basic GWR dengan LAD. Model RGWR merupakan pengembangan dari model GWR, sehingga persamaan model RGWR sama dengan model GWR, yaitu persamaan (2.1), perbedaannya hanya terletak pada pemilihan bandwidth optimum. Maka, persamaan model RGWR dalam pemograman linear adalah

meminimumkan $\sum_{i=1}^n |\varepsilon_i| w_i(u_0, v_0)$

Kendala $y_i = \beta_0(u_0, v_0) -$

$\sum_{k=1}^K \beta_k(u_0, v_0) x_{ik} + \varepsilon_i$

$\varepsilon_i = \varepsilon_i^+ - \varepsilon_i^-$ dan $|\varepsilon_i| = \varepsilon_i^+ + \varepsilon_i^-$

Dengan $\varepsilon_i^+ = |\varepsilon_i| I(\varepsilon_i > 0)$

$\varepsilon_i^- = |\varepsilon_i| I(\varepsilon_i < 0)$

Maka, persamaan model RGWR dalam pemograman linear adalah

Meminimumkan $\sum_{i=1}^n (\varepsilon_i^+ + \varepsilon_i^-) w_i(u_0, v_0)$

Dengan kendala $y_i = \beta_0(u_0, v_0) -$

$\sum_{k=1}^K \beta_k(u_0, v_0) x_{ik} + \varepsilon_i^+ - \varepsilon_i^-$

$\varepsilon_i^+, \varepsilon_i^- \geq 0$

METODE PENELITIAN

1. Sumber Data

Data yang digunakan pada penulisan tugas akhir ini berupa data sekunder tentang faktor-faktor yang mempengaruhi penyebaran kasus Covid-19 yang terdiri dari 34 provinsi di Indonesia. Data kasus Covid-19 diperoleh dari situs resmi penyebaran Covid-19 di Indonesia yang dihimpun dari tanggal 2 Maret 2020 hingga tanggal 24 Mei 2021 melalui www.covid.go.id.

2. Variabel Penelitian

Tabel 1 Variabel Penelitian

Jenis Variabel	Nama
Dependen (Y)	Kasus Positif Covid-19
Independen (X)	X ₁ : Kepadatan Penduduk
	X ₂ : Persentase Penduduk Lansia
	X ₃ : Persentase Tingkat Kemiskinan
	X ₄ : Persentase Rumah Tangga yang memiliki Sumber Air Minum Layak

3. Analisis Data

Tahapan analisis yang digunakan dalam penelitian ini adalah :

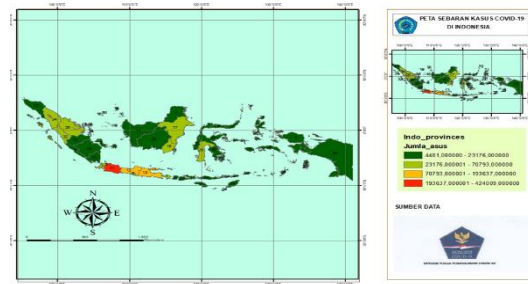
1. Analisis data yang digunakan pada penelitian ini ada beberapa langkah yang digunakan. Langkah-langkah yang akan digunakan antara lain:
2. Melakukan analisis deskriptif terhadap penyebaran kasus positif Covid-19 di Indonesia.
3. Melakukan analisis regresi linier berganda terhadap variabel-variabel bebas.
4. Melakukan uji heterogenitas spasial pada regresi linier berganda dengan menggunakan uji Breusch Pagan.
5. Melakukan pendektasian pencilan dengan menggunakan boxplot yang kemudian ditampilkan dengan boxmap.
6. Melakukan analisis Robust Geographically Weighted Regression, apabila terjadi heterogenitas spasial dan terdapat pencilan yang terdiri dari:
7. Menghitung jarak Euclidean
8. Menentukan bandwidth optimum dengan menggunakan metode Absolute Cross-Validation
9. Menghitung matriks pembobot dengan fungsi kernel Adaptif Gaussian Kernel
10. Melakukan estimasi parameter model Robust GWR dengan Least Absolute Deviation.
11. Menghitung ukuran ketepatan model dengan MAPE
12. Interpretasi hasil analisis dan menarik kesimpulan

HASIL DAN PEMBAHASAN

1. Analisis Deskriptif

Penelitian ini dilakukan untuk melihat ada atau tidaknya pengaruh lokasi terhadap penyebaran kasus Covid-19 di Indonesia. Indonesia merupakan salah satu negara yang berdampak Covid-19 yang memiliki 34 provinsi. Banyaknya data pengamatan pada penelitian ini yaitu 34 data yang masing-

masing dari 34 provinsi di Indonesia diperoleh dari situs resmi penyebaran Covid-19 di Indonesia dan publikasi lembaga pemerintahan resmi yakni Badan Pusat Statistik (BPS). Data kasus Covid-19 dapat dilihat pada peta tematik di bawah ini.



Berdasarkan terdapat 4 kategori yaitu warna merah dengan kategori sangat tinggi, warna orange dengan kategori tinggi, warna hijau muda dengan kategori sedang dan warna hijau tua untuk kategori rendah. Kasus positif Covid-19 terendah dan tertinggi adalah provinsi Maluku Utara dan DKI Jakarta sebesar 4.481 dan 424.009, dengan rata-rata kasus positif Covid-19 di Indonesia adalah 52.199,26. Variabel jumlah kasus positif Covid-19 dihubungkan dengan variabel bebas dengan menggunakan software R 4.0.3 didapatkan hasil estimasi parameter berikut:

2. Regresi Linier Berganda

Langkah selanjutnya adalah membentuk model regresi linier dengan metode Ordinary Least Square (OLS).

Tabel 1 Hasil Estimasi Parameter

Parameter	Nilai Estimasi	P-Value	R ²
Intercept	8.182×10^{-18}	1	
Kepadatan penduduk	0.4657	0.00481	
Penduduk Lansia	0.0432	0.79214	
Persentase Penduduk Miskin	0.3786	0.01929	0.5424
Sumber Air Minum Layak	0.1053	0.5006	

Berdasarkan Tabel 4.2 didapatkan model awal regresi dengan metode kuadrat terkecil yaitu:

$$\hat{Y} = 8.182 \times 10^{-18} + 0.4657X_1 + 0.04320X_2 + 0.378X_3 + 0.1053X_4$$

Namun, dari 4 variabel independent hanya 2 variabel bebas yang signifikan terhadap model yaitu variabel X1 (kepadatan penduduk) dan X3 (tingkat kemiskinan).

3. Uji Heterogenitas Spasial

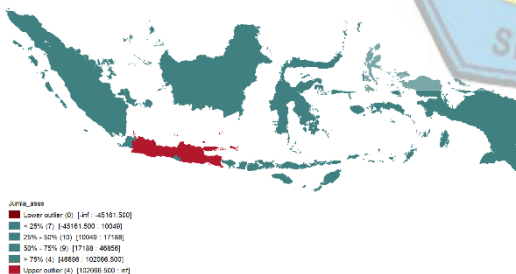
Tabel 2 Uji Heterogenitas Spasial

<i>Breusch-Pagan</i>	<i>df</i>	<i>P-Value</i>
11.072	4	0.02576

Berdasarkan Tabel 4.3 diatas dapat dilihat nilai P-value sebesar 0.02576 lebih kecil dari $\alpha = 0.05$, sehingga menunjukkan bahwa terdapat heterogenitas spasial.

4. Pendektesian Pencilan Spasial

Metode yang digunakan untuk mendeteksi pencilan pada penelitian ini adalah boxplot yang kemudian ditransformasi ke dalam bentuk peta yang disebut boxmap. Boxmap ini dapat memperlihatkan data pencilan secara spasial. Berikut adalah boxmap penyebaran kasus covid-19 di Indonesia.



Gambar 3. Provinsi yang Merupakan Pencilan Spasial

Berdasarkan hasil boxmap diatas dapat disimpulkan bahwa terdapat empat pencilan spasial yang ditandai dengan warna merah yaitu pada lokasi ke-11 (DKI Jakarta), lokasi ke - 12 (Jawa Barat), lokasi ke-13 (Jawa Tengah), dan lokasi ke-15 (Jawa Timur). Sehingga model RGWR dengan metode *least*

absolute deviation dapat digunakan pada penelitian ini.

5. Geographically Weighted Regression

a. Bandwith Optimum

Tabel 4 Nilai Cross Validation (CV) dan Bandwidth Optimum

Fungsi Pembobot	Bandwidth	CV
Gaussian	4929.933	22.71284

Berdasarkan tabel diatas menghasilkan nilai bandwidth optimum 4929.933 dan nilai CV yang paling minimum adalah 22.71284, maka dapat ditentukan matriks pembobot spasial.

b. Matriks Pembobot

Fungsi yang digunakan untuk menghitung matriks pembobot GWR adalah *fixed gaussian kernel* karna menghasilkan nilai CV terkecil.

c. Model GWR

Sebelum memodelkan penyebaran Covid-19 di Indonesia menggunakan RGWR, terlebih dahulu akan dimodelkan dengan GWR. Berikut merupakan ringkasan parameter model GWR dari variabel yang berpengaruh:

Tabel 5 Ringkasan Estimasi Parameter Model GWR

Parameter	Min	Max	Mean	Med	Standard Deviation
β_0	-0,000	0,004	0,002	0,00	0,1160
β_1	70691	30400	2	106	00
β_2	0,026	0,060	0,046	0,04	0,1626
β_3	0,363	0,392	0,383	0,38	0,1527
β_4	0,097	0,113	0,108	0,10	0,15440

Tabel 4.5 diatas dapat dilihat variabel yang memberikan pengaruh positif terhadap variabel respon adalah variabel kepadatan penduduk, persentase penduduk lanjut usia, persentase tingkat kemiskinan dan persentase sumber air minum yang layak.

6. RGWR

Dengan menggunakan metode boxmap, data penelitian memiliki pencilan, sehingga setiap kota/kabupaten harus menggunakan teknik yang robust, yaitu model RGWR. Adapun yang dilakukan dalam pemodelan RGWR adalah:

a. Bandwith Optimum

Tabel 6 Nilai Absolute Cross Validation (ACV) dan Bandwidth Optimum

Fungsi Pembobot	Bandwidth	ACV
Gaussian	30	12.22702

Berdasarkan kriteria ACV, dihasilkan bandwidth *Adaptive Gaussian Kernel* sebagai berikut:

b. Matriks Pembobot

Fungsi yang digunakan untuk menghitung matriks pembobot RGWR adalah *Adaptive Gaussian Kernel* karena menghasilkan nilai ACV terkecil.

c. Model RGWR

Ringkasan dari estimasi parameter untuk model RGWR yang diperoleh dengan menggunakan *software R* dan *package rq* adalah sebagai berikut:

Tabel 3 Ringkasan Estimasi Parameter Model RGWR

Parameter	Min	Q1	Med	Q3	Mean	Max
β_0	-	-	-	-	-	-
	0.11708	0.09758	0.09758	0.09329	0.09741	0.07082
β_1	0.5598	0.5598	0.5641	0.6020	0.5877	0.6593
β_2	0.01263	0.01263	0.10057	0.1026	0.0767	0.10267
β_3	0.02976	0.06398	0.08402	0.0840	0.07150	0.09001
β_4	-	-	0.002819	0.0031	-	0.004616
	0.032389	0.01485	0.002819	0.003163	0.006539	0.004616

Tabel 6 diatas dapat dilihat variabel yang memberikan pengaruh positif terhadap variabel respon adalah variabel kepadatan penduduk, persentase penduduk lanjut usia, persentase tingkat kemiskinan. Sedangkan persentase sumber air minum yang layak memberikan pengaruh negatif terhadap kasus positif Covid-19 di beberapa wilayah Indonesia. Berikut akan ditampilkan salah satu model yaitu model Provinsi Maluku Utara:

$$\hat{Y}_{32} = -0.11708119 + 0.6593263X_1 + 0.01263260X_2 + 0.02976439X_3 - 0.014854042X_4$$

Berdasarkan model RGWR diatas, maka dapat disimpulkan bahwa pada Provinsi Maluku Utara, variabel yang dapat meningkatkan kasus positif Covid-19 adalah variabel kepadatan penduduk, persentase tingkat kemiskinan, dan persentase penduduk lansia. Sedangkan persentase rumah tangga yang memiliki sumber air minum yang layak memberikan pengaruh negatif terhadap kasus positif Covid-19 di Provinsi Maluku Utara.

7. Ukuran Ketepatan Model

Perbandingan nilai estimasi dengan menggunakan model GWR, dan RGWR dapat dilihat pada MAPE.

Tabel 4 Nilai MAPE

Model	MAPE
GWR	162,2844
RGWR	92,08409

Berdasarkan table 4.8 diatas, dapat disimpulkan bahwa model RGWR memiliki nilai MAPE lebih kecil daripada model GWR yaitu sebesar 92,08409 sehingga pemodelan RGWR mampu menghasilkan estimasi parameter yang lebih baik dibandingkan model GWR.

KESIMPULAN

1. Rata-rata kasus positif Covid-19 di Indonesia adalah 52.199,26. Provinsi dengan kasus positif tertinggi dan terendah di Indonesia adalah provinsi DKI Jakarta dan Maluku Utara.
 2. Terdapat 34 model penyebaran kasus Covid-19 di Indonesia menggunakan RGWR dengan metode LAD dan memberikan hasil yang berbeda-beda untuk setiap provinsi.
 3. Berdasarkan besar dan kecilnya nilai koefisien variabel independent yang memiliki pengaruh terbesar adalah kepadatan penduduk, persentase penduduk lansia, persentase tingkat kemiskinan dan persentase rumah tangga yang memiliki sumber air minum yang layak.
 4. Model RGWR mampu memberikan hasil estimasi yang lebih baik dibandingkan dengan model GWR berdasarkan nilai MAPE sebesar 92,08409. Hal ini menunjukkan bahwa model ini cocok digunakan untuk mengatasi efek heterogenitas varias serta mampu mengatasi adanya pencilan.
- DAFTAR PUSTAKA**
- Afifah. 2017. Robust Geographically Weighted Regression with Least Absolute Deviation Method in Case of Poverty in Java Island. AIP Conference Proceedings.
- Anselin, L. 1988. Spatial Econometrics: Method and Models. Kluwer Academic Publisher: The Netherlands.
- Anselin, L. 2005. Exploring Spatial Data with GeoDaTM: A Workbook, Spatial Analysis Laboratory Department of Geography University of Illinois. Urbana-Champaign Urbana, IL 61801.
- Ari & Sengupta .2020. Nexus Between Population Density and Novel Coronavirus (Covid-19) Pandemic in The South Indian States
- A.S. Fotheringham, C. Brunsdon and M. Charlton. 2002. Geographically Weighted Regression. Jhon Wiley and Sons, Chicester UK
- D. Birkes & Y. Dodge.1993. Alternative Methods of Regression. New York.
- Fatihurizqi & Fitriah.2020. Pemodelan Kasus Covid-19 di Indonesia dengan Pendekatan regresi Poisson dan Regresi Binomial Negatif. Seminar Nasional Official Statistics 2020.
- Hanoatubun, S.2020.Dampak Covid-19 terhadap Perekonomian Indonesia. EduPsyCouns: Journal of Education, Psychology and Conseling.
- Hardianto.J.2020.Korelasi Kepadatan Penduduk dan Penyebaran Covid-19.Rujuk Center For Urban Studies.
- Kementerian Kesehatan Republik Indonesia.2020. Hindari Lansia dari Covid-19. Kementerian Kesehatan Republik Indonesia
- Kodoatie, R. J.2003.Manjemen dan Rekayasa Infrastruktur. Yogyakarta: Pustaka Pelajar.
- Levani, Y., Prasetya, A.D., & Mawaddatunnadila, S. 2021. Coronavirus Disease 2019 (Covid-19): Patogenesis, Manifestasi Klinis dan Pilihan Terapi. Jurnal Kedokteran dan Kesehatan, Volume 17, Nomor 1.
- Mahdy, Faisal Ilham .2020. Pemodelan Jumlah Kasus Covid-19 di Jawa Barat Menggunakan Geographically weighted Regression. Seminar Nasional Official Statistic.
- Myttenaere, D.A., Golden, B., Grand, B. L., & Rossi, F. 2016. Mean Absolute Percentage Error for Regression Models. Neurocomputing, Elsevier, 2016, 23rd European Symposium on Artificial Neural Networks (ESANN 2015), Volume 192.

- Nurhayati. 2018. Robust Geographically Weighted Regression dengan Metode Mutlak Simpangan terkecil pada Pemodelan Kejadian Diare di Kota Semarang. *Jurnal Gaussian*, Volume 7, Nomor 2.
- Perhimpunan Dokter Paru Indonesia. 2020. *Pneumonia Covid-19: Diagnosis & Penatalaksanaan di Indonesia*. Perhimpunan Dokter Paru Indonesia (PDPI)
- Suripin. 2002. *Pelestarian Sumber Daya Tanah dan Air*. Yogyakarta. Andi
- Wagner, H.M. 1959. Linear Programming Techniques for Regression Analysis. *Journal of the American Statistical Association*, Volume 54, Nomor 285.
- Wang, F.T. & Scott, D.W. 1994. The L1 Method for Robust Non Parametric Regression. *Journal of the American Statistical Association*, Vol. 89, No.425.
- Wang, Z., Qiang, W., Ke, H. 2020. *A Handbook of 2019-nCoV Pneumonia Control and Prevention*. Hubei Science and Technologi Press. China.
- World Health Organization. 2020. *Water, Sanitation, Hygienie, and Waste management for the Covid-19 Virus: interm guidance*, 23 April 2020. World Health Organization.
- Zhang, H. & Mei. C. 2011. Local Least Absolute Deviation Estimation of Spatially Varying Coefficient Models: Robust Geographically Weighted Regression Approachess. *Internasional Journal of Geographical Information Science*. Vol 25:9, 1467:1489.

