

Pemodelan *Geographically Weighted Regression* Menggunakan Pembobot Kernel *Fixed* dan *Adaptive* pada Kasus Tingkat Pengangguran Terbuka di Indonesia

Mila Rizki Ramadayani ^{1,a)}, Fariani Hermin Indiyah ^{1,b)}, Ibnu Hadi ^{1,c)}

¹Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta

Email: ^{a)}milarizki69@gmail.com, ^{b)}farianihermin828@gmail.com, ^{c)}ibnuhadiunj@gmail.com

Abstract

Unemployment Rate (UR) is an indicator for measuring the unemployment. Increase in the number of TPT in Indonesia by 1.84%, this is due to the impact of the covid-19 pandemic. analysis to find out the factors that affect TPT in Indonesia is by using multiple linear regression. The results showed that the data contained heterokedasticity and spatial aspects. Spatial data analysis continued with the point approach is by the Geographically Weighted Regression method (GWR). GWR is a weighted regression that results in a model that is local. GWR modeling uses weighting kernels Fixed Gaussian, Adaptive Gaussian, Fixed Bi-Square, and Adaptive Bi-Square produces that GWR Adaptive Bi-Square better, review value of the R², AIC and JKG. The ability of the GWR model explains the effect of UR on factors (Labor Force or economically active, Health Complaint and Poverty Percentage) by 89.1%.

Keywords: Unemployment Rate(UR), Spatial Data, GWR, and Kernel Weight.

Abstrak

Tingkat Pengangguran Terbuka (TPT) merupakan indikator untuk mengukur angka pengangguran. Peningkatan Angka TPT di Indonesia sebesar 1.84%, demikian terjadi karena dampak dari pandemi Covid-19 di Indonesia. Analisis untuk mengetahui faktor yang mempengaruhi TPT di Indonesia yaitu dengan Regresi Linear Berganda. Pada hasilnya menunjukkan bahwa data mengandung heterokedastisitas dan terdapat aspek spasial. Analisis data spasial dilanjutkan dengan Pendekatan titik yaitu metode Geographically Weighted Regression (GWR). GWR merupakan regresi terboboti sehingga menghasilkan model yang bersifat lokal. Pemodelan TPT dengan kernel pembobot Fixed Gaussian, Adaptive Gaussian, Fixed Bi-Square, dan Adaptive Bi-Square menghasilkan model GWR Adaptive Bi-Square yang terbaik. meninjau dari nilai R², AIC dan JKG. Kemampuan model GWR dalam menjelaskan pengaruh TPT terhadap faktor-faktor (Angkatan kerja, Keluhan Kesehatan dan Persentase Kemiskinan) sebesar 89.1%.

Kata kunci: TPT, Data Spasial, GWR, dan Pembobot Kernel.

PENDAHULUAN

Salah satu dampak yang dialami oleh beberapa negara saat pandemi corona yaitu sektor ketenagakerjaan, dimana selama pandemi banyak aktivitas diberhentikan sementara, pengurangan upah atau gaji, bahkan banyak juga pekerja yang terkena Pemutusan Hubungan Kerja (PHK). Hal ini menyebabkan Pencari Kerja Meningkat, Penyerapan Tenaga Kerja Yang Terbatas dan Angka Pengangguran Meningkat.

Untuk membuat suatu kesimpulan tentang suatu permasalahan maka cara preventif yang bisa

dilakukan dengan memprediksi faktor-faktor yang mempengaruhinya. Salah satu analisis yang digunakan yaitu regresi linear berganda. Pada model regresi biasa dengan Regresi OLS Menghasilkan estimasi parameter yang sama pada semua wilayah, namun pada kenyataan kondisi data lokasi satu tidak sama dengan lokasi lainnya. Dengan demikian, asumsi kehomogenan sulit dipenuhi dimana data mengandung efek spasial (lokasi). Penelitian data spasial dengan menggunakan pendekatan titik yaitu *Geographically Weighted Regression* (GWR).

GWR merupakan metode statistik yang menggunakan informasi geografis dan menjadi bagian regresi terboboti. Kelebihan model GWR dibanding dengan regresi klasik menurut jurnal-jurnal terkait yaitu dapat memberikan model secara lokal. Terdapat empat kernel pembobot berbeda pada GWR yaitu *fixed gaussian*, *adaptive gaussian*, *fixed bi-square* dan *adaptive bi-square*, sehingga mampu menemukan kernel GWR terbaik pada model TPT.

TINJAUAN PUSTAKA

Tingkat Pengangguran Terbuka

Pengangguran merupakan istilah untuk seseorang yang tidak bekerja sama sekali, mencari pekerjaan, dan bekerja kurang dari dua kali seminggu. Menurut BPS (2021) Tingkat pengangguran terbuka adalah indikator yang digunakan dalam mengukur tenaga kerja yang tidak diserap oleh pasar kerja dan menggambarkan kurang termanfaatkannya pasokan tenaga kerja. Saat ini Indonesia masih berada dalam masa bonus demografi. Memperbaiki dan meningkatkan kualitas pemuda di Indonesia adalah upaya yang sangat dibutuhkan dalam menghadapi angkatan kerja yang terus meningkat. Tingkat Pengangguran Terbuka (TPT) sebagai perbandingan banyaknya pengangguran terhadap banyaknya angkatan kerja. Diantara faktor yang mempengaruhi pengangguran yaitu angkatan kerja, keluhan kesehatan, dan persentase penduduk miskin.

Regresi Linear Berganda

Analisis regresi adalah metode umum untuk mendapatkan fungsi prediksi untuk memprediksi nilai variabel respon Y menggunakan variabel prediktor (x_1, \dots, x_k) . Regresi linear berganda digunakan untuk mendapati adanya pengaruh interaksi dua atau lebih variabel bebas terhadap variabel terikat. Berikut adalah model umum dari regresi linear berganda:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

Jika penelitian menggunakan sebanyak n pengamatan, model yang terbentuk adalah:

$$Y_i = \beta_0 + \sum_{k=1}^p \beta_k X_{ik} + \varepsilon_i$$

dengan $i = 1, 2, \dots, n$ pengamatan. bentuk di atas dapat juga disederhanakan dengan penulisan :

$$Y = X\beta + \varepsilon$$

Y merupakan vektor $n \times 1$, X merupakan vektor $n \times (k + 1)$ dan $\hat{\beta}$ merupakan vektor $(k + 1) \times 1$.

Metode yang digunakan untuk mendapatkan parameter pada model di sini yaitu *Ordinary Least Square* (OLS) atau biasa disebut metode kuadrat terkecil. Metode OLS dijalankan menggunakan nilai *error* kuadrat yang perhitungannya lebih mudah menggunakan matriks untuk menghitung taksiran kuadrat terkecil dari $\beta_0, \beta_1, \dots, \beta_k$. Solusi dari OLS ini adalah

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

Terdapat empat tahapan uji yang harus dilakukan untuk memeriksa validnya suatu model yaitu Uji homokedastisitas, uji normalitas, uji multikolinearitas, dan uji autokorelasi.

Pengujian hipotesis yang paling penting dalam regresi salah satunya adalah uji F. Setidaknya terdapat variabel independen yang berpengaruh terhadap variabel dependen, sehingga uji F untuk mengetahui bagaimanakah pengaruh semua variabel independen secara seluruh (simultan) terhadap

variabel dependen. Statistik uji F yaitu :

$$F_{hit} = \frac{\frac{JKR}{P}}{\frac{JKG}{n-p-1}}$$

dimana :

p = Banyaknya variabel bebas yang digunakan dalam penelitian

$$JKR = \hat{\beta}'X'y - n\bar{y}^2$$

$$JKG = y'y - \hat{\beta}'X'y$$

$$JKT = \sum_{i=1}^n (y_i - \bar{y})^2 = y'y - n\bar{y}^2$$

Dengan kriteria pengambilan keputusan model Signifikan ketika $F_{hitung} > F_{table}$ dan kolom signifikansi yaitu $sig < \alpha$ (selang kepercayaan).

Uji parsial atau bisa dikenal dengan uji T, fungsinya yaitu akan menguji pengaruh masing masing variabel independen terhadap variabel dependen. Statistik uji yang digunakan yaitu :

$$t_{hit} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}$$

$$SE(\hat{\beta}_k) = \sqrt{C_{(p+1)}S}$$

keterangan :

β_k : Nilai koefisien pengamatan ke- k ; $k = 0, 1, \dots, p$

$SE(\beta_k)$: Standar error koefisien pada pengamatan ke- k ; $\sqrt{C_{(p+1)}S}$ unsur ke $(p+1)$ diagonal $(xx)^{-1}$

S : Akar dari kuadrat rata-rata galat

Kriteria pengambilan keputusan yaitu Tolak H_0 jika $t_{hit} > t_{((n-p-1), \frac{\alpha}{2})}$

dengan α atau taraf nyata 5 %

Data Spasial

Data yang mengandung unsur geografis dan menggunakan koordinat lintang dan bujur sebagai dasar referensi disebut data spasial. Data spasial memiliki dua bagian penting yaitu informasi lokasi (*spatial*) dan informasi deskriptif (*attribute*). Pada penelian terhadap data spasial dibutuhkan konsen lebih karena menyelesaikan masalah berbasis lokasi sehingga memungkinkan terjadinya korelasi antar lokasi yang berdekatan. Pada data spasial adanya perbedaan kondisi wilayah, karena karakteristik lingkungan dan geografis memungkinkan terjadi heterogenitas spasial.

Geographically Weighted Regression

Geographically Weighted Regression merupakan metode regresi spasial dengan pendekatan titik yang berbasis wilayah. Model ini menghitung parameter atas setiap lokasi pengamatan atau dengan kata lain memperhitungkan lokasi data pengamatan (Ira 2016). Model GWR juga merupakan pengembangan regresi lokal dengan *Ordinary Least Square (OLS)* menjadi regresi terboboti dengan *Weighted Least Square (WLS)* yang memperhatikan efek spasial. Nilai yang dihasilkan setiap Parameter berbeda-beda untuk setiap wilayah. Bentuk umum dari model GWR yaitu :

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

keterangan :

$i = 1, \dots, n$

Y_i = Variabel dependen lokasi ke- i

(u_i, v_i) = Koordinat lintang bujur pada lokasi ke- i

$\beta_0(u_i, v_i)$ = Konstanta GWR

$\beta_k(u_i, v_i)$ = Parameter atau koefisien regresi ke- k

X_{ik} = Variabel prediktor ke- k pada lokasi ke- i

ε_i = Residual dengan asumsi $\varepsilon \sim IIDN(0, \sigma^2)$.

(Fotheringham, Brunsdon, dan Charlton, 2002)

Metode yang digunakan untuk mengetahui penaksiran parameter di setiap model GWR yaitu metode *Weighted Least Square* (WLS), dimana setiap lokasi diberikan bobot yang berbeda sesuai dengan titik pengamatan dimana data tersebut diambil. Notasi pembobot yang digunakan pada setiap lokasi adalah $w_j(u_i, v_i)$ dengan $j = 1, 2, \dots, n$ lokasi. Estimasi parameter GWR $\beta_k(u_i, v_i)$ untuk setiap variabel ke- k pada lokasi pengamatan ke- i , sebagai berikut:

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

Keterangan :

$\hat{\beta}(u_i, v_i)$ = Parameter duga lokasi ke- i

X = Matriks variabel prediktor berukuran $n \times (p + 1)$

$W(u_i, v_i)$ = Matriks bobot lokasi ke- i berukuran $n \times n$

Y = Vektor variabel respon berukuran $n \times 1$

Bandwidth GWR

bandwidth diartikan sebagai *radius* dari titik lokasi pusat sebagai acuan dalam penentuan bobot setiap wilayah untuk model regresi. Ketika pengamatan masih dalam radius artinya model pada lokasi tersebut masih dianggap memiliki pengaruh, maka model dapat diberi bobot. Untuk menghindari heterogen dari varian, haruslah nilai *bandwidth optimum*. Menurut (Fotheringham dkk. 2002) terdapat beberapa metode penentuan bandwidth optimum salah satunya yaitu metode *Cross Validation* (CV). Berikut persamaan CV secara matematis :

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

dimana $\hat{y}_{\neq i}$ merupakan penduga y_i dimana lokasi pengamatan (u_i, v_i) dihilangkan dari proses estimasi guna mendapat radius yang optimal, sehingga mendapat CV minimum.

Pembobot GWR

Analisis model yang memperhatikan aspek spasial memerlukan pembobotan dalam perhitungannya. Hubungan kedekatan (neighboring) antarlokasi pengamatan dinyatakan dalam bentuk matriks pembobot W_{ij} (Bekti 2012). Elemen pada matriks berisi W_{ij} yang menyatakan kedekatan lokasi i dengan lokasi j . Jika bobotnya semakin besar maka lebih berpengaruh. Sedangkan pembobot dengan fungsi kernel mampu memberikan bandwidth optimum dimana nilainya sesuai dengan keadaan data. Terdapat empat jenis kernel pembobot diantaranya yaitu :

1. Fungsi Kernel Fixed Bi-Square

$$W_j(u_i, v_i) = \begin{cases} (1 - (\frac{d_{ij}}{b})^2)^2 & : \text{jika } d_{ij} \leq b \\ 0 & : \text{lainnya} \end{cases}$$

dengan $W_j(u_i, v_i)$ merupakan matriks pembobot spasial untuk setiap lokasi, d_{ij} sebagai jarak euclidean dan b adalah *bandwidth* yang sama untuk semua lokasi

2. Fungsi Kernel Adaptive Bi-Square

$$W_j(u_i, v_i) = \begin{cases} (1 - (\frac{d_{ij}}{b})^2)^2 & : \text{jika } d_{ij} \leq b \\ 0 & : \text{lainnya} \end{cases}$$

dimana $W_j(u_i, v_i)$ merupakan matriks pembobot, d_{ij} merupakan jarak euclidean. Jadi Fungsi kernel adaptif ini memiliki nilai *bandwidth* (b) yang berbeda untuk setiap titik lokasi pengamatan (Lutfiani, 2017).

3. Fungsi Kernel Fixed Gaussian

$$W_j(u_i, v_i) = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right]$$

dimana $W_j(u_i, v_i)$ menyatakan sebagai matriks pembobot untuk fixed Gaussian, dan d_{ij} adalah jarak euclidean dan b merupakan *bandwidth* yang sama yang digunakan untuk setiap lokasi.

4. Fungsi Kernel Adaptive Gaussian $W_j(u_i, v_i) = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b_{i(q)}}\right)^2\right]$

dengan $W_j(u_i, v_i)$ merupakan matriks pembobot untuk setiap lokasi, d_{ij} sebagai jarak antar lokasi pengamatan, dan $b_{i(q)}$ merupakan *bandwidth* adaptif atau berbeda untuk setiap titik lokasi.

Pengujian Hipotesis

1. Pengujian Kelayakan Model

$H_0: \beta_k(u_i, v_i) = \beta_k, k = 1, 2, \dots, n$ (Tidak ada pengaruh faktor geografis pada model)

H_1 : Paling sedikit ada satu $\beta_k(u_i, v_i)$ yang berhubungan dengan lokasi (u_i, v_i) (Terdapat pengaruh geografis)

$$F_{hit} = \frac{SSE(H_0)/df_1}{SSE(H_1)/df_2}$$

Dengan kriteria pengambilan keputusan apabila $F_{hit} > F_{table}(F_{(\alpha; df_1, df_2)})$ maka tolak H_0 dengan kata lain terdapat pengaruh geografis sehingga termasuk kedalam model yang layak atau kategori *goodness of fit* lebih baik dari regresi global.

2. Pengujian Paramater

$H_0: \beta_k(u_i, v_i) = 0$ (Tidak ada pengaruh variabel bebas terhadap variabel respon)

$H_1: \beta_k(u_i, v_i) \neq 0$ (Terdapat pengaruh signifikan antara variabel bebas terhadap respon)

$$t_{hit} = \frac{\hat{\beta}_k(u_i, v_i)}{SE[\hat{\beta}_k(u_i, v_i)]}$$

kriteria pengambilan keputusan yaitu apabila $|t_{hit}| > t_{\frac{\alpha}{2}, df_2}$ maka tolak H_0 .

Pemilihan model terbaik

Penentuan model terbaik dapat ditentukan dari nilai R^2 terbesar dan nilai AIC yang terkecil.

1. Koefisien Determinasi (R^2)

$$R^2 = \frac{JKR}{JKT} = \frac{\hat{\beta}_k^T X^T y - n\bar{y}^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Sedangkan dalam model GWR koefisien determinasi dapat ditentukan dengan persamaan berikut :

$$R_i^2 = \frac{JKT_{GWR} - JKSGWR}{JKT_{GWR}}$$

2. Akaike Information Criterion (AIC)

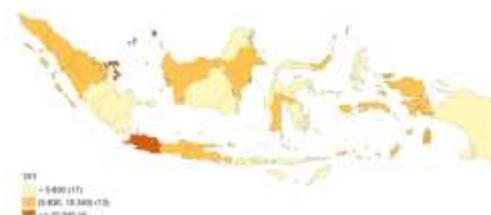
$$AIC = 2n \ln(\sigma) + \ln(2\pi) + n + tr(s)$$

PEMBAHASAN

Deskripsi Data

Pada penelitian analisis model Geographically Weighted Regression (GWR) di sini menggunakan 4 variabel yang terdiri dari satu variabel dependen yaitu Tingkat Pengangguran Terbuka (TPT) dan tiga variabel independen yaitu angkatan kerja (X1), keluhan kesehatan (X2) dan persentase penduduk miskin (X3). Gambaran umum mengenai informasi data disajikan dalam bentuk peta

1. Tingkat Pengangguran Terbuka (TPT)



GAMBAR 1. Tingkat Pengangguran Terbuka

Dari peta Indonesia pada masalah pengangguran di Indonesia masih banyak titik wilayah yang memiliki warna cukup gelap artinya masih memiliki pengangguran dengan jumlah yang tak sedikit. Persentase tingkat pengangguran terbuka yang memiliki kategori sangat tinggi di Indonesia pada 2020 terdapat 4 provinsi yaitu provinsi DKI Jakarta, Banten, Jawa Barat dan Kepulauan Riau.

2. Angkatan Kerja



GAMBAR 2. Angkatan Kerja

Persentase Angkatan Kerja Tertinggi yang tergolong tinggi terdapat di 5 provinsi di Indonesia, mayoritas terletak pada beberapa provinsi di Nusa Tenggara, persentase tertinggi yaitu provinsi Bali dengan persentase 74.32%

3. Keluhan Kesehatan



GAMBAR 3. Keluhan Kesehatan

Persentase Keluhan Kesehatan di Indonesia mayoritas berada di rata-rata yaitu sebanyak 21 provinsi di Indonesia. Namun masih beberapa yang memiliki persentase tinggi sebanyak 3 wilayah yaitu Bali, Aceh, dan Sumatera Barat dengan berturut-turut 61.32% , 57.92% , dan 57.25% .

4. Persentase kemiskinan



GAMBAR 4. Persentase Penduduk Miskin

Provinsi di Indonesia yang tergolong persentase penduduk miskin tertinggi yaitu terletak di Provinsi Papua, Papua Barat, Nusa Tenggara Timur dan Maluku dengan persentase berturut-turut 26.8% , 21.7% , 21.21% dan 17.99%. Penduduk miskin dengan kategori sedang terdapat 12 provinsi, sedangkan kategori rendah cukup banyak yaitu terdapat 18 provinsi dimana terendah terletak di Provinsi Bali dengan 4.45%.

Regresi Linear Berganda

Model dapat dibentuk dalam persamaan regresi linear berganda, yaitu :

$$Y = 29.760 - 0.395X_1 + 0.085X_2 - 0.058X_3$$

Model diatas memiliki nilai R^2 sebesar 0.461, artinya kemampuan model dalam menjelaskan pengaruh sebesar 46.1%. Untuk mengetahui apakah model yang terbentuk benar memiliki pengaruh, maka terlebih dahulu dilakukan pengujian. Untuk mengetahui pengujian model secara keseluruhan dengan menggunakan tabel Analisis of Variance (ANOVA) maka dapat diketahui nilai F_{hitung} .

TABEL 1. ANOVA

Sumber Variasi	DB	Jumlah Kuadrat	Kuadrat Total	F_{hit}	F_{tabel}
Regresi	3	61.522	20.507	8.544	2.92
Galat	30	71.923	2.397		
Total	33	133.445			

Berdasarkan hasil table 4.7 $F_{hit} = 8.554 > F_{(0.05;3;30)} = 2.92$ artinya setidaknya terdapat variabel independen yang memiliki pengaruh signifikan terhadap variabel TPT.

Pada hasil regresi global untuk mengukur validnya suatu model maka dilakukan uji asumsi klasik. Hasil Uji Asumsi Klasik memenuhi normalitas, multikolinearitas, dan autokorelasi . Namun tidak memenuhi homokedastisitas atau mengandung Heterokedastisitas. Ketika dilanjutkan dengan pengujian heterogenitas spasial dengan *Breusch-pagan* hasilnya terbukti mengandung heterogenitas spasial. Maka sesuai metode GWR penelitian dapat dilanjutkan pengujian spasial dan metode GWR dengan 4 pembobot kernel. Sehingga hasil dengan OLS nantinya mengandung presisi yang rendah disebabkan ketidakhomogenan ragam.

Pemodelan *Geographically Weighted Regression*

Ketika data mengandung heterogenitas spasial maka pemodelan dapat dilakukan dengan pendekatan titik yaitu dengan *Geographically weighted Regression* (GWR). Metode regresi spasial ini memang memiliki keunikan dimana model memerlukan pembobot dalam perhitungannya. Pembobotan pada model regresi dipengaruhi oleh *bandwidth* itu sendiri dengan sifat *bandwidth* yang *adaptive* dan *fixed*. Jenis dari pembobotan juga terbagi dua yaitu kernel Gaussian dan kernel Bi-square.

1. Fixed Bi-square

Bandwidth yang digunakan pada fixed bisquare ini merupakan bandwidth yang sama untuk setiap lokasi pengamatan yaitu 16.38876. Estimasi parameter pada model GWR untuk setiap

lokasi pengamatan yang berbeda-beda. Berikut nilai sebaran dari penduga parameter :

TABEL 2. Sebaran Estimasi Fixed Bi-square

	Min	Median	Max
Intercept	19.690109	36.229741	50.309221
X1	-0.725232	-0.489966	-0.061339
X2	-0.325141	0.086031	0.145159
X3	-0.184303	-0.054546	0.199935

Tabel 4.15 menunjukkan bahwa penduga parameter dari angkatan kerja (X1) dan persentase penduduk miskin (X3) bernilai negatif dengan masing- masing memiliki besaran pengaruh X1 sebesar -0.725232 sampai dengan -0.061339 dan pada X3 berkisar -0.184303 sampai 0.199935. Sedangkan pada variabel keluhan kesehatan memiliki nilai penduga yang negatif dan positif, dengan besaran pengaruh dari -0.184303 sampai 0.199935.

Berikut sampel model GWR yang terbentuk dengan menggunakan pembobot kernel fixed bi-square :

$$Y_{Jakarta} = 49.20784 + -0.71057268X1 + 0.13298866X2 + -0.016613989X3.$$

2. Adaptive Bi-Square

Pembobot pada adaptif bisquare memiliki langkah dan formulasi yang sama dengan fixed bi-square, namun memiliki perbedaan pada bandwidth dimana memiliki bandwidth sebanyak lokasi pengamatan. Berikut penulis sajikan nilai penduga parameter :

TABEL 3. Sebaran Estimasi Adaptive Bi-Square

	Min	Median	Max
Intercept	15.319670	36.101383	89.437373
X1	-1.398103	-0.482105	-0.188436
X2	-0.049349	0.080284	0.206029
X3	-0.266926	-0.058491	0.468898

Nilai penduga parameter pada angkatan kerja maupun persentase miskin selalu bernilai negatif dengan besaran pengaruh masing masing berkisar -1.398103 sampai -0.482105 dan -0.266926 sampai -0.058491. Sedangkan pada variabel keluhan kesehatan cukup bervariasi nilai penduga parameter nya dari nilai negatif sampai positif, dengan besaran pengaruh sebesar -0.049349 sampai 0.080284 .

Berikut sampel model GWR yang terbentuk dengan adaptive bi-square dari suatu provinsi :

$$Y_{Jakarta} = 87.30420 + -1.3357518X1 + 0.13494970X2 + 0.42269128X3$$

3. Fixed Gaussian

Pada kernel *Gaussian fixed* maka nilai bandwidth yang diperoleh hanya satu nilai yang optimum yaitu 5.334214. Estimasi parameter regresi untuk setiap wilayah, berikut nilai sebaran dari penduga parameter.

TABEL 4. Sebaran Estimasi Fixed Gaussian

	Min	Median	Max
Intercept	19.78295	36.186340	56.425761
X1	-0.823326	-0.473132	-0.210060
X2	-0.149969	0.090014	0.149627
X3	-0.193790	-0.055773	0.200615

Tabel menunjukkan bahwa variabel angkatan kerja (X1) memiliki koefisien yang selalu negatif dengan rentang -0.823326 sampai -0.210060. Sementara pada koefisien keluhan kesehatan terdapat nilai negatif dan positif. Besarnya pengaruh keluhan kesehatan terhadap pengangguran terbuka sekitar interval -0.149969 sampai 0.149627. Hal ini terjadi juga pada koefisien persentase penduduk miskin bernilai negatif dan positif, dengan besar pengaruh berkisar -0.193790 sampai 0.200615.

Setiap lokasi menghasilkan model yang berbeda, berikut sampel model GWR pada suatu provinsi :

$$Y_{Jakarta} = 54.78055 - 0.7976142X_1 + 0.132989749X_2 + 0.026983833X_3.$$

4. Adaptive Gaussian

Pada kernel Adaptif Gaussian nilai bandwidth yang diperoleh hanya Sebanyak lokasi pengamatan yaitu dalam penelitian ini 34. Berikut nilai sebaran dari penduga parameter :

TABEL 5. Sebaran Estimasi Adaptive Gaussian

	Min	Median	Max
Intercept	19.320341	31.284880	82.539618
X1	-1.268195	-0.396746	-0.239325
X2	0.023952	0.091884	0.152472
X3	-0.174613	-0.055882	0.369537

Tabel di atas menunjukkan bahwa variabel angkatan kerja (X1) memiliki koefisien regresi yang selalu negatif dengan rentang -1.268195 sampai -0.239325. Terjadi juga pada koefisien persentase penduduk miskin bernilai negatif, dengan besar pengaruh berkisar antara -0.174613 sampai -0.174613. Sementara kernel adaptif Gaussian pada koefisien keluhan kesehatan bernilai positif. Besarnya pengaruh keluhan kesehatan terhadap pengangguran terbuka sekitar interval 0.023952 sampai 0.152472 .

Berikut sampel model GWR yang terbentuk dari suatu provinsi :

$$Y_{Jakarta} = 71.66661 + -1.0666241X_1 + 0.12632881X_2 + 0.205416864X_3.$$

Uji hipotesis Model GWR

Pengujian ini dilakukan untuk mengetahui apakah model layak dan mengandung faktor geografis pada model. Pengujian ini terbagi dua uji yaitu dengan uji F dan Uji T pada GWR. Pertama dilakukan uji F, berikut hasil uji untuk setiap pembobot :

TABEL 6. Uji Kelayakan Model

Jenis Pembobot	F_{hit}	F_{tabel}	P-Value	Kesimpulan
Fixed Gaussian	2.6929	2.107143	0.01477	Tolak H_0
Adaptive Gaussian	2.7185	2.107143282	0.01435	Tolak H_0
Fixed Bi-square	2.132	2.010248	0.03643	Tolak H_0
Adaptive Bi-square	4.9882	2.380334	0.001468	Tolak H_0

Berdasarkan table 6 semua pembobot menunjukkan bahwa $F_{hit} > F_{table}$ maka H_0 ditolak artinya model-model yang dibentuk memiliki pengaruh geografis. Terdapat perbedaan signifikan antara model regresi global dengan model GWR pada penganalisisan menggunakan selang kepercayaan 95%.

Berikut disajikan contoh hasil uji dengan dua kernel pembobot.

1. Fixed Gaussian

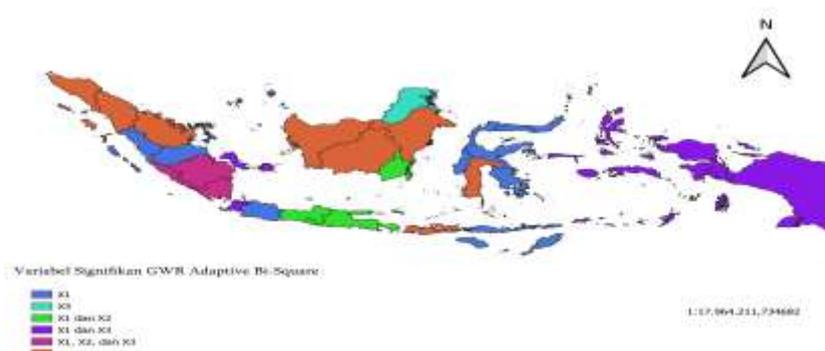


GAMBAR 5. Peta Kelompok Wilayah Berdasarkan Faktor yang Signifikan menggunakan Fixed Gaussian

Hasil parameter signifikan pada *Fixed Gaussian* terbagi menjadi 3 kelompok parameter. Terdapat 3 provinsi yang tidak memiliki pengaruh signifikan yaitu Aceh, Sumatera Utara, dan

Papua. Sedikitnya variasi dari klasifikasi signifikan, menunjukkan model GWR yang kurang identik dengan data spasial.

2. Adaptive Bi-Square



GAMBAR 6. Peta Kelompok Wilayah Berdasarkan Faktor yang Signifikan Menggunakan Adaptive Bi-Square

Variabel-variabel independen signifikan terhadap variabel TPT dengan adaptif bi-square terlihat bahwa model GWR terbagi ke dalam 5 kelompok sebaran model dan terdapat 8 provinsi yang tidak memiliki kesignifikan. Jika dibandingkan dengan sebelumnya pada fixed gaussian, maka kelompok model signifikan adaptif bi square lebih bervariasi.

Perbandingan Model GWR dengan Regresi Linear Berganda

Metode untuk mengetahui model yang terbaik antara model regresi global dan regresi spasial dapat menggunakan nilai koefisien determinasi (R^2) dan nilai *Akaike Information Criterion* (AIC). Hasil dari SSE/JKG bisa jadi pertimbangan juga dalam menentukan model terbaik. Penelitian ini terdiri dari 4 model GWR dengan pembobot berbeda dan 1 model regresi global.

TABEL 7. Nilai (R^2), AIC dan JKG

Model	R^2	AIC	JKG
Regresi Linear Berganda	0.461	131.9619	71.92322
Fixed Gaussian	0.7998555	101.2162	26.70827
Adaptive Gaussian	0.8017365	100.7926	26.45726
Fixed Bi-square	0.7471976	106.9829	33.73520
Adaptive Bi-square	0.8919505	84.61524	14.41866

Dari Tabel 4.18 terlihat bahwa pemodelan tingkat pengangguran terbuka di Indonesia dengan metode GWR lebih baik dari regresi linear berganda, dimana nilai R^2 pada model GWR lebih besar. Nilai R^2 yang terkecil terdapat di model regresi linear berganda dan R^2 yang terbesar dimiliki model adaptif bi-square dengan 0.8919505. Artinya model dengan GWR adaptif bisquare mampu menjelaskan pengaruh tingkat pengangguran terbuka terhadap faktor-faktor nya sebesar 89.1% dan sisanya dijelaskan oleh faktor lainnya yang tidak terdapat pada penelitian ini.

No	Provinsi	Model GWR
1	Aceh	$\hat{y} = 22.74388 - 0.2484283 X_1 + 0.04317026 X_2 - 0.16025001 X_3$
2	Sumatera Utara	$\hat{y} = 22.73267 - 0.2534034 X_1 + 0.04601183 X_2 - 0.14579995 X_3$
3	Sumatera Barat	$\hat{y} = 31.38618 - 0.3936257 X_1 + 0.05539786 X_2 - 0.10049762 X_3$
4	Riau	$\hat{y} = 26.29688 - 0.3176964 X_1 + 0.05429061 X_2 - 0.10837662 X_3$
5	Jambi	$\hat{y} = 51.81006 - 0.7804129 X_1 + 0.13522826 X_2 + 0.15376694 X_3$
6	Sumatera Selatan	$\hat{y} = 78.04460 - 1.2484397 X_1 + 0.20602924 X_2 + 0.41725232 X_3$
7	Bengkulu	$\hat{y} = 71.81507 - 1.1395849 X_1 + 0.19470670 X_2 + 0.35290602 X_3$
8	Lampung	$\hat{y} = 89.24538 - 1.3981035 X_1 + 0.17214502 X_2 + 0.46889757 X_3$
9	Kepulauan Bangka Belitung	$\hat{y} = 89.43737 - 1.3703735 X_1 + 0.13009045 X_2 + 0.44123459 X_3$
10	Kepulauan Riau	$\hat{y} = 58.25502 - 0.8139051 X_1 + 0.09698972 X_2 - 0.04763309 X_3$
11	DKI Jakarta	$\hat{y} = 87.30420 - 1.3357518 X_1 + 0.13494970 X_2 + 0.42269128 X_3$
12	Jawa Barat	$\hat{y} = 81.84833 - 1.2341937 X_1 + 0.12191480 X_2 + 0.34981671 X_3$
13	Jawa Tengah	$\hat{y} = 54.36831 - 0.7761690 X_1 + 0.13566691 X_2 - 0.06219566 X_3$
14	DI Yogyakarta	$\hat{y} = 53.02686 - 0.7526600 X_1 + 0.13566691 X_2 - 0.09747241 X_3$
15	Jawa Timur	$\hat{y} = 52.11825 - 0.7502448 X_1 + 0.15763620 X_2 - 0.11657477 X_3$
16	Banten	$\hat{y} = 88.73208 - 1.3659938 X_1 + 0.14424832 X_2 + 0.43858775 X_3$
17	Bali	$\hat{y} = 30.77764 - 0.4317816 X_1 + 0.12013226 X_2 - 0.06923491 X_3$
18	Nusa Tenggara Barat	$\hat{y} = 19.41945 - 0.2514782 X_1 + 0.08542788 X_2 - 0.06087760 X_3$
19	Nusa Tenggara Timur	$\hat{y} = 17.09164 - 0.2130726 X_1 + 0.07336279 X_2 - 0.05610348 X_3$
20	Kalimantan Barat	$\hat{y} = 33.93901 - 0.4648884 X_1 + 0.07513934 X_2 + 0.08940894 X_3$
21	Kalimantan Tengah	$\hat{y} = 15.31967 - 0.2427081 X_1 + 0.08809714 X_2 - 0.07751250 X_3$
22	Kalimantan Selatan	$\hat{y} = 22.23257 - 0.2844167 X_1 + 0.09090521 X_2 - 0.16606547 X_3$
23	Kalimantan Timur	$\hat{y} = 24.21718 - 0.2950676 X_1 + 0.06622593 X_2 - 0.21343800 X_3$
24	Kalimantan Utara	$\hat{y} = 19.85785 - 0.2427081 X_1 + 0.10027643 X_2 - 0.26692622 X_3$
25	Sulawesi Utara	$\hat{y} = 39.84583 - 0.4993217 X_1 - 0.4993217 X_2 + 0.03351404 X_3$
26	Sulawesi Tengah	$\hat{y} = 25.31618 - 0.2972933 X_1 + 0.02734510 X_2 - 0.13663053 X_3$
27	Sulawesi Selatan	$\hat{y} = 19.16758 - 0.2322670 X_1 + 0.06253639 X_2 - 0.09834028 X_3$
28	Sulawesi Tenggara	$\hat{y} = 21.28843 - 0.2616911 X_1 + 0.04692364 X_2 - 0.04920628 X_3$
29	Gorontalo	$\hat{y} = 33.92663 - 0.4109002 X_1 - 0.01410400 X_2 - 0.07178596 X_3$
30	Sulawesi Barat	$\hat{y} = 21.39798 - 0.2582015 X_1 + 0.05435806 X_2 - 0.11777837 X_3$
31	Maluku	$\hat{y} = 39.03120 - 0.5170755 X_1 - 0.03183004 X_2 + 0.19027366 X_3$
32	Maluku Utara	$\hat{y} = 38.26376 - 0.5137835 X_1 - 0.01848780 X_2 + 0.18424484 X_3$
33	Papua Barat	$\hat{y} = 41.36783 - 0.5456754 X_1 - 0.04301365 X_2 + 0.19410545 X_3$
34	Papua	$\hat{y} = 41.42779 - 0.5403669 X_1 - 0.04934939 X_2 + 0.18600791 X_3$

GAMBAR 7. Hasil Model GWR dengan Adaptive Bi-Square

KESIMPULAN

- Model Regresi Linear Berganda yang dihasilkan pada kasus pengangguran terbuka :

$$\hat{Y} = 29.760 - 0.395X_1 + 0.085X_2$$

Memiliki 2 faktor signifikan (angkatan kerja dan Keluhan Kesehatan) dari 3 faktor. Hasil uji asumsi klasik menunjukkan bahwa model regresi global memenuhi asumsi normalitas, non multikolinearitas, dan non autokorelasi, sedangkan model tidak memenuhi asumsi kehomogenan.
- Data yang digunakan merupakan data yang mengandung aspek spasial. Pengamatan menggunakan metode *Geographically Weighted Regression* menggunakan 4 pembobot berbeda yaitu *Fixed Gaussian*, *Adaptive Gaussian*, *Fixed Bi-Square*, dan *Adaptive Bi-Square* dimana menghasilkan kebaikan model yaitu memiliki pengaruh geografis menunjukkan perbedaan signifikan dengan regresi global.
- Model dengan metode *Geographically Weighted Regression* adalah model terbaik dalam penelitian ini jika dibanding dengan Regresi linear berganda dengan metode OLS (*Ordinary Least Square*). Ditinjau dari nilai R^2 , AIC dan Jumlah Kuadrat Galat (JKG) menunjukkan

hasil GWR kernel *Adaptive Bi-Square* merupakan model terbaik.

REFERENSI

- Badan Pusat Statistik. 2020. "Tenaga Kerja". Tersedia: <https://www.bps.go.id/subject/6/tenaga-kerja.html>. Diakses 26 Mei 2020.
- Badan Pusat Statistik. 2021. *Statistik Indonesia 2021*. Jakarta: BPS.
- Bekti, Rokhana Dwi 2012, 'Autokorelasi Spasial untuk Identifikasi Pola Hubungan Kemiskinan di Jawa Timur', ComTech. vol. 3, No.1. 217-227.
- Brundson C., Fotheringham, A.S., dan Charlton, M. E. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*. England : John Wiley and Sons Ltd.
- Caraka, Rezzy Eko dkk. 2017, "Geographically Weighted Regression (GWR) Sebuah Pendekatan Regresi Geografis. Edisi Pertama. Cet.Ke-1. Yogyakarta : Mobius.
- Ira, Farida. 2016. *Model Geographically Weighted Regression (GWR) dengan Pembobot Kernel Bi-Square*. Repository.upi.edu. 15-24.
- Lutfiani, Nurul.dkk. 2019. *Pemodelan Geographically Weighted Regression dengan Fungsi Pembobot Kernel Gaussian dan Bi-Square*. Unnes Journal of Mathematics. 8(1). 82-91.
- Nadya, Meila.dkk. 2017. *Analisis Geographically Weighted Regression pada Kasus Pneumonia Balita di Jawa Barat*. Jurnal Statistika dan Aplikasinya. Vol.1, No.1. 23-32.