

Pendekatan *Geographically Weighted Regression* (GWR) untuk Menganalisis Hubungan PDRB Sektor Pertanian, Kehutanan, dan Perikanan dengan Faktor Pencemaran Lingkungan di Jawa Timur

Nurul Aulya Bakri^{1*}, Suwardi Annas¹, Muhammad Kasim Aidid¹

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Keywords: GRDP, Spatial, Environment, Geographically Weighted Regression

Abstract:

The Geographically Weighted Regression (GWR) method is a method used to analyze spatial heterogeneity, where the same independent variable gives unequal responses at different locations in a research area. The purpose of this study was to determine the environmental pollution factors that affect GRDP in the agricultural, forestry and fisheries sectors in East Java. The data used in this study are the GRDP of the Agriculture, Forestry and Fisheries sectors in East Java in 2020 along with the environmental pollution factors that are thought to influence it. The results of this study obtained a different model for each district/city. The GWR model shows better results than the multiple linear regression model, as seen from the smallest AIC value and the largest R^2 .

1. Pendahuluan

Produk Domestik Regional Bruto (PDRB) merupakan salah satu indikator penting untuk mengukur kondisi ekonomi suatu daerah dalam suatu periode tertentu. Pertumbuhan ekonomi menggambarkan keberhasilan pembangunan. Jadi semakin tinggi pertumbuhan ekonomi, maka semakin tinggi pula kesejahteraan masyarakat (Suardin dkk., 2020).

PDRB pada dasarnya adalah jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha dalam suatu daerah tertentu. Berdasarkan data publikasi Badan Pusat Statistik (BPS), PDRB Indonesia terus mengalami pertumbuhan dari tahun ke tahun dengan Provinsi Jawa Timur sebagai penyumbang terbesar sektor Pertanian, Kehutanan, dan Perikanan. Salah satu pendekatan yang digunakan dalam menghitung PDRB adalah pendekatan produksi yang berkaitan dengan penggunaan sumber daya alam. Sehingga, besaran PDRB yang dihasilkan suatu daerah sangat bergantung pada pengelolaan sumber daya alam dan faktor produksinya. Namun, saat ini sumber daya alam terus mengalami penurunan akibat penggunaan berlebihan dengan tidak menggunakan prinsip optimal dan lestari. Selain itu faktor pencemaran lingkungan juga berpengaruh terhadap ketersediaan sumber daya alam. Daya dukung alam yang menurun mengakibatkan pertumbuhan perekonomian masyarakat terganggu kestabilannya.

Beberapa penelitian terkait PDRB dengan pencemaran lingkungan, diantaranya penelitian untuk menganalisis pengaruh Produk Domestik Regional Bruto (PDRB) sektor pertanian, industri, dan transportasi terhadap emisi karbondioksida di Provinsi Jawa Timur (Faris, 2022). Hasilnya menunjukkan bahwa variabel PDRB sektor pertanian, industri dan transportasi berpengaruh terhadap emisi karbondioksida. Sektor pertanian dan transportasi berpengaruh positif, sedangkan sektor industri berpengaruh negatif terhadap emisi karbondioksida. Penelitian lainnya yaitu untuk

* Corresponding author.

E-mail address: nurulauliyabakri@gmail.com



menentukan dinamika perubahan tutupan lahan dan hutan serta dampaknya terhadap pendapatan di sektor pertanian, kehutanan dan industri (Dienelly dkk., 2017). Namun, penelitian-penelitian tersebut tidak melihat pengaruh spasial.

Berdasarkan hal tersebut, penulis akan melakukan penelitian dengan metode GWR untuk memodelkan PDRB sektor Pertanian, Kehutanan, dan Perikanan di Jawa Timur serta mengetahui faktor pencemaran lingkungan yang berpengaruh secara signifikan di setiap kabupaten/kota.

2. Tinjauan Pustaka

2.1 Regresi Linear

Analisis regresi linear digunakan untuk memprediksi seberapa jauh perubahan nilai variabel dependen apabila nilai variabel independen berubah (Sugiyono, 2007). Bentuk umum model regresi linear untuk k variabel independen sebagai berikut (Tiro, 2010).

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

dimana $\beta_0, \beta_1, \dots, \beta_k$ merupakan parameter model dan ε merupakan error yang diasumsikan identik, independen, dan berdistribusi normal dengan mean nol dan varians konstan σ^2 atau ($\varepsilon_i \sim IIDN(0, \sigma^2)$).

2.2 Regresi Spasial

Regresi spasial merupakan pengembangan dari metode regresi linear berganda. Pengembangan ini didasari adanya pengaruh tempat atau spasial (*spatial effect*) dari data yang dianalisis (Nasir dkk., 2020). Efek lokasi (*spatial effect*) terdiri dari dependensi spasial dan heterogenitas spasial. Dependensi spasial dapat diartikan bahwa pengamatan di lokasi i bergantung pada pengamatan lain di lokasi j , $j \neq i$. Sedangkan heterogenitas spasial terjadi akibat efek lokasi acak, yaitu adanya perbedaan antara satu lokasi dengan lokasi yang lainnya (Yasin dkk., 2020).

Pemodelan regresi spasial dapat dibedakan menjadi pemodelan dengan pendekatan titik dan pendekatan area. Pemodelan spasial dengan pendekatan titik digunakan untuk menyelesaikan kasus yang mengandung heterogenitas spasial, sedangkan pendekatan area digunakan untuk menyelesaikan kasus yang mengandung dependensi spasial. Pada penelitian ini yang digunakan adalah pemodelan regresi spasial dengan pendekatan titik.

2.3 Uji Asumsi Klasik

2.3.1 Normalitas

Uji *Anderson-Darling* adalah uji kecocokan umum yang menguji apakah sampel berasal dari distribusi yang ditentukan. Metode ini menguji hipotesis bahwa sampel telah diambil dari populasi dengan fungsi distribusi kontinu $F(x)$ yang ditentukan (Anderson & Darling, 1954). Statistik uji Anderson-Darling didefinisikan sebagai berikut:

$$W_n^2 = -n - \frac{1}{n} \sum_{j=1}^n (2j-1) [\ln u_j + \ln(1 - u_{n-j+1})]$$

dengan $u_j = F(x_{(j)})$. Daerah kritis terima H_0 , jika nilai $W_n^2 \leq W_{n, \alpha}^2$ yang berarti data berdistribusi normal dengan α sebagai tingkat signifikansi.

2.3.2 Multikolinearitas

Uji multikolinearitas merupakan pengujian asumsi yang digunakan untuk melihat keeratan hubungan antara variabel independen. Pendeteksian kasus multikolinearitas dilakukan menggunakan kriteria nilai VIF. Jika nilai VIF (*Variance Inflation Factor*) lebih besar dari 10 menunjukkan adanya multikolinearitas antar variabel independen. Nilai VIF dinyatakan sebagai berikut (Hocking, 2003).

$$VIF = \frac{1}{1 - R_j^2}$$

dimana R_j^2 adalah koefisien determinasi antara satu variabel independen (X_j) dengan variabel independen lainnya.

2.3.3 Autokorelasi

Autokorelasi merupakan kondisi dimana pada model regresi ada korelasi antara residual pada periode t dengan residual pada periode sebelumnya ($t-1$). Model regresi yang baik adalah yang tidak adanya autokorelasi. Uji statistik

untuk mendeteksi autokorelasi adalah uji Durbin Watson (Trenggonowati dkk., 2020). Statistik uji Durbin Watson didefinisikan sebagai berikut.

$$DW = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2}$$

Daerah kritis tolak H_0 , jika $p\text{-value} < \alpha$ atau $d < D_U$ atau $d > 4 - D_U$.

2.4 Pengujian Aspek Spasial

2.4.1 Dependensi Spasial

Uji Moran's I merupakan metode pengujian yang digunakan untuk dependensi spasial (Lee & Wong, 2001). Statistik uji sebagai berikut.

$$Z_I = \frac{I - E(I)}{\sqrt{\text{Var}(I)}}$$

dimana, I : indeks Moran's I, Z_I : nilai statistik uji indeks Moran's I, $E(I)$: nilai ekspektasi indeks Moran's I, dan $\text{Var}(I)$: nilai variansi dari indeks Moran's I. Daerah kritis tolak H_0 jika $|Z_I \text{ hitung}| > Z_{\alpha/2}$ atau $p\text{-value} < \alpha$ yang berarti bahwa terdapat dependensi spasial antar lokasi dengan α adalah tingkat signifikansi.

2.4.2 Heterogenitas Spasial

Uji heterogenitas spasial merupakan pengujian untuk mengetahui apakah terdapat karakteristik atau keunikan sendiri di setiap lokasi pengamatan. Pengujian heterogenitas spasial menggunakan statistik uji *Breusch-Pagan* dengan hipotesis sebagai berikut (Anselin, 1988). Statistik uji:

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

dimana, $e_i = y_i - \hat{y}_i$ adalah *least square residual* untuk pengamatan ke- i dan \mathbf{Z} merupakan matriks berukuran $(n \times (p + 1))$ berisi vektor yang sudah dinormalisasi untuk tiap pengamatan. Daerah kritis tolak H_0 jika $BP > \chi_p^2$ atau $p\text{-value} < \alpha$ dengan p adalah banyaknya variabel independen.

2.5 Geographically Weighted Regression (GWR)

Geographically Weighted Regression (GWR) merupakan metode statistika yang digunakan untuk menganalisis heterogenitas spasial (Fotheringham dkk., 2002). Heterogenitas spasial adalah kondisi dimana satu variabel independen yang sama memberikan respon yang tidak sama pada lokasi yang berbeda dalam satu wilayah penelitian. Pada model GWR, variabel dependen ditaksir dengan variabel independen yang setiap koefisien regresinya tergantung pada lokasi dimana data tersebut diamati (Caraka & Yasin, 2017). Model GWR dapat ditulis sebagai berikut.

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

dimana, y_i merupakan nilai observasi variabel dependen ke- i , $\beta_0(u_i, v_i)$ adalah konstanta/*intercept* pada pengamatan ke- i , $\beta_k(u_i, v_i)$ ialah nilai observasi variabel independen ke- k pada lokasi ke- i , x_{ik} adalah nilai observasi variabel independen ke- k pada lokasi ke- i , dan ε_i adalah error pengamatan ke- i yang diasumsikan identik, independen dan berdistribusi normal dengan mean nol dan varian konstan σ^2 .

2.5.1 Pendugaan Parameter

Pendugaan parameter pada model GWR menggunakan metode *Weighted Least Square* (WLS) yaitu dengan memberikan pembobot yang berbeda untuk setiap lokasi dimana data diamati. Bentuk penduga parameter dari model GWR untuk setiap lokasi adalah:

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

2.5.2 Pembobot Spasial

Proses pendugaan parameter model GWR di titik (u_i, v_i) membutuhkan pembobot spasial dimana pembobot yang digunakan pada penelitian ini adalah fungsi kernel gaussian sebagai berikut.

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2} \left(\frac{d_{ij}}{h}\right)^2\right)$$

dimana, d_{ij} adalah jarak Euclidean antara lokasi (u_i, v_i) ke lokasi (u_j, v_j) dan h adalah nilai parameter non negatif yang biasanya disebut parameter penghalus (*bandwidth*) (Lee & Wong, 2001). Ada beberapa metode yang dapat digunakan untuk memilih *bandwidth* optimum, salah satunya adalah metode *Cross Validation* (CV) yang didefinisikan sebagai berikut.

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

dengan $\hat{y}_{\neq i}(h)$ adalah nilai penduga y_i dimana pengamatan di lokasi (u_i, v_i) dihilangkan dari proses estimasi. Untuk mendapatkan nilai *bandwidth* (h) yang optimal maka diperoleh dari h yang menghasilkan nilai CV yang minimum (Caraka & Yasin, 2017).

2.5.3 Uji Hipotesis Model GWR

Uji hipotesis model GWR terdiri dari 2 jenis sebagai berikut (Caraka & Yasin, 2017).

1. Uji kesesuaian model (*Goodness of Fit*)

Pengujian kesesuaian model dilakukan dengan hipotesis sebagai berikut:

$H_0 : \beta_k(u_i, v_i) = \beta_k, k = 1, 2, \dots, p$ (tidak ada perbedaan yang signifikan antara model regresi linear dengan GWR)

$H_1 : \text{paling sedikit ada satu } \beta_k(u_i, v_i) \neq \beta_k$ (ada perbedaan yang signifikan antara model regresi linear dan GWR)

Statistik uji:

$$F^* = \frac{SSE(H_0)/df_1}{SSE(H_1)/df_2}$$

dengan:

$$SSE(H_0) = \mathbf{Y}^T(\mathbf{I} - \mathbf{H})\mathbf{Y} \text{ dimana } \mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$$

$$df_1 = n - p - 1$$

$$SSE(H_1) = \mathbf{Y}^T(\mathbf{I} - \mathbf{S})^T(\mathbf{I} - \mathbf{S})\mathbf{Y}$$

$$df_2 = (n - 2tr(\mathbf{S}) + tr(\mathbf{S}^T\mathbf{S}))$$

\mathbf{S} adalah matriks $n \times n$ yang memproyeksikan model GWR dan \mathbf{I} adalah matriks identitas ordo n . Jika diberikan tingkat signifikansi sebesar α , maka H_0 ditolak ($F^* > F_{\alpha, df_1, df_2}$). Artinya model GWR mempunyai *goodness of fit* yang lebih baik daripada model regresi linear.

2. Uji parameter model

Pengujian parameter model dilakukan untuk mengetahui parameter mana saja yang signifikan terhadap variabel dependen. Hipotesisnya sebagai berikut.

$$H_0 : \beta_k(u_i, v_i) = 0$$

$$H_1 : \beta_k(u_i, v_i) \neq 0 ; k = 1, 2, \dots, p$$

Statistik uji:

$$T = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma} \sqrt{g_{kk}}}$$

dengan g_{kk} adalah elemen diagonal ke- k dari matriks $\mathbf{G}\mathbf{G}^T$ dimana $\mathbf{G} = (\mathbf{X}^T\mathbf{W}(u_i, v_i)\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}(u_i, v_i)$ sehingga didapatkan $\frac{\hat{\beta}_k(u_i, v_i) - \beta_k(u_i, v_i)}{\sigma \sqrt{g_{kk}}} \sim N(0, 1)$

Daerah kritis tolak H_0 jika $|T_{hit}| > t_{\frac{\alpha}{2}, df_2}$ yang berarti bahwa parameter $\beta_k(u_i, v_i)$ signifikan terhadap model dengan α adalah tingkat signifikansi.

2.6 Pemilihan Model Terbaik

2.6.1 Akaike Information Criterion (AIC)

Model terbaik ditentukan dengan memiliki nilai AIC yang terkecil (Fotheringham dkk., 2002). Rumus AIC seperti berikut ini.

$$AIC = e^{\frac{2k}{n} \frac{\sum_{i=1}^n \hat{u}_i^2}{n}}$$

dengan k jumlah parameter yang diestimasi dalam model regresi, n jumlah pengamatan, e yang bernilai 2,718, dan u adalah residual.

2.6.2 Koefisien determinasi (R^2)

Kriteria ini digunakan sebagai informasi kecocokan suatu model yang berkisar antara 0 sampai dengan 1. Nilai R^2 digunakan untuk melihat seberapa besar pengaruh yang diberikan variabel dependen terhadap variabel independen. Rumusnya dituliskan seperti berikut (Ndruru, E, M.Situmorang, 2014).

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \times 100\%$$

dengan:

y_i : data aktual pada pengamatan ke- i

\hat{y} : hasil prediksi pada pengamatan ke- i

\bar{y} : nilai rata-rata variabel dependen

n : jumlah pengamatan

2.7 Produk Domestik Regional Bruto (PDRB)

Produk Domestik Regional Bruto (PDRB) adalah nilai barang dan jasa yang dihasilkan oleh seluruh unit ekonomi di suatu wilayah. PDRB atas dasar harga konstan menunjukkan nilai tambah barang dan jasa yang dihitung berdasarkan harga yang berlaku pada satu tahun tertentu. PDRB konstan ini digunakan untuk mengetahui pertumbuhan ekonomi secara riil dari tahun ke tahun atau pertumbuhan ekonomi yang tidak dipengaruhi oleh faktor harga.

Perhitungan PDRB menggunakan tiga jenis pendekatan, yaitu pendekatan produksi, pengeluaran, dan pendapatan. Mengenai pemanfaatan sumber daya alam, pendekatan yang digunakan adalah pendekatan produksi. Pada penelitian ini akan berfokus pada unit produksi pertanian, kehutanan, dan perikanan.

3. Metode Penelitian

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur yaitu PDRB sektor Pertanian, Kehutanan, dan Perikanan atas dasar harga konstan tahun 2020 dan faktor-faktor pencemaran lingkungan yang diperoleh dari Kementerian Lingkungan Hidup & Kehutanan. Unit penelitian ini adalah 38 kabupaten/kota di Jawa Timur.

3.2 Teknik Analisis Data

Pada penelitian ini, pengujian data PDRB sektor pertanian, kehutanan, dan perikanan dengan factor pencemaran lingkungan di Jawa Timur pada tahun 2020 menggunakan teknik analisis linear berganda dan *Geographically Weighted Regression* (GWR). Aplikasi pengolahan data yang digunakan yaitu R-Studio versi 4.3.0.

4. Hasil dan Pembahasan

4.1. Statistika Deskriptif

Sebelum analisis lebih lanjut, data PDRB (Pertanian, Kehutanan, dan Perikanan) dan faktor-faktor pencemaran lingkungan akan dideskripsikan dengan statistika deskriptif. Statistik deskriptif dari setiap variabel dalam penelitian dapat dilihat pada Tabel 1.

Tabel 1 Statistika Deskriptif Tiap Variabel

Statistika Deskriptif	Y	X ₁	X ₂	X ₃	X ₄	X ₅
Minimum	27.210	39,33	50,00	25,66	295	23.790
Median	3.971.162	58,87	84,47	55,99	945	130.111
Rata-rata	4.647.731	60,33	83,37	59,02	1.923	174.749
Maksimum	14.950.320	84,66	98,22	102,76	8.200	811.255

4.2 Regresi Linear Berganda

Analisis regresi linear berganda dilakukan untuk melihat hubungan antara PDRB (Pertanian, Kehutanan, dan Perikanan) di Jawa Timur dengan faktor-faktor pencemaran lingkungan tanpa melihat pengaruh spasial. Nilai estimasi parameter dapat dilihat pada Tabel 2.

Tabel 2 Nilai Estimasi Parameter Model Regresi Linear Berganda

Variabel Independen	Estimasi	Std.Error	t value	Pr(> t)	Keterangan
Intersep	-1.455.000	5.741.000	-0,253	0,8015090	Tidak Signifikan
X ₁	41.150	41.310	0,996	0,3266970	Tidak Signifikan
X ₂	40.600	48.830	0,831	0,4118930	Tidak Signifikan
X ₃	14.910	23.480	0,635	0,5299980	Tidak Signifikan
X ₄	-1.178	202,9	-5,806	0,0000019	Signifikan
X ₅	9,273	2,413	3,843	0,0005440	Signifikan

Model regresi linear yang terbentuk sebagai berikut.

$$\hat{Y} = -1.455.000 + 41.150X_1 + 40.600X_2 + 14.910X_3 - 1.178X_4 + 9,273X_5$$

Menentukan variabel yang signifikan terhadap PDRB (Pertanian, Kehutanan, dan Perikanan) dapat dilakukan menggunakan metode *forward selection*. Setelah didapatkan model signifikan terbaik, maka dibentuk kembali persamaan regresinya. Sehingga model regresi linear yang terbentuk menjadi seperti berikut.

$$\hat{Y} = 5.276.000 - 1.192X_4 + 9,517X_5$$

Berdasarkan hasil statistik uji pada Tabel 2, dengan taraf signifikan 10% didapatkan bahwa faktor pencemaran lingkungan yang berpengaruh terhadap PDRB (Pertanian, Kehutanan, dan Perikanan) adalah kepadatan penduduk (X₄) dan timbulan sampah (X₅). Hal ini dilihat dari nilai Pr(> | t |) yang lebih kecil dari 0,1. Kemudian, hasil pengujian secara simultan untuk melihat kecocokan model regresi linear berganda dapat dilihat pada Tabel 3.

Tabel 3 Hasil Pengujian Simultan Model Regresi Linear Berganda

R ²	F	DF ₁	DF ₂	P - value
0,6317	10,98	5	32	0,000003261

Berdasarkan Tabel 3, menunjukkan bahwa semua variabel independen berpengaruh secara simultan terhadap variabel dependen yaitu PDRB (Pertanian, Kehutanan, dan Perikanan). Hal ini dilihat dari nilai $F_{hitung}(10,98) > F_{tabel}(2,51)$ dengan p -value 0,000003261. Kemudian, diperoleh juga nilai $R^2 = 0,6317$ yang artinya sebesar 63,17% PDRB (Pertanian, Kehutanan, dan Perikanan) dapat dijelaskan dengan kepadatan penduduk dan timbulan sampah. Sedangkan 36,83% sisanya dijelaskan oleh faktor-faktor lain yang tidak diamati.

Selanjutnya, pengujian asumsi klasik untuk regresi linear berganda sebagai berikut.

4.2.1 Normalitas

Hasil pengujian didapatkan nilai A = 0,481 dan p -value = 0,219 yang menunjukkan bahwa residual data menyebar secara normal karena p -value = 0,219 > $\alpha = 0,1$.

4.2.2 Multikolinearitas

Nilai VIF dari variabel-variabel independen dapat dilihat pada tabel berikut.

Tabel 4 Nilai VIF Variabel Independen

	X_1	X_2	X_3	X_4	X_5
VIF	1,141	1,210	1,272	1,271	1,024

Berdasarkan Tabel 4, terlihat bahwa semua nilai VIF dari variabel independen kurang dari 10 yang artinya tidak ada multikolinearitas antar variabel independen.

4.2.3 Autokorelasi

Penelitian ini menggunakan uji Durbin Watson. Hasil uji diperoleh nilai DW = 1,4756 dan $p\text{-value} = 0,02708$, menunjukkan bahwa ada autokorelasi pada residual karena nilai $p\text{-value} = 0,02708 < \alpha = 0,1$.

4.2. Pengujian Aspek Spasial

Hasil pengujian aspek spasial dapat dilihat pada tabel berikut.

Tabel 5 Nilai Moran's I dan Statistik Uji *Breusch-Pagan*

Pengujian	Statistik Uji	$P\text{-value}$
Moran's I	0,1013	0,1578
<i>Breusch-Pagan</i>	9,6488	0,0858

Berdasarkan Tabel 5, didapatkan bahwa nilai Moran's I sebesar 0,1013 dengan $p\text{-value} 0,1578 > \alpha = 0,1$ yang artinya tidak ada autokorelasi spasial pada data. Hal ini menjelaskan bahwa peristiwa atau keadaan di setiap kabupaten/kota di Jawa Timur tidak dipengaruhi oleh wilayah yang saling berdekatan. Kemudian, dari pengujian heterogenitas spasial diperoleh nilai statistik uji *Breusch-Pagan* sebesar 9,6488 dengan $p\text{-value} 0,0858 < \alpha = 0,1$. Hasil ini menunjukkan bahwa terdapat heterogenitas pada data penelitian, artinya ada karakteristik atau keunikan sendiri di setiap kabupaten/kota di Jawa Timur.

4.3. Analisis Geographically Weighted Regression (GWR)

Nilai minimum dan maksimum dari pemodelan GWR dapat dilihat pada Tabel 6.

Tabel 6 Nilai Minimum dan Maksimum Parameter Model GWR

Variabel Independen	Minimum	Maksimum
Intersep	-5.092.400	8.641.500
X_1	-138.310	51.589
X_2	-73.739	109.630
X_3	-38.893	41.117
X_4	-1232	-498,1
X_5	2,9963	30,166

Selanjutnya, dilakukan uji kesesuaian model GWR untuk mengetahui apakah ada perbedaan signifikan antara model regresi linear berganda dan model GWR. Hasil pengujian didapatkan nilai $F = 4,9554$ dengan $DF_1 = 32$ dan $DF_2 = 14,125$, serta $p\text{-value} = 0,001267$. Jadi, dapat disimpulkan ada perbedaan yang signifikan antara model regresi linear berganda dengan model GWR, ini dilihat dari $p\text{-value} = 0,001267 < \alpha = 0,1$.

Kemudian, pengujian parameter model GWR dilakukan untuk mengetahui parameter variabel independen apa saja yang signifikan terhadap variabel dependen di tiap kabupaten/kota di Provinsi Jawa Timur. Hasil pengujian diperoleh nilai t_{hitung} yang berbeda pada tiap kabupaten/kota, apabila nilai $t_{hitung} > t_{tabel(\frac{0,1}{2};32)} = 1,693$ maka diambil keputusan tolak H_0 yang artinya ada pengaruh variabel independen terhadap variabel dependen. Variabel yang signifikan disetiap kabupaten/kota dapat dilihat pada Tabel 7.

Tabel 7 Variabel yang Signifikan di Setiap Kabupaten/Kota

Kabupaten/Kota	Variabel Signifikan	Kabupaten/Kota	Variabel Signifikan
Pacitan	-	Magetan	X ₄ , X ₅
Ponorogo	X ₄ , X ₅	Ngawi	X ₄ , X ₅
Trenggalek	X ₄ , X ₅	Bojonegoro	X ₄ , X ₅
Tulungagung	X ₄ , X ₅	Tuban	X ₂ , X ₄
Blitar	X ₄ , X ₅	Lamongan	X ₂ , X ₄
Kediri	X ₄ , X ₅	Gresik	X ₂ , X ₄
Malang	X ₃ , X ₄ , X ₅	Bangkalan	X ₃ , X ₄ , X ₅
Lumajang	X ₃ , X ₄ , X ₅	Sampang	X ₄ , X ₅
Jember	X ₄ , X ₅	Pamekasan	X ₄ , X ₅
Banyuwangi	X ₅	Sumenep	X ₁ , X ₄ , X ₅
Bondowoso	X ₄ , X ₅	Kota Kediri	X ₄ , X ₅
Situbondo	X ₄ , X ₅	Kota Mojokerto	X ₂ , X ₄ , X ₅
Probolinggo	X ₃ , X ₄ , X ₅	Kota Blitar	X ₄ , X ₅
Pasuruan	X ₃ , X ₄ , X ₅	Kota Malang	X ₃ , X ₄ , X ₅
Sidoarjo	X ₃ , X ₄ , X ₅	Kota Probolinggo	X ₃ , X ₄ , X ₅
Mojokerto	X ₃ , X ₄ , X ₅	Kota Pasuruan	X ₃ , X ₄ , X ₅
Jombang	X ₂ , X ₄ , X ₅	Kota Madiun	X ₄ , X ₅
Nganjuk	X ₄ , X ₅	Kota Surabaya	X ₃ , X ₄ , X ₅
Madiun	X ₄ , X ₅	Kota Batu	X ₃ , X ₄ , X ₅

Model GWR yang terbentuk untuk tiap kabupaten/kota berbeda-beda. Sebagai contoh, berikut ini model GWR untuk Kota Surabaya.

$$\hat{Y}_{Surabaya} = -3.763.865,04 + 35.498,28X_3 - 1.003,13X_4 + 3,20X_5$$

Model diatas menunjukkan bahwa Indeks Kualitas Tutupan Lahan (X₃), Kepadatan Penduduk (X₄), dan Timbulan Sampah (X₅) berpengaruh terhadap PDRB (Pertanian, Kehutanan, dan Perikanan) di Kota Surabaya. Indeks Kualitas Tutupan Lahan (X₃) berpengaruh sebesar 35.498,28. Artinya, jika Indeks Kualitas Tutupan Lahan naik sebesar satu satuan, maka akan menaikkan PDRB (Pertanian, Kehutanan, dan Perikanan) sebesar 35.498,28. Kepadatan Penduduk (X₄) memiliki pengaruh sebesar 1.003,13. Artinya, jika kepadatan penduduk naik sebesar satu satuan, maka akan menurunkan PDRB (Pertanian, Kehutanan, dan Perikanan) sebesar 1.003,13. Timbulan Sampah (X₅) memiliki pengaruh sebesar 3,20. Artinya, jika timbulan sampah naik sebesar satu satuan, maka akan menaikkan PDRB (Pertanian, Kehutanan, dan Perikanan) sebesar 3,20.

4.4. Pemilihan Model Terbaik

Penentuan model terbaik ditentukan dengan kriteria nilai AIC dan R² yang ditampilkan pada Tabel 8.

Tabel 8 Nilai Kriteria AIC dan R²

Model	AIC	R ²
Regresi Linear Berganda	1228,82	63,17%
<i>Geographically Weighted Regression</i>	1173,30	92,56%

Berdasarkan kriteria AIC dan R² yang ditunjukkan pada Tabel 8 maka model terbaik yang digunakan untuk memodelkan PDRB (Pertanian, Kehutanan, dan Perikanan) dengan faktor pencemaran lingkungan Provinsi Jawa Timur tahun 2020 adalah model Geographically Weighted Regression (GWR), karena memiliki nilai AIC terkecil dan R² terbesar.

5. Kesimpulan

PDRB (Pertanian, Kehutanan, dan Perikanan) Jawa Timur tahun 2020 yang tertinggi di Kabupaten Banyuwangi dan terendah di Kota Mojokerto. Indeks Kualitas Air (IKA) tertinggi berada di Kabupaten Madiun dan indeks terendah berada di Kabupaten Bangkalan. Indeks Kualitas Udara (IKU) tertinggi pada Kabupaten Blitar dan indeks terendah berada di Kabupaten Madiun. Indeks Kualitas Tutupan Lahan (IKTL) tertinggi berada di Kota Batu dan Kabupaten Sampang yang terendah. Kota Surabaya mempunyai kepadatan penduduk tertinggi dan terendah di Kabupaten Banyuwangi. Jumlah timbulan sampah tertinggi berada di Kota Surabaya dan terendah di Kota Mojokerto.

Model yang terbentuk dari pemodelan regresi linear berganda dan *Geographically Weighted Regression* (GWR) adalah sebagai berikut:

$$\hat{Y} = 5.276.000 - 1.192X_4 + 9,517X_5$$

Model yang terbentuk dengan pemodelan GWR berbeda untuk setiap kabupaten/kota di Jawa Timur karena estimasi parameter yang berlaku secara lokal. Salah satu contoh model yang terbentuk sebagai berikut.

$$\hat{Y}_{Surabaya} = -3.763.865,04 + 35.498,28X_3 - 1.003,13X_4 + 3,20X_5$$

Berdasarkan nilai kriteria AIC dan R^2 , model *Geographically Weighted Regression* (GWR) lebih baik digunakan dibandingkan model regresi linear berganda.

References

- Anderson, T. W., & Darling, D. A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268), 765–769. <https://doi.org/10.1080/01621459.1954.10501232>
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publisher.
- Caraka, R. E., & Yasin, H. (2017). *Geographically Weighted Regression (GWR) ; Sebuah Pendekatan Regresi Geografis* (Pertama, hal. 160). Mobius.
- Dienelly, U., Bakri, S., & Santoso, T. (2017). Pengaruh Perubahan Tutupan Hutan Dan Lahan Terhadap Produk Domestik Regional Bruto (PDRB) Di Sektor Pertanian, Kehutanan Dan Industri : Studi Di Provinsi Lampung. *Jurnal Sylva Lestari*, 5(1), 61. <https://doi.org/10.23960/jsl1561-70>
- Faris, N. A. (2022). Analisis Pengaruh PDRB Sektor Pertanian, Industri, Dan Transportasi Terhadap Emisi Karbondioksida Di Provinsi Jawa Timur. *Jurnal Ilmiah Mahasiswa FEB*. <https://jimfeb.ub.ac.id/index.php/jimfeb/article/view/8176>
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons, Inc.
- Hocking, R. R. (2003). *Methods and Applications of Linear Models* (Kedua). John Wiley & Sons, Inc. <https://doi.org/10.2307/1271138>
- Lee, J., & Wong, D. W. S. (2001). *Statistical Analysis with ArcView GIS*. John Wiley & Sons, Inc.
- Nasir, R., Annas, S., & Nusrang, M. (2020). Pemodelan dengan Spatial Autoregressive (SAR) pada Angka Putus Sekolah Bagi Anak Usia Wajib Belajar di Provinsi Sulawesi Selatan. *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, 3(1), 44. <https://doi.org/10.35580/variansium9358>
- Ndruru, E, M.Situmorang, G. T. (2014). Analisa Faktor-Faktor yang Mempengaruhi Hasil Produksi Padi di Deli Serdang. *Saintia Matematika*, 2(1), 71–83.
- Suardin, M., Bustan, M. N., & Ahmar, A. S. (2020). Pemodelan Pertumbuhan Ekonomi Provinsi Sulawesi Selatan dengan Menggunakan Regresi Data Panel. *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, 2(3), 158. <https://doi.org/10.35580/variansium14637>
- Sugiyono. (2007). *Statistika Untuk Penelitian*. CV Alfabeta.
- Tiro, M. A. (2010). *Analisis Korelasi dan Regresi* (Ketiga). Andira Publisher.

Trenggonowati, D. L., Kulsum, K., & Salma, F. A. (2020). Model Persamaan Regresi Pengoptimalan Waktu Pelayanan Bongkar Kargo Menggunakan Multiple Regression Analysis di PT. XYZ. *Journal Industrial Services*, 6(1), 13. <https://doi.org/10.36055/jiss.v6i1.9484>

Yasin, H., Warsito, B., & Hakim, A. R. (2020). *Regresi Spasial (Aplikasi dengan R)* (Pertama). WADE Publish.