

***Geographically Weighted Regression* dalam Menganalisis Faktor-Faktor yang Mempengaruhi Kasus Tuberkulosis di Sumatera Utara**

Vera Maya Santi^{1*}, Afifah Nur Mutia², Qorry Meidianingsih³
*e-mail: vmsanti@unj.ac.id

^{1,2,3}*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Jakarta*

ABSTRACT

Geographically Weighted Regression (GWR) is a method used to analyze a case by considering spatial effects. Spatial effects indicate differences in environmental characteristics between regions. A spatial effect can be seen from the violation of the assumption of the multiple linear regression homoscedasticity test. Research in the health sector often shows that the spread of disease, especially infectious diseases, is influenced by spatial effects, one of which is tuberculosis. In 2020, North Sumatra will become one of the areas with high tuberculosis cases, so it is necessary to study tuberculosis cases and determine the factors that influence it by involving spatial effects. This study aims to model tuberculosis cases and analyze the ten factors thought to influence it using the GWR method with the adaptive kernel bi-square weighting function. The results showed that 23 regencies/cities had local GWR models and different factors influencing tuberculosis cases for each region, and ten other regencies/cities were influenced by other factors outside the variables studied. Based on the variables that affect it, 13 regional groups are formed. Modeling with the GWR method can explain tuberculosis cases in North Sumatra by 99.4%, and other predictor variables influence the rest.

Keywords: adaptive kernel bisquare, spatial effects, geographically weighted regression, tuberculosis

ABSTRAK

Geographically Weighted Regression (GWR) adalah metode yang digunakan untuk menganalisis suatu kasus dengan mempertimbangkan efek spasial. Efek spasial menunjukkan adanya perbedaan karakteristik lingkungan antar wilayah. Adanya efek spasial dapat dilihat dari pelanggaran pada uji asumsi homoskedastisitas regresi linier berganda. Penelitian di bidang kesehatan, seringkali menunjukkan bahwa penyebaran suatu penyakit terutama penyakit menular dipengaruhi oleh efek spasial salah satunya penyakit tuberkulosis. Pada tahun 2020 Sumatera Utara menjadi salah satu wilayah dengan kasus tuberkulosis tinggi, maka diperlukan kajian mengenai kasus tuberkulosis serta mengetahui faktor-faktor yang mempengaruhinya dengan melibatkan efek spasial. Penelitian ini bertujuan memodelkan kasus tuberkulosis dan analisis terhadap 10 faktor yang diduga mempengaruhinya menggunakan metode GWR dengan fungsi pembobot *adaptive kernel bi-square*. Hasil penelitian menunjukkan terdapat 23 Kabupaten/Kota memiliki model GWR lokal dan faktor-faktor yang berbeda dalam mempengaruhi kasus tuberkulosis untuk setiap wilayahnya dan 10 Kabupaten/Kota lainnya dipengaruhi faktor lain di luar dari variabel yang diteliti. Berdasarkan variabel yang mempengaruhinya terbentuk 13 kelompok wilayah. Pemodelan dengan metode

GWR mampu menjelaskan kasus tuberkulosis di Sumatera Utara sebesar 99,4%, sisanya dipengaruhi oleh variabel prediktor lain.

Kata Kunci: *adaptive kernel bi-square*, efek spasial, *geographically weighted regression*, tuberkulosis

PENDAHULUAN

Pada tahun 2020 tuberkulosis menjadi penyakit menular dengan angka kematian tertinggi nomor dua setelah Covid-19. Tuberkulosis disebabkan oleh *Mycobacterium Tuberculosis* (WHO, 2020), yang ditularkan antara manusia. Pada tahun 2020 Indonesia menjadi penyumbang kasus tuberkulosis tertinggi ke-3 dunia. Berdasarkan laporan Kemenkes RI pada tahun 2020 penderita tuberkulosis di Indonesia mencapai angka 351.936 kasus. Salah satu wilayah yang memiliki angka tuberkulosis yang tinggi di Indonesia adalah Provinsi Sumatera Utara dengan jumlah kasus 33.779 (Dinas Kesehatan Provinsi Sumatera Utara, 2020).

Menurut *Universal Health Coverage* (UHC) jumlah orang yang terinfeksi, penderita tuberkulosis aktif, dan jumlah kematian yang disebabkan oleh tuberkulosis dapat dikurangi melalui tindakan pencegahan multisektoral. Oleh karena itu, salah satu upaya yang dapat dilakukan dalam pencegahan kasus tuberkulosis dengan memberikan informasi berupa faktor-faktor yang mempengaruhi tuberkulosis di Sumatera Utara.

Berdasarkan kajian UHC dapat dilakukan analisis dengan melibatkan efek spasial untuk mengetahui faktor-faktor yang berpengaruh dalam peningkatan tuberkulosis di Sumatera Utara. Efek spasial merupakan perbedaan karakteristik lingkungan dan geografis antar daerah sehingga masing-masing daerah memiliki variasi yang berbeda disebut dengan heterogenitas spasial (Caraka & Yasin, 2017). Adanya efek spasial dapat dilihat dari pelanggaran

pada uji asumsi homoskedastisitas regresi linier berganda.

Metode yang dapat digunakan salah satunya dengan *Geographically Weighted Regression* (GWR). GWR merupakan pengembangan dari regresi linier dengan menambahkan faktor letak geografis yang menghasilkan estimator parameter yang hanya dapat digunakan untuk memprediksi setiap titik atau lokasi dimana data diamati (Fotheringham *et.al.*, 2002). Proses pembentukan model GWR membutuhkan pembobot spasial yang menjelaskan letak lokasi amatan. Pembobot pada GWR terdiri dari dua fungsi, yaitu fungsi *fixed kernel* dan fungsi *adaptive kernel*. Fungsi *fixed kernel* memiliki satu nilai *bandwidth* optimum tetap pada seluruh wilayah pengamatan dan fungsi *adaptive kernel* memiliki *bandwidth* optimum berbeda pada setiap wilayah pengamatan (Fotheringham *et.al.*, 2002).

Penerapan GWR dalam penelitian dengan menggunakan pembobot *fixed gaussian kernel* telah dilakukan diantaranya oleh Nadya *et.al.* (2017) yang meneliti kasus pneumonia balita di provinsi Jawa Barat. Ulfie Safitri & Amaliana (2021) melakukan penelitian dalam menentukan model GWR terbaik dengan pembobot *fixed kernel* dan *adaptive kernel* pada kasus kematian ibu di Jawa Timur tahun 2018 yang menghasilkan fungsi pembobot terbaik adalah *adaptive bi-square kernel*.

Berdasarkan pemaparan diatas, maka penelitian ini fokus pada pemodelan dan menganalisis faktor-faktor yang berpengaruh terhadap penyebaran penyakit tuberkulosis di Sumatera Utara pada tahun 2020

menggunakan metode regresi yang diberikan pembobot yaitu *Geographically Weighted Regression* dengan fungsi pembobot yang digunakan adalah *adaptive kernel bi-square*.

METODE

Pada bagian metode ini, akan menguraikan terkait Analisis regresi linier berganda, uji signifikansi, asumsi pada analisis regresi linier berganda, dan pemodelan menggunakan *Geographically Weighted Regression* (GWR).

Data Penelitian

Pemilihan variabel prediktor yang digunakan dalam penelitian ini mengacu pada penelitian Pelissari & Diaz-Quijano (2017), Maziyah (2016), Lestari dkk (2014), Zulaikha (2018), dan Karima & Hayati (2021). Data pada penelitian bersumber dari BPS Provinsi Sumatera Utara dan data publikasi Dinas Kesehatan Provinsi Sumatera Utara (2020), terdiri dari satu variabel respon yaitu jumlah kasus tuberkulosis (Y) dan sepuluh variabel prediktor yaitu persentase penduduk miskin (X_1), persentase Tempat-Tempat Umum (TTU) yang memenuhi syarat kesehatan (X_2), persentase Tempat Pengelolaan Makanan (TPM) yang memenuhi syarat kesehatan (X_3), kepadatan penduduk (X_4), jumlah puskesmas (X_5), persentase balita yang pernah mendapat imunisasi BCG (X_6), persentase rumah tangga yang memiliki akses terhadap sanitasi layak (X_7), tingkat pengangguran terbuka (X_8), jumlah gizi buruk pada balita (X_9), dan jumlah penduduk usia produktif (X_{10}).

Uji Asumsi

Pengujian asumsi pada regresi digunakan untuk melihat kebaikan dari data agar dapat dianalisis lebih lanjut.

Uji Normalitas

Uji normalitas digunakan untuk mengetahui nilai residual pada model regresi yang digunakan berdistribusi normal atau tidak, dengan hipotesis sebagai berikut:

H_0 : Galat berdistribusi normal

H_1 : Galat tidak berdistribusi normal

Statistik uji yang digunakan:

$$A^2 = -n - \left(\sum_{i=1}^n \frac{(2i-1)}{n} [\log F(Z_i) + \log (1 - F(Z_{n+1-i}))] \right)$$

Keterangan:

A^2 = Statistik uji *anderson darling*

Z_i = Data x_i yang distandarisasi

$F(Z_i)$ = Fungsi distribusi kumulatif dari normal baku di Z_i

Keputusan yang diambil atau kriteria daerah kritis akan terima hipotesis nol jika nilai $A^2 \geq AD_{tabel}$ atau jika nilai *p-value* > α (Anderson & Darling, 1954).

Uji Multikolinearitas

Uji multikolinearitas digunakan untuk melihat kondisi antar variabel prediktor apakah terdapat hubungan linier atau korelasi, dilakukan dengan melihat kriteria nilai *Variance Inflation Factor* (VIF) (Montgomery *et.al.*, 2012). Nilai yang diharapkan kurang 10, dengan rumus:

$$VIF_k = \frac{1}{1 - R_k^2}$$

Uji Autokorelasi

Uji autokorelasi digunakan untuk mengidentifikasi apakah terdapat ketergantungan spasial antar amatan terhadap variabel responnya atau tidak, dimana amatan pada penelitian ini adalah Kabupaten/Kota. Pengujian autokorelasi spasial dilakukan dengan menggunakan uji *Moran's I* (Anselin, 1988)

Statistik uji yang digunakan:

$$Z_{hit} = \frac{I - E(I)}{\sqrt{var(I)}}$$

$$I = \frac{n \sum_{i=j}^n \sum_{j=1}^n W_{ij}^* (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=j}^n \sum_{j \neq 1}^n W_{ij}^* \sum_{i=1}^n (y_i - \bar{y})^2}$$

$$E(I) = -\frac{1}{(n-1)}$$

$$\text{Var}(I) = \frac{n^2(n-1)V_1 - n(n-1)V_2 - 2V_0^2}{(n+1)(n-1)^2V_0^2}$$

Dimana

$$V_0 = \sum_{i=1}^n \sum_{j \neq i}^n W_{ij}^*$$

$$V_0 = \frac{1}{2} \sum_{i=1}^n \sum_{j \neq i}^n (W_{ij}^* + W_{ji}^*)^2$$

$$V_2 = \sum_{k=1}^n (\sum_{j=1}^n W_{kj}^* + \sum_{i=1}^n W_{ik}^*)^2, i, j, k = 1, 2, 3, \dots, n.$$

keputusan yang diambil adalah tolak hipotesis nol dengan taraf signifikan α jika nilai $|Z_{hit}| > Z_{\alpha/2}$ atau jika nilai p_{value} kurang dari α .

Uji Heteroskedastisitas

Uji heteroskedastisitas digunakan untuk melihat kondisi antar variabel dalam model apakah memiliki kesamaan varian dari residual atau tidak, hipotesis yang digunakan:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1: \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2$$

Statistika uji yang digunakan:

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

Keterangan:

BP = Nilai uji Breusch-Pagan

e_i = Galat bagi pengamatan ke- i

σ_i^2 = Ragam dari e_i

\mathbf{X} = Matriks berukuran $n \times (k+1)$

\mathbf{f} = Elemen vector \mathbf{f} ($f_i = \frac{e_i^2}{\sigma_i^2} - 1$),

dimana $i = 1, 2, \dots, n$

Keputusan yang diambil atau kriteria daerah kritis akan tolak hipotesis nol jika nilai $BP \geq \chi_k^2$ atau jika nilai $p\text{-value} < \alpha$ (Anselin, 1988).

Geographically Weighted Regression (GWR)

GWR adalah salah satu metode regresi yang digunakan untuk memodelkan data dengan mempertimbangan efek spasial atau lokasi. Model GWR menurut (Fotheringham *et.al.*, 2002) dituliskan sebagai berikut:

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

Penaksiran parameter pada GWR adalah dengan metode Weighted Least Square (WLS) (Fotheringham *et.al.*,

2002). Metode WLS memberikan pembobot yang berbeda setiap lokasi pengamatan. Bentuk penaksirannya adalah sebagai berikut:

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

Langkah awal dalam pemodelan GWR adalah menentukan nilai *bandwidth* optimum berdasarkan kriteria Cross Validation (CV), dengan rumus sebagai berikut:

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

Proses pendugaan parameter model GWR di suatu lokasi membutuhkan pembobot spasial, fungsi pembobot yang digunakan dalam penelitian ini adalah fungsi pembobot *adaptive kernel bisquare*, dimana masing-masing pembobot memerlukan nilai *bandwidth* optimum, sehingga nilai *bandwidth* yang digunakan akan berbeda-beda untuk setiap wilayahnya, dengan rumus sebagai berikut:

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_i} \right)^2 \right]^2, & \text{jika } d_{ij} < b_i \\ 0, & \text{lainnya} \end{cases}$$

Konstanta b merupakan *bandwidth* dan d_{ij} menunjukkan jarak euclid.

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$$

Pada analisis dengan metode GWR terdapat dua uji signifikansi parameter yang pertama adalah uji signifikansi secara simultan, dengan hipotesis sebagai berikut:

$$H_0: \beta_k(u_i, v_i) = \beta_k, k = 1, 2, \dots, p$$

$$H_1: \text{Paling sedikit ada satu } \beta_k(u_i, v_i)$$

yang berhubungan dengan lokasi (u_i, v_i)

Statistik uji yang digunakan:

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

Dimana:

$$SE(H_0) = \mathbf{Y}^T (\mathbf{I} - \mathbf{H}) \mathbf{Y},$$

$$\mathbf{H} = \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$$

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

$$df_1 = n - p - 1$$

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

Keputusan yang diambil atau kriteria daerah kritis akan tolak hipotesis

nol jika $F^* > F_{tabel}$ ($F_{tabel} = F_{\alpha;df_1;df_2}$). F^* akan mengikuti distribusi F dengan derajat bebas df_1 dan df_2 .

Pengujian signifikansi yang kedua adalah uji signifikansi secara parsial, hipotesis yang digunakan:

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

$$H_1: \beta_k(u_i, v_i) \neq 0, \text{dimana } i = 1, 2, \dots, n; k = 1, 2, \dots, p$$

Statistika uji yang digunakan:

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

Keterangan:

$\hat{\beta}_k(u_i, v_i)$ = Nilai penaksiran parameter pada lokasi (u_i, v_i)

g_{kk} = Elemen diagonal ke- k dari matriks \mathbf{GG}^T ,

dimana $\mathbf{G} = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T(u_i, v_i)$

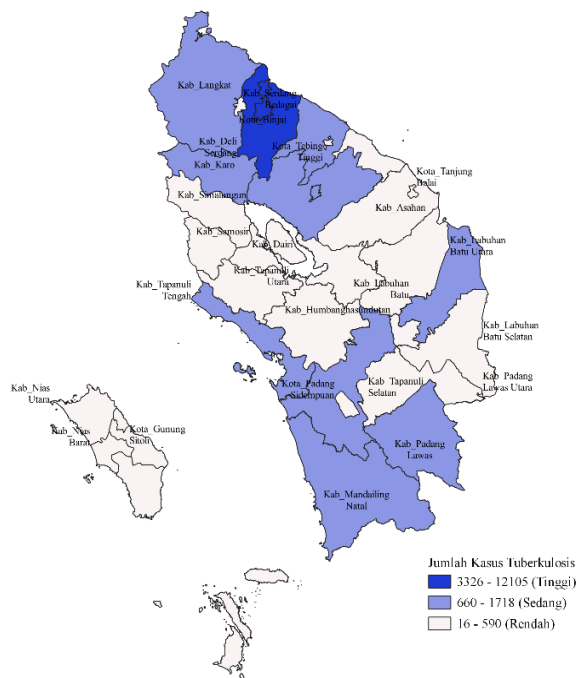
Pemilihan Model Terbaik

Menentukan model terbaik dengan membandingkan nilai *Akaike's Information Criterion* (AIC) serta *R Square* pada metode *Ordinary Least Square* dan *Geographically Weighted Regression*. Jika nilai AIC lebih rendah maka model lebih baik. Rumus AIC adalah sebagai berikut:

$$AIC = 2p - 2\ln\hat{L}(\hat{\theta})$$

HASIL DAN PEMBAHASAN

Hasil penelitian dan pembahasan akan menguraikan hasil dari analisis yang telah dilakukan.



Gambar 1. Sebaran Kasus Tuberkulosis

Data yang digunakan dalam penelitian ini bersumber dari Badan Pusat Statistik Provinsi Sumatera Utara dan data publikasi Dinas Kesehatan Provinsi Sumatera Utara yaitu “Profil Kesehatan Provinsi Sumatera Utara Tahun 2020”. Data terdiri atas 33 amatan Kabupaten/Kota yang meliputi 25

Kabupaten dan 8 Kota di Provinsi Sumatera Utara.

Gambar 1 menunjukkan peta sebaran jumlah kasus tuberkulosis di Sumatera Utara pada tahun 2020 yang menunjukkan bahwa sebagian besar Kabupaten/Kota memiliki kategori jumlah kasus tuberkulosis rendah sebanyak 21 Kabupaten/Kota, kategori

sedang sebanyak 10 Kabupaten/Kota. Kota Medan dan Kabupaten Deli Serdang menjadi wilayah dengan kasus tuberkulosis kategori tinggi.

Uji Normalitas

Uji normalitas menggunakan uji anderson darling, dihasilkan sebagai berikut:

Tabel 1. Uji Normalitas

Nilai D	p-value
0,367	0,412

Hasil pada Tabel 1. menunjukkan nilai p-value sebesar 0,412, menunjukkan bahwa $p\text{-value} > \alpha$ yang artinya terima hipotesis H_0 sehingga diperoleh kesimpulan bahwa galat berdistribusi normal.

Uji Multikolinearitas

Uji multikolinearitas menggunakan kriteria nilai VIF yang dihasilkan sebagai berikut:

Tabel 2. Uji Multikolinearitas

Variabel	VIF
X_1	1,953
X_2	1,414
X_3	1,517
X_4	2,546
X_5	3,206
X_6	1,494
X_7	1,875
X_8	2,797
X_9	1,267
X_{10}	4,295

Berdasarkan nilai VIF yang dihasilkan pada setiap variabel menunjukkan kurang dari 10. Hal ini menyatakan bahwa semua variabel tidak ada hubungan linier atau korelasi antara variabel prediktor pada model.

Uji Autokorelasi

Pemeriksaan autokorelasi spasial pada penelitian ini menggunakan uji Moran's I, dihasilkan sebagai berikut:

Tabel 3. Uji Autokorelasi

Moran I statistic standard deviate	p-value
4,6693	1,511e-06

hasil analisis yang telah dilakukan dengan menggunakan taraf signifikansi 5% pada uji Moran's I diperoleh nilai $p\text{-value}$ sebesar 1,511e-06. Maka, dengan $\alpha = 0,05 > p\text{-value}$ terjadi penolakan terhadap hipotesis no. Hal ini menunjukkan bahwa terdapat ketergantungan spasial yang signifikan antar Kabupaten/Kota terkait kasus tuberkulosis di Sumatera Utara.

Uji Heteroskedastisitas

Uji heteroskedastisitas dengan uji Breusch-Pagan, dihasilkan sebagai berikut:

Tabel 4. Uji Heteroskedastisitas

Nilai uji BP	Df	p-value
19,089	10	0,039

Tabel 5 menunjukan hasil yang diperoleh adalah nilai p-value sebesar $0,039 < 0,05$ diperoleh keputusan tolak hipotesis nol karena $p\text{-value} < \alpha$, hal ini menunjukkan bahwa varians dari residual dalam model memiliki perbedaan.

Pemodelan GWR

Langkah awal yang dapat dilakukan adalah dengan menentukan matriks jarak antar 33 Kabupaten/Kota menggunakan rumus jarak euclid, matriks jarak yang dihasilkan berukuran 33×33 sebagai berikut:

$$D_{33 \times 33} = \begin{bmatrix} 0 & 1,662 & 1,613 & \dots & 0,233 \\ 1,665 & 0 & 0,839 & \dots & 1,838 \\ 1,613 & 0,839 & 0 & \dots & 1,697 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0,233 & 1,838 & 1,697 & \dots & 0 \end{bmatrix}$$

Langkah kedua, menentukan nilai *bandwidth* optimum berdasarkan kriteria *Cross Validation* (CV) dengan fungsi pembobot *adaptive kernel bi-square*, nilai *bandwidth* yang digunakan akan berbeda-beda untuk 33 Kabupaten/Kota.

Tabel 5. Nilai *Bandwidth*

No	Kabupaten/Kota	Nilai <i>Bandwidth</i>
1.	Kab Nias	2,434
2.	Kab Mandailing Natal	2,206
⋮	⋮	⋮
32.	Kota Padangsidimpuan	1,671
33.	Kota Gunungsitoli	2,400

Setelah mendapatkan matriks jarak dan nilai *bandwidth*, maka dapat menentukan matriks pembobot untuk 33 Kabupaten/Kota diperoleh matriks ordo 33×33 untuk data tuberkulosis di Sumatera Utara, sebagai berikut:

$$W_{33 \times 33} = \begin{bmatrix} 1 & 0,285 & \dots & 0,982 \\ 0,187 & 1 & \dots & 0,093 \\ \vdots & \vdots & \ddots & \vdots \\ 0,981 & 0,171 & \dots & 1 \end{bmatrix}$$

Uji Kesesuaian Model

Berdasarkan Tabel 7 berikut, hasil penelitian menunjukkan diperoleh keputusan tolak hipotesis nol karena $F^* > F_{tabel}$. Hal ini menunjukkan terdapat perbedaan yang signifikan antara model regresi global dan GWR.

Table 6. Uji Kesesuaian Model

Sumber	Df	F value	F tabel
<i>OLS Residual</i>	11,000		
<i>GWR Improvement</i>	16,689		
<i>GWR Residual</i>	5,3101	3.487	2,456

Uji Signifikansi Parameter

Taraf signifikan yang digunakan sebesar 5% dengan pengujian parameter

dilakukan dengan uji T dapat menunjukkan setiap Kabupaten/Kota memiliki faktor-faktor yang berbeda, hal ini dilakukan dengan melihat dari nilai statistik uji t jika variabel penjelas memiliki $|T_{hitung}| > t_{(0,025;31)} = 2,040$. Berikut adalah variabel yang signifikan pada masing-masing Kabupaten/Kota:

Tabel 7. Uji Signifikansi Parameter

No	Kabupaten/Kota	Hasil Uji
1	Kabupaten Nias, Kabupaten Humbang Hasundutan, Kabupaten Nias Utara, Kabupaten Nias Barat, dan Kota Gunungsitoli	X_{10}
2	Kabupaten Labuhanbatu	X_6, X_8, X_{10}
3	Kabupaten Asahan dan Kota Tanjungbalai	$X_4, X_5, X_6, X_7, X_8, X_{10}$
4	Kabupaten Simalungun	$X_3, X_4, X_5, X_6, X_7, X_8, X_{10}$
5	Kabupaten Dairi dan Kabupaten Samosir	$X_1, X_4, X_6, X_7, X_8, X_{10}$
6	Kabupaten Karo, Kabupaten Deli Serdang, Kabupaten Langkat, Kota Medan, Kota Binjai	$X_4, X_6, X_7, X_8, X_{10}$
7	Kabupaten Pakpak Bharat	X_1, X_4, X_8, X_{10}
8	Kabupaten Serdang Bedagai	$X_3, X_4, X_6, X_7, X_8, X_{10}$
9	Kabupaten Batu Bara dan Kota Tebing Tinggi	$X_3, X_4, X_7, X_8, X_{10}$
10	Kabupaten Labuhanbatu Selatan	X_8
11	Kabupaten Labuhanbatu Utara	X_8, X_{10}
12	Kota Pematangsiantar	$X_1, X_3, X_4, X_5, X_6, X_7, X_8, X_{10}$

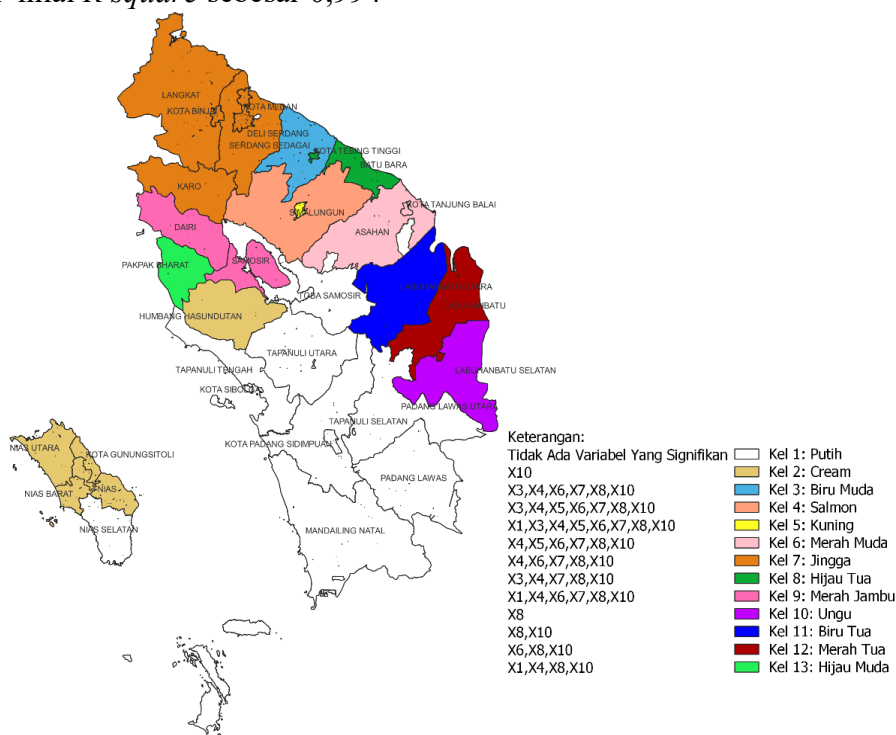
Pemilihan Model Terbaik

Tabel 8. Pemilihan Model Terbaik

Metode	AIC	R Square
Ordinary Least Square	528,362	0,912
Geographically Weighted Regression	447,463	0,994

hal ini menunjukkan bahwa metode *Geographically Weighted Regression* mampu menjelaskan kasus penyakit tuberkulosis di Sumatera Utara sebesar 99,4%, sedangkan sisanya dipengaruhi oleh variabel lainnya.

Pada Tabel 9 di atas menunjukkan bahwa model terbaik adalah model GWR, hal ini berdasarkan kriteria nilai AIC terkecil dan nilai *R square* terbesar. Model GWR memiliki nilai AIC sebesar 447,463 dan nilai *R square* sebesar 0,994



Gambar 2. Pengelompokan Wilayah

Berdasarkan hasil uji signifikansi parameter dengan hasil seperti yang tersaji pada Tabel 8, terbentuk pengelompokan wilayah Kabupaten/Kota berdasarkan variabel yang signifikan terhadap kasus tuberkulosis pada setiap Kecamatan/Kota di Sumatera Utara yang ditunjukkan pada Gambar 2 di atas.

KESIMPULAN

Model dari jumlah kasus tuberkulosis di Sumatera Utara menggunakan GWR menghasilkan 23 model lokal untuk setiap Kabupaten/Kota yang memiliki variabel yang signifikan mempengaruhi kasus tuberkulosis, 10 Kabupaten/Kota lainnya tidak memiliki variabel yang berpengaruh signifikan dari kesepuluh variabel yang digunakan. Pemodelan GWR menggunakan fungsi

adaptive kernel bi-square pada kasus tuberkulosis di Sumatera Utara menghasilkan 13 kelompok Kabupaten/Kota yang berbeda, perbedaan dari setiap kelompoknya berdasarkan karakteristik variabel yang mempengaruhinya secara signifikan.

DAFTAR PUSTAKA

- Anderson, T. W., & Darling, D. A. (1954). A TEST OF GOODNESS OF FIT. *Journal Of the American Statistical Association*, 49, 765–769.
- Anselin, L. (1988). Spatial Econometrics: Methods and Models. In *Economic Geography* (Vol. 65, Issue 2). <https://doi.org/10.2307/143780>.
- Caraka, R. E., & Yasin, H. (2017). *Geographically Weighted Regression; Sebuah Pendekatan Regresi Geografis*. Yogyakarta: Mobius. <https://doi.org/10.4135/9781412953962.n81>.
- Dinas Kesehatan Provinsi Sumatera Utara. (2020). Profil Kesehatan Provinsi Sumatera Utara. *Dinas Kesehatan Provinsi Sumatera Utara*, 1–422. <http://dinkes.sumutprov.go.id/unduh>.
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). Geographically Weighted Regression: the analysis of spatially varying relationships. In *Syria Studies*. John Wiley & Sons, Ltd. https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civilwars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625
- Indonesia, K. K. R. (2021). Profil Kesehatan Indonesia 2020. In *Jakarta*.
- Karima, N. A., & Hayati, N. (2021). Analisis Faktor-Faktor yang Mempengaruhi Jumlah Kasus Tuberkulosis di Indonesia Menggunakan Model Geographically Weighted Poisson Regression Analysis of the Factors Affecting the Number of Tuberculosis Cases in Indonesia By Using Geographically Weighted P. 12(2016).
- Lestari, R. D., Wulandari, S. P., & Purhadi. (2014). Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Kasus Penyakit Tuberkulosis di Jawa Timur dengan Pendekatan Generalized Poisson Regression dan Geographically Weighted Poisson Regression. *Jurnal Sains Dan Seni Pomits*, 3(2), D-188-D-193.
- Maziyah, R. (2016). *Pemodelan Prevalensi Penyakit Tuberkulosis (TB) Beserta Faktor-Faktor yang Berpengaruh Menggunakan Metode Geographically Weighted Regression (Studi Kasus Penyakit Tuberkulosis di Kota Surabaya Tahun 2014) [Skripsi]*.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis* (5th ed.). John Wiley & Sons, Inc.
- Nadya, M., Rahayu, W., & Santi, V. M. (2017). Analisis Geographically Weighted Regression (Gwr) Pada Kasus Pneumonia Balita Di Provinsi Jawa Barat. *Jurnal Statistika Dan Aplikasinya*, 1(1), 23–32. <https://doi.org/10.21009/jsa.01103>.
- Pelissari, D. M., & Diaz-Quijano, F. A. (2017). Household crowding as a potential mediator of socioeconomic determinants of tuberculosis incidence in Brazil. *PLoS ONE*, 12(4), 1–13. <https://doi.org/10.1371/journal.pone>.

0176116.

- Ulfie Safitri, & Amaliana, L. (2021). Model Geographically Weighted Regression dengan Fungsi Pembobot Adaptive dan Fixed Kernel pada Kasus Kematian Ibu di Jawa Timur. *Jurnal Statistika Dan Aplikasinya*, 5(2), 208–220. <https://doi.org/10.21009/jsa.05209>
- WHO. (2020). Tuberculosis Reports. In *The Lancet* (Vol. 188, Issue 4870). [https://doi.org/10.1016/S0140-6736\(00\)58733-9](https://doi.org/10.1016/S0140-6736(00)58733-9).
- Zulaikha, E. (2018). Pemetaan Dan Analisis Faktor-Faktor Yang Mempengaruhi Tuberkulosis Menggunakan Geographically Weighted *Skripsi*. <https://dspace.uui.ac.id/handle/123456789/7428>.