

**Pemodelan *Mixed Geographically Weighted Regression* (MGWR)  
(Studi Kasus: Jumlah Penderita Diare di Provinsi Kalimantan Timur Tahun 2015)**

***The Model of Mixed Geographically Weighted Regression (MGWR)  
(Case Study: The Number of Diarrhea Sufferers in East Kalimantan Province on 2015)***

**Nur Fajar Apriyani<sup>1</sup>, Desi Yuniarti<sup>2</sup>, dan Memi Nor Hayati<sup>3</sup>**

<sup>1,3</sup>Laboratorium Statistika Terapan FMIPA Universitas Mulawarman

<sup>2</sup>Laboratorium Statistika Ekonomi dan Bisnis Jurusan Matematika FMIPA Universitas Mulawarman

E-mail: [nurfajarapriyani53@gmail.com](mailto:nurfajarapriyani53@gmail.com)

**Abstract**

*Diarrhea disease is one of the conditions which a person has soft or liquid defecate consistency, even can be water and frequency more often in one day. The province of East Kalimantan includes areas where the percentage of diarrhea tends to increase annually. Therefore, as one of the efforts to handle cases of diarrhea in East Kalimantan Province, so that the research using Mixed Geographically Weighted Regression (MGWR) model which is a modeling method that combines global regression model and Geographically Weighted Regression (GWR) model. Modeling MGWR aim to find out the factors that affect the number of diarrhea sufferers, where factors are differentiated into factors that affect locally in each District/City and factors that affect globally throughout the District/City. The result of the research using the MGWR method, the variable of the number of households that live clean and healthy and the number of food management places do not meet the criteria affect globally. The number of communal latrine facilities affect locally.*

*Keywords : diarrhea, global, GWR, local, MGWR*

## PENDAHULUAN

Pembangunan di bidang kesehatan masih menjadi perhatian utama bagi pemerintah di Indonesia. Timbulnya penyakit akibat tempat tinggal yang kurang ramah lingkungan merupakan permasalahan bagi negara-negara berkembang. Penyakit yang sering dialami oleh masyarakat di negara berkembang terutama Indonesia adalah penyakit diare.

Penyakit diare adalah salah satu kondisi di mana seseorang buang air besar dengan konsistensi lembek atau cair, bahkan dapat berupa air saja dan frekuensinya lebih sering (biasanya tiga kali atau lebih) dalam satu hari (Depkes, 2011).

Provinsi Kalimantan Timur termasuk wilayah dengan jumlah persentase diare cenderung meningkat setiap tahunnya. Kabupaten/Kota di Provinsi Kalimantan Timur memiliki potensi yang berbeda dari segi ekonomi, pendidikan, kesehatan dan pelayanan masyarakat yang diindikasikan adanya efek spasial.

Analisis regresi adalah metode analisis statistika yang digunakan untuk memodelkan hubungan kebergantungan yang mungkin ada antara variabel respon  $Y$  dan variabel prediktor  $X$ . Jika metode regresi diterapkan pada data spasial dengan asumsi *error* identik independen dan berdistribusi normal, maka akan diperoleh suatu model taksiran untuk semua data. Hal inilah yang menyebabkan ketidaksesuaian model pada data

spasial karena pada analisis regresi diasumsikan bahwa lokasi geografis tidak mempengaruhi respon model.

Metode lain yang bisa digunakan adalah *Geographically Weighted Regression* (GWR). Model GWR adalah pengembangan dari model regresi di mana setiap parameter dihitung pada setiap lokasi pengamatan mempunyai nilai parameter regresi yang berbeda-beda atau bersifat lokal (Fotheringham, et al, 2002).

Pada saat pengujian parameter prediktor GWR ada beberapa variabel yang tidak signifikan atau tidak mempunyai pengaruh lokasi, namun bila dikaji lebih lanjut ternyata variabel-variabel ini ada yang berpengaruh secara global. Maka dari itu dikembangkan lagi metode *Mixed Geographically Weighted Regression* (MGWR). Estimasi parameter pada model MGWR dapat dilakukan dengan metode *Weighted Least Square* (WLS) seperti halnya pada model GWR. Dengan langkah awal yaitu dengan membentuk matriks pembobot untuk setiap lokasi pengamatan (Fotheringham, et al, 2002).

Penelitian sebelumnya yang mengkaji pemodelan GWR dan MGWR telah dilakukan oleh Purhadi dan Yasin (2012) yang melakukan pemodelan persentase rumah tangga miskin di Mojokerto dengan menggunakan pendekatan regresi global, GWR, dan MGWR dimana pada penelitian tersebut didapatkan model MGWR

untuk kasus persentase rumah tangga miskin di Mojokerto.

Penelitian ini dibatasi pada pemodelan MGWR dengan fungsi pembobot *Adaptive Kernel Gaussian* pada data jumlah penderita diare di Provinsi Kalimantan Timur tahun 2015. Adapun tujuan dari penelitian ini yaitu untuk mengetahui model MGWR pada studi kasus jumlah penderita diare di Provinsi Kalimantan Timur tahun 2015 serta mengetahui faktor-faktor apa saja yang mempengaruhi jumlah penderita diare di Provinsi Kalimantan Timur tahun 2015.

**Data Spasial**

Creese (1991) menyatakan bahwa data spasial merupakan data yang dikumpulkan dari lokasi spasial berbeda dan memiliki sifat ketergantungan antara pengukuran data dengan lokasi. Data spasial berasumsikan berdistribusi normal dan memiliki hubungan secara spasial untuk dapat dianalisis secara spasial. Pada saat ini data spasial menjadi media yang penting dalam pengambilan kebijakan perencanaan pembangunan dan pengelolaan sumber daya alam. Pemanfaatan data spasial semakin berkembang yang dikarenakan adanya teknologi dan pemanfaatannya pada Sistem Informasi Geografis (SIG). Pada umumnya gambaran/deskripsi yang digunakan adalah berupa peta atau gambar dengan format digital yang memiliki titik koordinat tertentu.

**Model Regresi Linier**

Regresi linier merupakan metode yang memodelkan hubungan linier antara variabel respon  $Y$  dan variabel prediktor  $X_1, X_2, \dots, X_p$ . Model regresi linier untuk pengamatan ke- $i$  dengan  $p$  variabel prediktor jika diambil sebanyak  $n$  pengamatan ditulis sebagai berikut:

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

dengan  $\beta_0, \beta_1, \dots, \beta_p$  adalah parameter model dan  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  adalah *error* yang diasumsikan identik, independen, dan berdistribusi normal dengan mean nol dan varians konstan  $\sigma^2$  atau  $\varepsilon_i \sim IIDN(0, \sigma^2)$ .

Estimator dari parameter model didapat dengan meminimumkan jumlah kuadrat residual atau yang dikenal dengan *Ordinary Least Square* (OLS) (Rencher, 2000). Dengan taksiran sebagai berikut:

$$\hat{\beta} = (X^T X)^{-1} X^T y \tag{2}$$

**Metode Geographically Weighted Regression**

*Geographically Weighted Regression* (GWR) adalah salah satu model spasial dengan vektor

titik. GWR merupakan pengembangan dari model regresi linier OLS menjadi model regresi terboboti dengan memperhatikan efek spasial, sehingga menghasilkan penduga parameter yang hanya dapat digunakan untuk memprediksi setiap titik atau lokasi dimana data tersebut diamati dan disimpulkan (Fortheringham, 2002).

Asumsi yang digunakan pada model GWR adalah *error* berdistribusi normal dengan *mean* nol dan varians konstan  $\sigma^2$ . Model GWR dapat ditulis sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon_i; i = 1, 2, \dots, n \tag{3}$$

Di mana:

- $y_i$  : nilai observasi variabel respon ke- $i$
- $x_{ik}$  : nilai observasi variabel prediktor ke- $k$  pada lokasi pengamatan ke- $i$
- $\beta_0(u_i, v_i)$  : konstanta / *intercept* pada pengamatan ke- $i$
- $(u_i, v_i)$  : menyatakan koordinat letak geografis (*longitude, latitude*) dari lokasi pengamatan ke- $i$
- $\beta_k(u_i, v_i)$  : nilai observasi variabel prediktor ke- $k$  pada lokasi pengamatan ke- $i$
- $\varepsilon_i$  : error pengamatan ke- $i$

Estimasi parameter model GWR menggunakan metode *Weighted Least Square* (WLS) yaitu dengan memberikan pembobot yang berbeda untuk setiap lokasi pengamatan.

estimasi parameter pada  $(u_i, v_i)$  adalah :

$$\hat{\beta}(u_i, v_i) = [X^T W(u_i, v_i) X]^{-1} X^T W(u_i, v_i) y \tag{4}$$

(Mei, 2006).

**Pemilihan Pembobot**

Fungsi dari pembobot adalah untuk memberikan hasil estimasi parameter yang berbeda pada lokasi yang berbeda. Pada analisis spasial, estimasi parameter di suatu lokasi akan lebih dipengaruhi oleh titik-titik yang dekat dengan lokasi tersebut daripada titik-titik yang lebih jauh (Leung, *et.al* 2000). Pada jenis data titik, pembobot untuk setiap lokasi ke- $i$  pada koordinat  $(u_i, v_i)$  dinyatakan dengan  $w(u_i, v_i)$ .

Ada beberapa cara yang dapat digunakan untuk menentukan besarnya pembobot untuk masing-masing lokasi yang berbeda pada model GWR, diantaranya adalah Fungsi Kernel dimana fungsi pembobotnya masing-masing dapat ditulis sebagai berikut:

- Fungsi Kernel Gaussian :

$$w_j(u_i, v_i) = \exp \left\{ -\frac{1}{2} \left( \frac{d_{ij}}{h} \right)^2 \right\} \quad (5)$$

dengan  $d_{ij}$  jarak Euclidian antara lokasi  $(u_i, v_i)$  ke lokasi  $(u_j, v_j)$  dan  $h$  adalah parameter penghalus (*bandwidth*).

$$\text{dengan } d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (6)$$

Bandwidth ( $h$ ) dapat ditentukan menggunakan *Cross Validation* (CV). Misalkan

$$CV = \sum_{i=1}^n (y_i - \hat{y}_{(i)}(h))^2 \quad (7)$$

di mana  $\hat{y}_{(i)}(h)$  adalah *fitted value* dari  $y_i$  dengan pengamatan di lokasi  $(u_i, v_i)$  dihilangkan dari proses *fitting*. Untuk mendapatkan nilai  $h$  yang optimal maka diperoleh dari  $h$  yang menghasilkan CV yang minimum.

- Fungsi Kernel *Bisquare* :

$$w_j(u_i, v_i) = \begin{cases} \left( 1 - \left( \frac{d_{ij}}{h} \right)^2 \right)^2, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \quad (8)$$

- Fungsi Kernel *Tricube* :

$$w_j(u_i, v_i) = \begin{cases} \left( 1 - \left( \frac{d_{ij}}{h} \right)^3 \right)^3, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \quad (9)$$

- Fungsi Kernel *Adaptive Gaussian*

$$: w_j(u_i, v_i) = \exp \left\{ -\frac{1}{2} \left( \frac{d_{ij}}{h_i} \right)^2 \right\} \quad (10)$$

- Fungsi Kernel *Adaptive Bisquare*:

$$w_j(u_i, v_i) = \begin{cases} \left( 1 - \left( \frac{d_{ij}}{h_i} \right)^2 \right)^2, & \text{untuk } d_{ij} \leq h_i \\ 0, & \text{untuk } d_{ij} > h_i \end{cases} \quad (11)$$

- Fungsi Kernel *Adaptive Tricube* :

$$w_j(u_i, v_i) = \begin{cases} \left( 1 - \left( \frac{d_{ij}}{h_i} \right)^3 \right)^3, & \text{untuk } d_{ij} \leq h_i \\ 0, & \text{untuk } d_{ij} > h_i \end{cases} \quad (12)$$

pengujian parameter model GWR secara parsial Bentuk hipotesisnya adalah sebagai berikut:

Hipotesis:

$$H_0 : \beta_k(u_i, v_i) = 0 ; k=1,2,\dots,p$$

$$H_1 : \text{paling tidak ada satu } \beta_k(u_i, v_i) \neq 0$$

Statistik uji yang digunakan adalah sebagai berikut:

$$T_{hit} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma} \sqrt{C_{kk}}} \quad (13)$$

Kriteria ujinya adalah tolak  $H_0$  jika

nilai  $|T_{hit}| > t_{(\alpha/2; df)}$  di mana  $df = \delta_1^2 / \delta_2$

$$\delta_i = \text{tr} \left( \left[ (\mathbf{I} - \mathbf{L})^T (\mathbf{I} - \mathbf{L}) \right]^i \right), i = 1, 2$$

### Metode *Mixed Geographically Weighted Regression*

Metode *Mixed Geographically Weighted Regression* (MGWR) adalah suatu metode pemodelan yang menggabungkan model regresi global dengan model regresi yang terboboti. Model MGWR dengan  $p$  variabel prediktor dan  $q$  variabel prediktor diantaranya bersifat lokal, dengan mengasumsikan bahwa *intercept* model bersifat lokal dapat dituliskan sebagai berikut

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

$$i = 1, 2, \dots, n ; k = q + 1, q + 2, \dots, p \quad (14)$$

Estimasi parameter model MGWR dapat dituliskan sebagai berikut:

$$\hat{\beta}_g = \left[ \mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l) \mathbf{X}_g \right]^{-1} \mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l) \mathbf{y} \quad (15)$$

$$\hat{\beta}_l(u_i, v_i) = \left[ \mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l \right]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i) (\mathbf{y} - \mathbf{X}_g \hat{\beta}_g) \quad (16)$$

$$\mathbf{S}_l = \begin{bmatrix} \mathbf{x}_l^T [\mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i) \\ \mathbf{x}_l^T [\mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i) \\ \dots \\ \mathbf{x}_l^T [\mathbf{X}_l^T \mathbf{W}(u_i, v_i) \mathbf{X}_l]^{-1} \mathbf{X}_l^T \mathbf{W}(u_i, v_i) \end{bmatrix}_{n \times n}$$

### Pengujian Hipotesis Model *Mixed Geographically Weighted Regression*

Uji F dilakukan pertama kali untuk uji kesesuaian model regresi global dan model MGWR, maka bentuk hipotesisnya adalah sebagai berikut:

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

$$H_1 : \text{minimal ada satu}$$

$$\beta_k(u_i, v_i) \neq \beta_k \text{ dengan}$$

$$k = 0, 1, 2, \dots, q, \text{ dan } i = 1, 2, \dots, n$$

Statistik Uji

$$F_1 = \left( \frac{\mathbf{y}^T [(\mathbf{I} - \mathbf{H}) - (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S})] \mathbf{y} / v_i}{\mathbf{y}^T (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S}) \mathbf{y} / u_i} \right) \quad (17)$$

dengan:

$$\mathbf{S} = \mathbf{S}_l + (\mathbf{I} - \mathbf{S}_l) \mathbf{X}_g [\mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l) \mathbf{X}_g]^{-1} \mathbf{X}_g^T (\mathbf{I} - \mathbf{S}_l)^T (\mathbf{I} - \mathbf{S}_l)$$

$$u_i = tr([(I-S)^T (I-S)^i]), i = 1, 2$$

$$v_i = tr([(I-H) - (I-S)^T (I-S)]^i), i = 1, 2$$

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

$$df_1 = \left(\frac{v_1}{v_2}\right) \text{ dan } df_2 = \left(\frac{u_1^2}{u_2}\right). \text{ Tolak } H_0 \text{ jika}$$

$$F_1 \geq F_{(\alpha, df_1, df_2)} \text{ (Mei, et al, 2006).}$$

Yang pertama adalah pengujian serentak pada parameter variabel prediktor global. Dengan hipotesis sebagai berikut:

$$H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_p = 0$$

$$H_1 : \text{Minimal ada satu } \beta_k \neq 0$$

$$k = q + 1, q + 2, \dots, p$$

Statistik Uji:

$$F_2 = \left(\frac{y^T [(I-S_i)^T (I-S)^T (I-S)] y / r_i}{y^T (I-S)^T (I-S) y / u_i}\right)$$

(22)

dengan:

$$u_i = tr([(I-S)^T (I-S)^i]), i = 1, 2$$

$$r_i = tr([(I-S_i)^T (I-S_i) - (I-S)^T (I-S)]^i),$$

$$i = 1, 2$$

Dengan derajat bebas  $df_1 = \left(\frac{r_1}{r_2}\right)$  dan

$$df_2 = \left(\frac{u_1^2}{u_2}\right). \text{ Tolak } H_0 \text{ jika } F_2 \geq F_{(\alpha, df_1, df_2)}.$$

Selanjutnya uji serentak yang kedua adalah uji hipotesis serentak pada parameter variabel prediktor lokal. Dengan bentuk hipotesis sebagai berikut:

$$H_0 :$$

$$\beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_q(u_i, v_i) = 0,$$

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

$$H_1 : \text{minimal ada satu } \beta_k(u_i, v_i) \neq 0$$

Statistik Uji:

$$F_3 = \left(\frac{y^T [(I-S_g)^T (I-S_g) - (I-S)^T (I-S)] y / t_i}{y^T (I-S)^T (I-S) y / u_i}\right) \quad (18)$$

dengan:

$$t_i = tr([(I-S_g)^T (I-S_g) - (I-S)^T (I-S)]^i),$$

$$i = 1, 2.$$

$$S_g = X_g [X_g^T X_g]^{-1} X_g^T$$

$$df_1 = \left(\frac{t_1}{t}\right) \text{ dan } df_2 = \left(\frac{u_1^2}{u_2}\right). \text{ Tolak } H_0 \text{ jika}$$

$$F_3 \geq F_{(\alpha, df_1, df_2)}.$$

Uji selanjutnya adalah pengujian parsial parameter model MGWR. Untuk pengujian signifikansi pada variabel global digunakan hipotesis sebagai berikut:

$$H_0 : \beta_k = 0 ; k = q + 1, q + 2, \dots, p \text{ (Variabel global } X_k \text{ tidak signifikan)}$$

$$H_1 : \beta_k \neq 0 ; k = q + 1, q + 2, \dots, p \text{ (Variabel global } X_k \text{ signifikan)}$$

Statistik Uji:

$$T_{g\_hit} = \frac{\hat{\beta}_k}{\hat{\sigma} \sqrt{g_{kk}}} \quad (19)$$

$$G = [X_g^T (I-S_g)^T (I-S_g) X_g]^{-1} X_g^T (I-S_g)^T (I-S_g),$$

$$\hat{\sigma}^2 = \frac{y^T (I-S)^T (I-S) y}{tr((I-S)^T (I-S))}$$

dengan  $g_{kk}$  adalah elemen diagonal ke- $k$  dari matriks  $GG^T$  signifikansi sebesar  $\alpha$ , maka dapat diambil keputusan tolak  $H_0$  jika  $|T_{g\_hit}| > t_{\alpha/2, df}$ ,

$$\text{dimana } df = \left(\frac{u_1^2}{u_2}\right).$$

Uji hipotesis selanjutnya ditunjukkan untuk mengetahui variabel lokal yang berpengaruh signifikan terhadap respon pada model MGWR digunakan hipotesis sebagai berikut:

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

$$H_1 : \beta_k(u_i, v_i) \neq 0 \text{ Untuk } k = 1, 2, \dots, q \text{ dan } i = 1, 2, \dots, n$$

Statistik Uji:

$$T_{i\_hit} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma} \sqrt{m_{kk}}} \quad (20)$$

dengan  $m_{kk}$  adalah elemen diagonal ke- $k$  dari matriks  $MM^T$  dan  $M = \frac{X_i^T W(u_i, v_i) (I - X_g G)}{X_i^T W(u_i, v_i) X_i}$  (21)

Tolak  $H_0$  jika  $|T_{i\_hit}| > t_{\alpha/2, df}$  dimana  $df = \left(\frac{u_1^2}{u_2}\right)$ .

### Penyakit Diare

Menurut Departemen Kesehatan Republik Indonesia (Depkes RI) (2005), diare adalah suatu penyakit dengan tanda-tanda adanya perubahan bentuk dan konsistensi dari tinja, yang melembek sampai mencair dan bertambahnya frekuensi BAB biasanya tiga kali atau lebih dalam sehari.

Penyakit diare merupakan salah satu penyakit yang berbasis lingkungan, dua faktor yang sangat mendominasi adalah sarana air bersih dan pembuangan tinja. Kedua faktor ini akan berinteraksi bersama perilaku manusia, apabila faktor lingkungan yang tidak sehat karena tercemar bakteri atau virus serta berakumulasi dengan perilaku manusia yang tidak sehat pula,

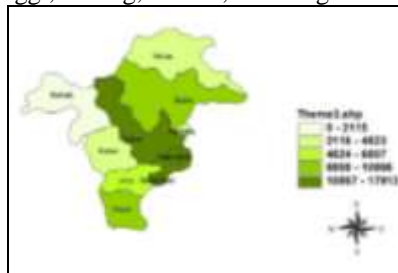
maka dapat menimbulkan penyakit diare (Depkes RI, 2005).

**Hasil dan Pembahasan**

Data yang digunakan dalam penelitian ini merupakan data persentase kejadian penyakit diare dari tiap Kabupaten/Kota di Provinsi Kalimantan Timur tahun 2015. Terdiri dari 1 variabel dependen (Y) yaitu jumlah penderita diare dan 5 variabel independen (X) yaitu jumlah rumah tangga berperilaku hidup bersih dan sehat, jumlah desa yang melaksanakan sanitasi total berbasis masyarakat, jumlah tempat pengelolaan makanan tidak memenuhi syarat, jumlah sarana jamban komunan, dan jumlah sarana jamban cemplung beserta letak koordinat (*latitude* dan *longitude*) untuk masing-masing wilayah.

**Analisis Deskriptif**

Penyebaran jumlah penderita diare dan faktor-faktor yang diduga mempengaruhinya di setiap Kabupaten/Kota, maka disajikan peta penyebaran untuk jumlah penderita diare dengan pembagian menjadi 5 klasifikasi, yaitu sangat tinggi, tinggi, sedang, rendah, dan sangat rendah.



Gambar 1. Persebaran Jumlah Penderita Diare

Pada Gambar 1 jumlah penderita diare sangat tinggi berkisar antara 10.367 sampai 17.913 orang yaitu Kota Samarinda, Balikpapan dan Kabupaten Kutai Kartanegara. Sedangkan untuk jumlah penderita diare sangat rendah berkisar antara 0 sampai 2116 orang, yaitu Kabupaten Mahakam Ulu.

**Pendeteksian Non Multikolinieritas**

Pendeteksian ini dilakukan untuk melihat adanya hubungan linier yang kuat diantara beberapa variabel independen dalam suatu model regresi. Jika nilai  $VIF_k < 10$ , maka mengindikasikan tidak terjadinya multikolinieritas antar variabel independen.

Tabel 1. Pendeteksian Non Multikolinieritas

Variabel	Nilai VIF
$X_1$	1,322
$X_2$	1,244
$X_3$	1,741
$X_4$	1,543
$X_5$	1,210

Berdasarkan Tabel 1 menunjukkan bahwa tidak terjadi multikolinieritas antar variabel independen pada data jumlah penderita diare di Provinsi Kalimantan Timur.

**Uji Heterogenitas Spasial**

Uji asumsi ini dilakukan untuk mengetahui apakah dalam model regresi terjadi ketidaksamaan varian dari *error* satu pengamatan ke pengamatan lainnya karena adanya efek spasial. Uji *Breusch-Pagan* (BP) merupakan salah satu uji yang digunakan untuk mengetahui terjadinya heterogenitas pada residual.

Tabel 2. Uji Heterogenitas Spasial

BP	Keputusan
0,0185	Gagal Menolak $H_0$

Berdasarkan Tabel 2 didapatkan nilai  $BP(0,0185) < \chi^2_{(0,1;5)}(9,236)$  maka diputuskan gagal menolak  $H_0$  dan dapat disimpulkan bahwa tidak terjadi heteroskedastisitas pada data jumlah penderita diare di Provinsi Kalimantan Timur.

**Model Awal Geographically Weighted regression**

Analisis dapat dilanjutkan ke model GWR apabila terjadinya kasus heterogenitas spasial yang mengindikasikan data jumlah penderita diare per Kabupaten/Kota di Provinsi Kalimantan Timur dipengaruhi oleh faktor lokasi pengamatan. Analisis regresi linier sebelumnya menunjukkan bahwa tidak terjadinya heterogenitas spasial pada data jumlah penderita diare per Kabupaten/Kota di Provinsi Kalimantan Timur tetapi analisis tetap dilanjutkan ke model GWR untuk melihat model regresi di tiap Kabupaten/Kota di Provinsi Kalimantan Timur. Model awal GWR untuk data jumlah penderita diare di Provinsi Kalimantan Timur adalah sebagai berikut:

$$\hat{y}_i = \hat{\beta}_0(u_i, v_i) + \hat{\beta}_1 X_1(u_i, v_i) + \hat{\beta}_2 X_2(u_i, v_i) + \hat{\beta}_3 X_3(u_i, v_i) + \hat{\beta}_4 X_4(u_i, v_i) + \hat{\beta}_5 X_5(u_i, v_i)$$

**Keterangan :**

- $\hat{y}$  : Jumlah Penderita Diare (Orang)
- $X_1$  : Jumlah Rumah Tangga Berperilaku Hidup Bersih dan Sehat (Rumah Tangga)
- $X_2$  : Jumlah Desa yang Melaksanakan Sanitasi Total Berbasis Masyarakat (Desa)
- $X_3$  : Jumlah Tempat Pengelolaan Makanan Tidak Memenuhi Syarat (Unit)
- $X_4$  : Jumlah Sarana Jamban Komunal (Unit)
- $X_5$  : Jumlah Sarana Jamban Cemplung (Unit)

**Estimasi Parameter Model Geographically Weighted Regression**

Estimasi model GWR untuk data jumlah penderita diare per Kabupaten/Kota di Provinsi Kalimantan Timur, yaitu :

$$\begin{aligned} \hat{y}_1 &= 978,5 + 0,283X_1 + 34,28X_2 + 4,109X_3 + 1,302X_4 + 0,115X_5 \\ \hat{y}_2 &= 1081,1 + 0,284X_1 + 30,57X_2 + 3,997X_3 + 1,408X_4 + 0,105X_5 \\ \hat{y}_3 &= 1140,9 + 0,280X_1 + 32,54X_2 + 3,990X_3 + 1,450X_4 + 0,098X_5 \\ \hat{y}_4 &= 948,27 + 0,278X_1 + 37,88X_2 + 4,088X_3 + 1,455X_4 + 0,093X_5 \\ \hat{y}_5 &= 1052,9 + 0,278X_1 + 36,60X_2 + 4,085X_3 + 1,387X_4 + 0,104X_5 \\ \hat{y}_6 &= 913,42 + 0,282X_1 + 36,08X_2 + 4,105X_3 + 1,380X_4 + 0,107X_5 \\ \hat{y}_7 &= 1182,3 + 0,282X_1 + 30,69X_2 + 3,997X_3 + 1,360X_4 + 0,109X_5 \\ \hat{y}_8 &= 851,96 + 0,280X_1 + 38,64X_2 + 4,139X_3 + 1,411X_4 + 0,101X_5 \\ \hat{y}_9 &= 1112,5 + 0,278X_1 + 34,98X_2 + 4,014X_3 + 1,469X_4 + 0,092X_5 \\ \hat{y}_{10} &= 1178,9 + 0,278X_1 + 34,03X_2 + 3,990X_3 + 1,467X_4 + 0,093X_5 \end{aligned}$$

Model GWR merupakan model yang tepat untuk menggambarkan kejadian diare di Provinsi Kalimantan Timur pada tahun 2015, karena mempunyai nilai R<sup>2</sup> sebesar 0,955 yang berarti bahwa kelima variabel independen (X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>, X<sub>4</sub>, X<sub>5</sub>) memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi jumlah penderita diare di Kalimantan Timur yaitu sebesar 95% sedangkan sisanya 5% dijelaskan oleh variabel lain yang tidak diketahui.

**Pengujian Signifikansi Parameter Model Geographically Weighted Regression**

Pengujian parsial model GWR dapat dilihat melalui prosedur pengujian:

Hipotesis

$$H_0 : \beta_k(u_i, v_i) = 0 ; k = 1, 2, 3, 4, 5 ,$$

$$H_1 : \text{paling tidak ada satu } \beta_k(u_i, v_i) \neq 0 ;$$

$$k = 1, 2, 3, 4, 5$$

Karena nilai  $t_{\frac{0,1}{2}; 4,00} = 2,776$  maka

diputuskan menolak H<sub>0</sub> dan dapat disimpulkan parameter yang signifikan adalah jumlah rumah tangga berperilaku hidup bersih dan sehat (X<sub>1</sub>), jumlah tempat pengelolaan makanan tidak memenuhi syarat (X<sub>3</sub>), dan jumlah sarana jamban komunal (X<sub>4</sub>). Dari semua hasil pengujian parameter diperoleh bahwa terdapat tiga variabel yang berpengaruh secara signifikan di beberapa wilayah Kabupaten/Kota.

Tabel 3. Parameter Signifikan dalam Model GWR

Kabupaten/Kota	Variabel Signifikan
Paser	X <sub>1</sub> , X <sub>3</sub>
Kutai Barat	X <sub>1</sub> , X <sub>3</sub>
Kutai Timur	X <sub>1</sub> , X <sub>3</sub> , X <sub>4</sub>
Berau	X <sub>1</sub> , X <sub>3</sub>
Penajam Paser Utara	X <sub>1</sub> , X <sub>3</sub>
Mahakam Ulu	X <sub>1</sub> , X <sub>3</sub>
Balikpapan	X <sub>1</sub> , X <sub>3</sub>
Samarinda	X <sub>1</sub> , X <sub>3</sub> , X <sub>4</sub>
Bontang	X <sub>1</sub> , X <sub>3</sub> , X <sub>4</sub>

**Model Mixed Geographically Weighted regression**

Berdasarkan analisis model GWR dengan pembobot adaptive kernel gaussian terdapat satu variabel prediktor yang berpengaruh secara lokal, yaitu X<sub>4</sub>. Sedangkan dua variabel prediktor berpengaruh secara global adalah X<sub>1</sub> dan X<sub>3</sub>. Oleh karena itu dengan variabel tersebut dibentuk model MGWR jumlah penderita diare dengan pembobot Adaptive Kernel Gaussian di Provinsi Kalimantan Timur tahun 2015 sebagai berikut:

$$\hat{y}_i = \hat{\beta}_0(u_i, v_i) + \hat{\beta}_1X_1 + \hat{\beta}_3X_3 + \hat{\beta}_4X_4(u_i, v_i)$$

**Estimasi Parameter Model Mixed Geographically Weighted Regression**

Estimasi model MGWR untuk data jumlah penderita diare per Kabupaten/Kota di Provinsi Kalimantan Timur, yaitu :

$$\begin{aligned} \hat{y}_1 &= 3741,9 + 0,269X_1 + 3,874X_3 + 1,048X_4 \\ \hat{y}_2 &= 3500,9 + 0,269X_1 + 3,874X_3 + 1,174X_4 \\ \hat{y}_3 &= 3510,7 + 0,269X_1 + 3,874X_3 + 1,206X_4 \\ \hat{y}_4 &= 3711,6 + 0,269X_1 + 3,874X_3 + 1,125X_4 \\ \hat{y}_5 &= 3701,04 + 0,269X_1 + 3,874X_3 + 1,108X_4 \\ \hat{y}_6 &= 3753,6 + 0,269X_1 + 3,874X_3 + 1,079X_4 \\ \hat{y}_7 &= 3372,1 + 0,269X_1 + 3,874X_3 + 1,173X_4 \\ \hat{y}_8 &= 3751,9 + 0,269X_1 + 3,874X_3 + 1,095X_4 \\ \hat{y}_9 &= 3615,4 + 0,269X_1 + 3,874X_3 + 1,179X_4 \\ \hat{y}_{10} &= 3482,1 + 0,269X_1 + 3,874X_3 + 1,245X_4 \end{aligned}$$

Model MGWR merupakan model yang tepat untuk menggambarkan jumlah penderita diare di Provinsi Kalimantan Timur pada tahun 2015, karena mempunyai nilai R<sup>2</sup> sebesar 0,878 yang berarti bahwa ketiga variabel independen (X<sub>1</sub>, X<sub>3</sub>, X<sub>4</sub>) memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variasi jumlah penderita diare di Kalimantan Timur yaitu sebesar 87% sedangkan sisanya 13% dijelaskan oleh variabel lain yang tidak diketahui.

**Pengujian Signifikansi Parameter Model Mixed Geographically Weighted Regression**

Pengujian kesesuaian model antara model MGWR dengan regresi global.

F <sub>1</sub>	Keputusan
0,599	Gagal Menolak H <sub>0</sub>

Berdasarkan Tabel 4 didapatkan nilai F<sub>1</sub>(0,599) < F<sub>(0,1;6,551;7,515)</sub>(2,83) maka diputuskan gagal menolak H<sub>0</sub> dan dapat disimpulkan bahwa model MGWR tidak berbeda model regresi global.

Pengujian serentak parameter prediktor global dilakukan untuk mengetahui signifikansi parameter prediktor global di seluruh Kabupaten/Kota di Provinsi Kalimantan Timur secara serentak terhadap variabel terikat.

Tabel 5. Uji Serentak Parameter Prediktor Global Model MGWR

$F_2$	Keputusan
16,419	Menolak $H_0$

Berdasarkan Tabel 5 diperoleh nilai  $F_2(16,419) > F_{(0,1;2;7,515)}(3,26)$  maka diputuskan menolak  $H_0$  dan dapat disimpulkan bahwa variabel prediktor global yang signifikan adalah jumlah rumah tangga yang berperilaku hidup bersih dan sehat ( $X_1$ ) dan jumlah tempat pengelolaan makanan tidak memenuhi syarat ( $X_3$ ).

Pengujian serentak parameter prediktor lokal dilakukan untuk mengetahui signifikansi parameter prediktor lokal di tiap Kabupaten/Kota di Provinsi Kalimantan Timur secara serentak terhadap variabel terikat.

Tabel 6. Uji Serentak Parameter Prediktor Lokal Model MGWR

$F_3$	Keputusan
5,579	Menolak $H_0$

Berdasarkan Tabel 6 diperoleh nilai  $F_3(5,579) > F_{(0,1;8,33;7,515)}(2,75)$  maka diputuskan menolak  $H_0$  dan dapat disimpulkan bahwa variabel prediktor lokal secara serentak berpengaruh terhadap pemodelan jumlah penderita diare di Provinsi Kalimantan Timur.

Pengujian parsial parameter prediktor global dilakukan untuk mengetahui signifikansi parameter prediktor global di seluruh Kabupaten/Kota di Provinsi Kalimantan Timur secara parsial terhadap variabel terikat.

Tabel 7. Uji Parsial Parameter Prediktor Global Model MGWR

Variabel	$ T_{g\_hit} $	Keputusan
$X_1$	5,335	Menolak $H_0$
$X_3$	4,350	Menolak $H_0$

Berdasarkan Tabel 7 didapatkan nilai  $|T_{g\_hit}| > t_{(0,1/2;7,515)}(2,365)$  maka diputuskan menolak  $H_0$  dan dapat disimpulkan parameter yang signifikan adalah jumlah rumah tangga yang berperilaku hidup bersih dan sehat ( $X_1$ ) dan jumlah tempat pengelolaan makanan tidak memenuhi syarat ( $X_3$ ).

Pengujian parsial parameter prediktor lokal dilakukan untuk mengetahui signifikansi parameter prediktor lokal di tiap Kabupaten/Kota di Provinsi Kalimantan Timur secara parsial terhadap variabel terikat. Dapat dilihat melalui prosedur pengujian:

Hipotesis

$$H_0 : \beta_k(u_i, v_i) = 0 ; k = 4 ; i = 1, 2, \dots, 10$$

$$H_1 : \beta_k(u_i, v_i) \neq 0$$

Karena nilai  $t_{(0,1/2;7,515)} = 2,365$  maka

diputuskan menolak  $H_0$  dan dapat disimpulkan parameter prediktor lokal ( $X_4$ ) signifikan di 6 Kabupaten/Kota yaitu Kutai Barat, Kutai Kartanegara, Kutai Timur, Mahakam Ulu, Samarinda, dan Bontang.

Dari semua hasil pengujian parameter diperoleh bahwa parameter global dan parameter lokal berpengaruh secara signifikan di tiap Kabupaten/Kota yang dapat disajikan pada Tabel 8.

Tabel 8. Parameter yang Signifikan dalam Model MGWR

Kabupaten/Kota	Variabel Signifikan
Paser	$X_1, X_3$
Kutai Barat	$X_1, X_3, X_4$
Kutai Kartanegara	$X_1, X_3, X_4$
Kutai Timur	$X_1, X_3, X_4$
Berau	$X_1, X_3$
Penajam Paser Utara	$X_1, X_3$
Mahakam Ulu	$X_1, X_3, X_4$
Balikpapan	$X_1, X_3$
Samarinda	$X_1, X_3, X_4$
Bontang	$X_1, X_3, X_4$

### Interpretasi Model Mixed Geographically Weighted Regression

Berikut adalah contoh model MGWR Kota Samarinda:

$$\hat{y}_9 = 3615,4 + 0,269X_1 + 3,874X_3 + 1,179X_4$$

Dari model tersebut dapat dijelaskan bahwa, jika jumlah rumah tangga yang berperilaku hidup bersih dan sehat, jumlah tempat pengelolaan makanan tidak memenuhi syarat, dan jumlah sarana jamban komunal bernilai 0 maka jumlah penderita diare di Kota Samarinda adalah sebanyak 3625 orang. Jika terjadi peningkatan jumlah rumah tangga yang berperilaku hidup bersih dan sehat sebanyak 1 rumah tangga, maka jumlah penderita diare akan bertambah sebanyak 1 orang. Selanjutnya jika terjadi penambahan sebanyak 1 unit tempat pengelolaan makanan tidak memenuhi syarat, maka jumlah penderita diare akan bertambah sebanyak 4 orang. Selanjutnya jika terjadi penambahan sarana jamban komunal 1 unit maka jumlah penderita diare akan bertambah sebanyak 1 orang.

### Kesimpulan

Berdasarkan analisis yang telah dilakukan, maka dapat diperoleh kesimpulan bahwa dengan menggunakan pemodelan MGWR terdapat faktor-faktor yang mempengaruhi jumlah penderita diare

di Provinsi Kalimantan Timur dimana faktor dibedakan menjadi faktor yang berpengaruh secara global di seluruh Kabupaten/Kota yaitu jumlah rumah tangga yang berperilaku hidup bersih dan sehat ( $X_1$ ) dan jumlah tempat pengelolaan makanan tidak memenuhi syarat ( $X_3$ ) serta faktor yang berpengaruh secara lokal di tiap Kabupaten/Kota yaitu jumlah sarana jamban komunal ( $X_4$ ).

#### Daftar Pustaka

- Cressie, N. A. C. (1991). *Statistics for Spatial Data*. USA: John Wiley & Sons.
- Depkes, R. I. (2005). *Pedoman Pemberantasan Penyakit Diare*. Jakarta: Ditjen PPM dan PL.
- \_\_\_\_\_. (2011). *Buku Saku Diare Edisi 2011*. Jakarta: Departemen Kesehatan RI.
- Fotheringham, A. S., Brunson, C., dan Charlton, M. (2002). *Geographically Weighted Regression the Analysis of Spatial Varying Relationship*. UK: John Wiley & Sons.
- Mei, C. L., Hes, Y., Fang, K. T. (2004). *A Note on the Mixed Geographically Weighted Regression Model*. *Journal of Regional Science*, 44, 143-157.
- Widoyono. (2011). *Penyakit Tropis*. Jakarta: Penerbit Erlangga.