ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 10, Nomor 2, Tahun 2021, Halaman 250-258
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

PEMODELAN *SEMIPARAMETRIC GEOGRAPHICALLY WEIGHTED REGRESSION* PADA KASUS PNEUMONIA BALITA PROVINSI JAWA TENGAH

Putri Fajar Utami[1], Agus Rusgiyono[2], Dwi Ispriyanti[3]
[1,2,3]Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
e-mail: putrifajarutami@students.undip.ac.id

ABSTRACT

Geographical and inter-regional differences have contributed to the diversity of child pneumonia cases in Central Java, so a spatial regression modelling is formed that is called Geographically Weighted Regression (GWR). GWR is a development of linear regression by involving diverse factors geographical location, so that local parameters are produced. Sometimes, there are non-local GWR parameters. To overcome some non-local parameters, Semiparametric Geographically Weighted Regression (SGWR) is formed to develop a GWR model with local and global influences simultaneously. SGWR Model is used to estimate the model of percentage of children with pneumonia in Central Java with population density, average temperature, percentage of children with severe malnutrition, percentage of children with under the red line weight, percentage of households behave in clean and healthy lives, and percentage of children who measles immunized. SGWR models on percentage of children with pneumonia in Central Java produce locally significant variables that is population density, average temperature, and percentage of households behave in clean and healthy lives. Variable that globally significant is percentage of children with severe malnutrition. Based on Akaike Information Criterion (AIC), SGWR is a better model to analize percentage of children with pneumonia in Central Java because of smallest AIC.

Keywords: Akaike Information Criterion, Geographically Weighted Regression, Semiparametric Geographically Weighted Regression

1. PENDAHULUAN

Menurut Prihaningtyas (2014), pneumonia merupakan infeksi yang disebabkan oleh virus, bakteri, jamur, atau infeksi campuran. Di Indonesia, pneumonia merupakan penyebab terbesar kematian pada balita. Angka balita penderita pneumonia di Jawa Tengah menduduki peringkat ke-3 terbanyak di Indonesia. Banyaknya kasus pneumonia balita pada setiap kabupaten/kota di Jawa Tengah sangat beragam.

Aspek lokasi (spasial) perlu diperhatikan dalam menentukan faktor yang signifikan mempengaruhi kasus pneumonia balita di suatu daerah. Berdasarkan uraian tersebut, perlu adanya suatu pemodelan statistik yaitu *Geographically Weighted Regression* (GWR). Pada model GWR, semua variabel prediktor diasumsikan bervariasi spasial dengan distribusi spasial parameter lokal berbeda untuk setiap lokasi pengamatan (Ma dan Gopal, 2018).

Pada saat pengujian parameter prediktor GWR terkadang terdapat variabel yang tidak mempunyai pengaruh lokasi, maka dikembangkan metode *Semiparametric Geographically Weighted Regression* (SGWR). Model SGWR merupakan pengembangan dari model GWR yang menggabungkan model lokal dan model global menjadi satu bentuk model. Model SGWR akan menghasilkan estimator parameter yang sebagian bersifat global dan yang lain bersifat lokal sesuai dengan lokasi pengamatan (Fotheringham *et al*, 2002).

2. TINJAUANPUSTAKA

2.1. Pneumonia

Menurut Hidayat (2008), pneumonia merupakan infeksi yang menyebabkan paru-paru meradang. Kantung-kantung kemampuan menyerap oksigen berkurang.Kekurangan oksigen membuat sel-sel tubuh tidak bisa bekerja. Penyebab dari pneumonia diketahui ada 30 sumber infeksi, dengan sumber utama bakteri, virus, mikroplasma, jamur, berbagai senyawa kimia maupun partikel. Pneumonia pada balita paling sering disebabkan oleh virus pernafasan dan puncaknya terjadi pada umur 2-3 tahun.Pneumonia merupakan penyebab utama mortalitas balita di Indonesia.

2.2. Model Regresi Linier

Analisis regresi merupakan analisis untuk mendapatkan hubungan dan model matematis antara variabel respon (Y) dan satu atau lebih variabel prediktor (X). Hubungan antara satu variabel respon dengan satu atau lebih variabel prediktor dapat dinyatakan dalam model regresi linier. Secara umum hubungan antara k variabel respon dengan p variabel prediktor dapat dinyatakan sebagai berikut (Gujarati, 2003):

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i, \quad = 1,2,3,\ldots,n \qquad (1)$$

Metode yang digunakan untuk menaksir parameter model regresi adalah dengan meminimumkan jumlah kuadrat error atau yang sering dikenal dengan *Ordinary Least Square* (OLS). Dalam OLS errornya diasumsikan identik, independen dan berdistribusi normal dengan mean nol dan varians konstan. Dengan taksiran sebagai berikut:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\ (\mathbf{X'Y})$$

dengan $\hat{\boldsymbol{\beta}}$ merupakan vektor dari parameter yang ditaksir berukuran (p+1)x1, nilai **X** merupakan matriks variabel prediktor berukuran nx(p+1) dan **Y** adalah vektor dari variabel respon berukuran (n x 1).

2.3. Model *Geographically Weighted Regression*

Model *Geographically Weighted Regression* (GWR) merupakan salah satu pengembangan metode regresi dengan mempertimbangkan unsur spasial di dalamnya. GWR digunakan untuk mengatasi heteroskedastisitas pada data. Model GWR menerapkan hubungan regresi spasial non-stasioner untuk masalah heteroskedastisitas dengan menghasilkan koefisien model regresi pada masing-masing lokasi. Model GWR dapat ditulis sebagai berikut (Fotheringham *et al.*, 2002):

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

dengan $\boldsymbol{y_i}$ merupakan nilai observasi variabel respon untuk lokasi ke-i, (u_i, v_i) adalah titik koordinat (*longitude*, *latitude*) lokasi i, $\beta_k(u_i,v_i)$ merupakan koefisien regresi variabel prediktor ke-k untuk lokasi ke-i, x_{ik} merupakan nilai observasi variabel prediktor ke-k pada pengamatan ke-i, dan ε_i merupakan *error* ke-i.

Estimator parameter model GWR untuk setiap lokasinya adalah:

$$\hat{\boldsymbol{\beta}}(u_i,v_i) = (\mathbf{X}^\mathbf{T}\mathbf{W}(u_i,v_i)\mathbf{X})^{-1}\mathbf{X}^\mathbf{T}\mathbf{W}(u_i,v_i)\mathbf{y} \qquad (3)$$

Dalam menentukan besarnya pembobot untuk masing-masing lokasi yang berbeda pada model GWR, diantaranya dengan menggunakan fungsi kernel (*kernel function*). Pembobot yang terbentuk dengan menggunakan fungsi kernel ini salah satunya fungsi Gaussian. Fungsi pembobot dapat ditulis sebagai berikut (Fotheringham *et al*, 2002):

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

dimana $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak Euclid antara lokasi (u_i, v_i) ke titik lokasi (u_j, v_j) dan h adalah parameter non negatif yang diketahui dan biasanya disebut parameter penghalus *(bandwidth)*.

Ada beberapa metode yang digunakan untuk memilih *bandwidth* optimum, salah satu diantaranya adalah metode *Cross Validation* (CV) yang secara matematis didefinisikan sebagai berikut:

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

dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir y_i dimana pengamatan di titik lokasi (u_i, v_i) tidak dilibatkan dari proses estimasi. Nilai h yang optimal diperoleh dari h yang menyebabkan nilai CV minimum.

Untuk pengujian kesesuaian model GWR, dapat dilakukan dengan hipotesis sebagai berikut (Leung dkk, 2000):

H_0: tidak ada perbedaan yang signfikan antara model regresi global dengan GWR

H_1:ada perbedaan yang signifikan antara model regresi global dengan GWR

dengan statistik uji

$$F_1 = \frac{(SSE(H_0) - SSE(H_1))/v_1}{SSE(H_1)/\delta_1}$$

dimana

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

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

$df_1 = \frac{v_1^2}{v_2}$ dengan $v_i = tr\left([(\mathbf{I-H}) - (\mathbf{I-L})^{\mathbf{T}}(\mathbf{I-L})]^i\right)$,

$df_2 = \frac{\delta_1^2}{\delta_2}$ dengan $\delta_i = tr\left([(\mathbf{I-L})^{\mathbf{T}}(\mathbf{I-L})]^i\right)$, i=1,2.

H_0 ditolak jika $\boldsymbol{F_1 \geq F_{\alpha,df_1,df_2}}$ pada taraf signifikansi sebesar α.

Selanjutnya, untuk mengidentifikasi variabel prediktor yang berpengaruh secara lokal dilakukan pengujian dengan hipotesis sebagai berikut:

H_0: $\beta_k(u_1, v_1) = \beta_k(u_2, v_2) = \ldots = \beta_k(u_n, v_n)$ (tidak ada perbedaan pengaruh yang signifikan dari variabel prediktor x_k antara satu lokasi dengan lokasi lainnya)

H_1:Minimal ada satu $\beta_k(u_i, v_i)$ yang berbeda (ada perbedaan pengaruh yang signifikan dari variabel prediktor x_k antara satu lokasi dengan lokasi lainnya)

dengan statistik uji

$$F_2(k) = \frac{V_k^2 / tr(\frac{1}{n}\boldsymbol{B}_k^T[\boldsymbol{I} - \frac{1}{n}\mathbf{J}]\boldsymbol{B_k}}{SSE(H_1)/\delta_1}$$

dimana

$$V_k^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_k(u_i, v_i) - \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_k(u_i, v_i)\right)^2 = \frac{1}{n}\hat{\boldsymbol{\beta}}_k^T[\boldsymbol{I} - \frac{1}{n}\mathbf{J}]\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$$

$$\hat{\boldsymbol{\beta}}_k(u_i, v_i) = \begin{bmatrix} \hat{\beta}_k(u_1, v_1) \\ \hat{\beta}_k(u_2, v_2) \\ \vdots \\ \hat{\beta}_k(u_n, v_n) \end{bmatrix} \text{ dan } \mathbf{B}_k = \begin{bmatrix} \mathbf{e}_k^T(\mathbf{X^TW}(u_1, v_1)\mathbf{X})^{-1}\mathbf{X^TW}(u_1, v_1) \\ \mathbf{e}_k^T(\mathbf{X^TW}(u_2, v_2)\mathbf{X})^{-1}\mathbf{X^TW}(u_2, v_2) \\ \vdots \\ \mathbf{e}_k^T(\mathbf{X^TW}(u_n, v_n)\mathbf{X})^{-1}\mathbf{X^TW}(u_n, v_n) \end{bmatrix}, k=1,2,,\ldots,p$$

Matriks $\boldsymbol{J}$ merupakan matriks berukuran n x n yang semua elemennya adalah 1 dan $\mathbf{e}_k$ adalah vektor kolom berukuran (p+1) yang bernilai 1 untuk elemen ke-k dan nol untuk lainnya. $F_2(k)$ akan mengikuti distribusi F dengan derajat bebas $df_1 = \frac{\gamma_1^2}{\gamma_2}$ dan $df_2 = \frac{\delta_1^2}{\delta_2}$ dengan $\gamma_i = tr(\frac{1}{n}\boldsymbol{B_k^T}\left[\boldsymbol{I} - \frac{1}{n}\mathbf{J}\right]\boldsymbol{B_k})^i$, i=1,2. H_0 ditolak jika $F_2(k) \geq F_{\alpha,df_1,df_2}$ pada taraf signifikansi α.

Pengujian signifikansi parameter dilakukan untuk mengetahui apakah parameter $\beta_k(u_i, v_i)$ signifikan mempengaruhi variabel responnya. Pengujiannya ialah sebagai berikut:
$H_0: \beta_k(u_i, v_i) = 0$
$H_1: \beta_k(u_i, v_i) \neq 0$ dengan k=1,2,...,p dan i=1,2,...,n
dengan statistik uji:

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

dimana c_{kk} adalah elemen diagonal ke-k dari matriks $\boldsymbol{C_i C_i^T}$ dengan $\boldsymbol{C_i} = (\boldsymbol{X^T W(u_i, v_i) X})^{-1}\boldsymbol{X^T W(u_i, v_i)}$. T akan mengikuti distribusi t dengan derajat bebas $df = \frac{\boldsymbol{\delta_1^2}}{\boldsymbol{\delta_2}}$ dan $\hat{\sigma}^2 = \frac{SSE(H_1)}{\delta_1}$. Jika diberikan tingkat signifikansi sebesar α, maka diambil keputusan dengan menolak H_0 jika $|T_{hitung}| \geq t_{\frac{\alpha}{2};df}$.

2.4. Model *Semiparametric Geographically Weighted Regression* (SGWR)

Pada model SGWR, beberapa koefisien pada model GWR diasumsikan konstan untuk seluruh lokasi pengamatan sedangkan yang lain bervariasi sesuai lokasi pengamatan data. Model SGWR dengan p variabel prediktor dan q variabel prediktor diantaranya bersifat lokal, dengan mengasumsikan bahwa intersep model bersifat lokal dapat dituliskan sebagai berikut (Fotheringham *et al*, 2002):

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{q} \beta_k (u_i, v_i) x_{ik} + \sum_{k=q+1}^{p} \beta_k \, x_{ik} + \varepsilon_i \quad , i = 1,2, \dots, n \tag{4}$$

Estimasi parameter pada model SGWR dapat dilakukan dengan metode WLS seperti halnya model GWR. Estimasi parameter untuk variabel global ialah

$$\widehat{\boldsymbol{\beta}}_g = [\mathbf{X}_g^T(\mathbf{I} - \mathbf{S}_l)^T(\mathbf{I} - \mathbf{S}_l)\mathbf{X}_g]^{-1}\mathbf{X}_g^T(\mathbf{I} - \mathbf{S}_l)^T(\mathbf{I} - \mathbf{S}_l)\mathbf{y} \tag{5}$$

Sedangkan estimasi parameter untuk variabel lokal ialah

$$\widehat{\boldsymbol{\beta}}_l(u_i, v_i) = (\mathbf{X}_l^T \mathbf{W}(u_i, v_i)\mathbf{X}_l)^{-1}\mathbf{X}_l^T \mathbf{W}(u_i, v_i)(\, \mathbf{y} - \mathbf{X}_g \boldsymbol{\beta}_g) \tag{6}$$

dimana $\mathbf{S}_l = \begin{bmatrix} \mathbf{x}_{l1}^T(\mathbf{X}_l^T \mathbf{W}(u_1, v_1)\mathbf{X}_l)^{-1}\mathbf{X}_l^T \mathbf{W}(u_1, v_1) \\ \mathbf{x}_{l2}^T(\mathbf{X}_l^T \mathbf{W}(u_2, v_2)\mathbf{X}_l)^{-1}\mathbf{X}_l^T \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_{ln}^T(\mathbf{X}_l^T \mathbf{W}(u_n, v_n)\mathbf{X}_l)^{-1}\mathbf{X}_l^T \mathbf{W}(u_n, v_n) \end{bmatrix}$.

Pengujian kesesuaian model SGWR dilakukan dengan hipotesis sebagai berikut (Purhadi danYasin, 2012):
H_0:tidak ada perbedaan yang signfikan antara model regresi global dengan SGWR
H_1:ada perbedaan yang signifikan antara model regresi global dengan SGWR
dengan statistik uji:

$$F_3 = \frac{(\frac{v_1 DDS_1}{v_2 \sigma^2}) / \frac{v_1^2}{v_2}}{(\frac{u_1 SSE(H_1)}{u_2 \sigma^2}) / \frac{u_1^2}{u_2}} = \frac{DDS_1/v_1}{SSE(H_1)/u_1}$$

dimana $\boldsymbol{S} = \boldsymbol{S}_l + (\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\left[\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)\boldsymbol{X}_g\right]^{-1}\boldsymbol{X}_g^T(\boldsymbol{I} - \boldsymbol{S}_l)^T(\boldsymbol{I} - \boldsymbol{S}_l)$,
$DDS_1 = \mathbf{y}^T[(\boldsymbol{I} - \boldsymbol{H}) - (\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})]\mathbf{y}$, $SSE(H_1) = \mathbf{y}^T(\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})\mathbf{y}$,
$u_i = tr\left(\left[(\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})^i\right]\right)$, dan $v_i = tr\left(\left[(\boldsymbol{I} - \boldsymbol{H}) - (\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})^i\right]\right)$, $i = 1,2$. H_0 ditolak jika $F_3 \geq F_{\alpha, df_1, df_2}$ pada taraf signifikansi sebesar α, dengan $df_1 = \frac{v_1^2}{v_2}$ dan $df_2 = \frac{u_1^2}{u_2}$.

Pengujian signifikansi parameter dilakukan dua kali, yang pertama ialah menguji signifikansi pada variabel global $x_k(q + 1 \leq k \leq p)$ yakni:

$H_0: \beta_k = 0$ (variabel global x_k tidak signifikan)
$H_1: \beta_k \neq 0$ (variabel global x_k signifikan)
dengan statistik uji:

$$T_{g\_hitung} = \frac{\hat{\beta}_k}{\hat{\sigma}\sqrt{g_{kk}}}$$

dengan g_{kk} adalah elemen diagonal ke-k dari matriks $\boldsymbol{GG^T}$ dan
$\boldsymbol{G} = \left[\boldsymbol{X_g^T}\left(\boldsymbol{I}-\boldsymbol{S_g}\right)^{\boldsymbol{T}}\left(\boldsymbol{I}-\boldsymbol{S_g}\right)\boldsymbol{X_g}\right]^{-\mathbf{1}}\boldsymbol{X_g^T}\left(\boldsymbol{I}-\boldsymbol{S_g}\right)^{\boldsymbol{T}}\left(\boldsymbol{I}-\boldsymbol{S_g}\right)$, dimana
$\boldsymbol{S_g} = \boldsymbol{X_g}(\boldsymbol{X_g}^{\boldsymbol{T}}\boldsymbol{X_g})^{-\mathbf{1}}\boldsymbol{X_g}^{\boldsymbol{T}}$ dan $\hat{\boldsymbol{\sigma}}^{\mathbf{2}} = \frac{\boldsymbol{y^T}(\boldsymbol{I}-\boldsymbol{S})^{\boldsymbol{T}}(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{y}}{\boldsymbol{tr}(((\boldsymbol{I}-\boldsymbol{S})^{\boldsymbol{T}}(\boldsymbol{I}-\boldsymbol{S}))}$. Pada signifikansi sebesar α, maka dapat mengambil keputusan tolak H_0 atau parameter β_k signifikan terhadap model jika $\left|T_{g\_hitung}\right| \geq t_{\frac{\alpha}{2};df}$, dimana $df = \frac{u_1^2}{u_2}$.

Selanjutnya, menguji signifikansi suatu variabel prediktor lokal $x_k (1 \leq k \leq q)$digunakan hipotesis sebagai berikut:
$H_0: \beta_k(u_i, v_i) = 0$ (variabel lokalx_k pada lokasi ke-i tidak signifikan)
$H_1: \beta_k(u_i, v_i) \neq 0$ (variabel lokal x_k pada lokasi ke-i signifikan)
dengan k =1,2,...,q dan i=1,2,...,n
statistik uji yang digunakan:

$$T_{l\_hitung} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma}\sqrt{m_{kk}}}$$

dengan m_{kk} adalah elemen diagonal ke-k dari matriks $\boldsymbol{M_i M_i}^{\boldsymbol{T}}$dan
$\boldsymbol{M_i} = (\boldsymbol{X_l^T W}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{X_l})^{-\mathbf{1}}\boldsymbol{X_l^T W}(\boldsymbol{u_i}, \boldsymbol{v_i})(\boldsymbol{I} - \boldsymbol{X_g G})$. Pada tingkat signifikansi α, maka dapat diambil keputusan menolak H_0 atau parameter $\beta_k(u_i, v_i)$ signifikan terhadap model jika $\left|T_{l\_hitung}\right| \geq t_{\frac{\alpha}{2};df}$, dimana $df = \frac{u_1^2}{u_2}$.

2.5. Pemilihan Model Terbaik

Ada beberapa metode yang digunakan untuk memilih model terbaik, salah satunya adalah Akaike Information Criterion (AIC) yang didefinisikan sebagai berikut (Fotheringham *et al*, 2002):

$$AIC_c = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n\left\{\frac{n + tr(\mathbf{S})}{n - 2 - tr(\mathbf{S})}\right\}$$

dengan

$\hat{\sigma}$: Nilai estimator standar deviasi dari error hasil estimasi maksimum likelihood, yaitu $\hat{\sigma}^2 = \frac{SSE}{n}$

S : Matriks proyeksi dimana $\hat{\mathbf{y}} = \mathbf{Sy}$

Pemilihan model terbaik dilakukan dengan menentukan model dengan nilai AIC terkecil.

3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder yang berasal dari Profil Kesehatan Provinsi Jawa Tengah tahun 2016, 2017, dan 2018. Variabel respon dalam penelitian ini adalah data persentase balita penderita pneumonia pada tahun 2016, 2017, dan 2018 di 35 kabupaten dan kota wilayah Provinsi Jawa Tengah. Variabel prediktor dalam penelitian ini adalah kepadatan penduduk (X1), suhu rata-rata (X2), persentase balita gizi buruk (X3), persentase balita BGM (X4), persentase rumah tangga ber-PHBS (X5), dan persentase balita imunisasi campak (X6).

Langkah-langkah yang dilakukan pada penelitian ini adalah:

1. Mengumpulkan data variabel respon dan variabel prediktor.
2. Menguji asumsi regresi linier.
3. Menganalisis model GWR.
4. Menganalisis model SGWR.
5. Menarik kesimpulan.

4. HASIL DAN PEMBAHASAN

4.1. Model Regresi Linier

Penaksiran nilai koefisien parameter dilakukan dengan menggunakan metode *Ordinary Least Square* (OLS). Model awal regresi yang terbentuk adalah:

$$\hat{y} = 8{,}206 + 0{,}0002X_1 - 0{,}256X_2 + 0{,}092X_3 - 0{,}080X_4 + 0{,}005X_5 - 0{,}001X_6$$

Asumsi pada model regresi linier tersebut yaitu normalitas, non autokorelasi, dan non multikolinieritas terpenuhi, sedangkan asumsi homoskedastisitas tidak terpenuhi. Asumsi homoskedastisitas tidak terpenuhi dengan nilai Uji *Breush-Pagan* sebesar 20,090 dengan sig=0,003 pada taraf signifikansi α=5%.

4.2. Model GWR Persentase Balita Penderita Pneumonia

Pengujian statistik kesesuaian model GWR dapat dilihat pada Tabel 1 berikut ini.

Tabel 1. Uji Kesesuaian Model GWR

Source	SS	df	MS	F_1
Improvement	301,003	8,732	8,831	3,520984
GWR	77,114	86,268	2,508	
Regresi	223,889	98		

Pada Tabel 1 tersebut, terlihat bahwa nilai F_1 (3,520984) $>F_{0,05;12,026;92,467}$ (1,859). Jadi, pada taraf signifikansi 5%, H_0 ditolak sehingga dapat disimpulkan bahwa ada perbedaan yang signifikan antara model regresi global dengan GWR.

Selanjutnya, dilakukan pengujian pengaruh lokasi secara parsial dengan hipotesis yaitu:
H_0: $\beta_k(u_1, v_1) = \beta_k(u_2, v_2) = ... = \beta_k(u_n, v_n)$, untuk k=0,1,2,...,6 (tidak ada perbedaan pengaruh yang signifikan dari variabel prediktor x_k antara satu lokasi dengan lokasi lainnya)
H_1:Minimal ada satu $\beta_k(u_i, v_i)$, untuk i=1,2,...,35 yang berbeda (ada perbedaan pengaruh yang signifikan dari variabel prediktor x_k antara satu lokasi dengan lokasi lainnya)
dimana nilai uji pengaruh lokasi dapat dilihat pada Tabel 2 berikut.

Tabel 2. Uji Pengaruh Lokasi pada Setiap Parameter GWR

Variabel	F_2	Df1	Df2	$F_{0,05;df1;df2}$
Intersep	11,293501	1,832	92,467	3,945
X_1	5,743534	1,856	92,467	3,945
X_2	3,740826	2,028	92,467	3,095
X_3	3,436318	1,133	92,467	3,945
X_4	-0,350320	1,205	92,467	3,945
X_5	4,362877	1,855	92,467	3,945
X_6	0,366153	1,105	92,467	3,945

Dengan taraf signifikansi 5%, maka menolak H_0 jika $F_2 \geq F_{0,05;df1;df2}$. Pada Tabel 2, dapat disimpulkan bahwa variabel prediktor X_1, X_2, dan X_5 ada perbedaan pengaruh yang signifikan dari variabel prediktor antara satu lokasi dengan lokasi yang lainnya atau dengan kata lain variabel prediktor tersebut bersifat lokal. Dikarenakan tidak semua variabel prediktor berpengaruh secara lokal, maka sebaiknya dibentuk model *Semiparametric* GWR.

Pengujian signifikansi parameter GWR pada setiap lokasi dilakukan dengan hipotesis:
$\boldsymbol{H_0}:\boldsymbol{\beta_k}(\boldsymbol{u_i},\boldsymbol{v_i})=\mathbf{0}$ untuk setiap k=1,2,...,6 dan i=1,2,...,35
$\boldsymbol{H_1}:\boldsymbol{\beta_k}(\boldsymbol{u_i},\boldsymbol{v_i})\neq\mathbf{0}$

Dengan menggunakan taraf signifikansi α =5%, maka menolak H_0 jika $|t_{hitung}| \geq t_{0,025;92,467}$(2,279). Ringkasan untuk faktor-faktor yang signifikan pada kabupaten/kota dapat dilihat pada Tabel 3.

Tabel 3. Faktor yang Signifikan Pada Kabupaten/kota Model GWR

Variabel	Kabupaten/kota
X_1	Wonosobo, Temanggung, Semarang, Purworejo, Pekalongan, Magelang, Kudus, Kota Semarang, Kota Pekalongan, Kota Magelang, Kendal, Kebumen, Jepara, Demak, Batang, dan Banjarnegara
X_2	Wonosobo, Temanggung, Purworejo, Purbalingga, Magelang, Kota Magelang, Kendal, Kebumen, Cilacap, Banyumas, dan Banjarnegara.
X_3	Semua kabupaten dan kota di Jawa Tengah
X_4	-
X_5	Semarang, Pekalongan, Magelang, Kota Semarang, Kota Salatiga, Kota Pekalongan, Kota Magelang, Klaten, Boyolali, dan Banjarnegara.
X_6	-

4.3. Model SGWR Persentase Balita Penderita Pneumonia

Pengujian kesesuaian model SGWR diperoleh nilai statistik uji sebagai berikut:

Tabel 4. Uji Kesesuaian Model SGWR

Source	SS	df	MS	F_3
Improvement	301,003	19,298	8,898	5,415944
SGWR	171,709	78,702	1,643	
Regresi	129,294	98		

Dengan menggunakan taraf signifikansi α sebesar 5%, H_0 ditolak karena nilai F_3 (5,415944) > $F_{0,05;25,011;83,371}$(1,639) sehingga disimpulkan bahwa ada perbedaan antara model SGWR dengan model regresi global.

Pengujian parsial parameter SGWR dilakukan dua kali, yakni pada variabel global dengan hipotesis:
$H_0: \beta_k = 0$ (variabel global x_k tidak signifikan)
$H_1: \beta_k \neq 0$ (variabel global x_k signifikan)
Pengujian dilakukan dengan kriteria menolak H_0 apabila nilai $|t_{hitung}|>t_{0,025;83,371}$(2,28). Hasil uji statistik parsial parameter variabel global terdapat pada Tabel 5.

Tabel 5. Uji Statistik Parsial Parameter Global SGWR

Variabel global	Beta	t_{hitung}	$t_{0,025;83,371}$
X_3	-0,069483	-2,30634	2,28
X_4	-0,062480	-0,452062	
X_6	0,045784	0,305059	

Berdasarkan Tabel 5, diperoleh bahwa dengan menggunakan tingkat signifikansi α sebesar 5%, H_0 ditolak pada variabel prediktor X_3 karena $|t_{hitung}|$ (2,30634)> $t_{0,025;83,371}$ (2,28). Jadi, dapat disimpulkan variabel prediktor global yang berpengaruh signifikan terhadap persentase balita penderita pneumonia adalah X_3 (persentase balita gizi buruk).

Pengujian signifikansi suatu variabel prediktor lokal $x_k(k=1,2,5)$ digunakan dengan hipotesis:

$H_0: \beta_k(u_i, v_i) = 0$ (variabel lokal x_k pada lokasi ke-i tidak signifikan)

$H_1: \beta_k(u_i, v_i) \neq 0$ (variabel lokal x_k pada lokasi ke-i signifikan)

Pengujian dilakukan dengan kriteria menolak H_0 apabila nilai $|t_{hitung}|>t_{0,025;83,371}$ (2,28).

Tabel 6. Faktor Lokal yang Signifikan Pada Kabupaten/kota Model SGWR

Variabel	Kabupaten/kota
X_1	Wonosobo, Temanggung, Semarang, Purworejo, Pekalongan, Magelang, Kudus, Kota Semarang, Kota Pekalongan, Kota Magelang, Kendal, Kebumen, Jepara, Demak, Batang, dan Banjarnegara
X_2	Wonosobo, Temanggung, Purworejo, Purbalingga, Magelang, Kota Magelang, Kendal, Kebumen, Cilacap, Banyumas, dan Banjarnegara.
X_5	Semarang, Pekalongan, Magelang, Kota Semarang, Kota Salatiga, Kota Pekalongan, Kota Magelang, Klaten, Boyolali, dan Banjarnegara.

4.4. Pemilihan Model yang Lebih Baik

Pemilihan model terbaik dilakukan dengan membandingkan nilai AIC pada model GWR dan SGWR. Model yang lebih baik ialah model dengan AIC terkecil. Nilai AIC pada model GWR sebesar 408,897 sedangkan AIC model SGWR sebesar 378,230. Berdasarkan nilai tersebut, diketahui bahwa nilai AIC terkecil ialah model SGWR. Jadi, dari hasil tersebut dapat disimpulkan bahwa model yang lebih baik untuk memodelkan data persentase balita penderita pneumonia yaitu model SGWR.

5. KESIMPULAN

Dari hasil dan pembahasan, diperoleh kesimpulan yaitu:

1. Model GWR yang diregresikan diperoleh variabel X_1(kepadatan penduduk), variabel X_2(suhu rata-rata), dan variabel X_5(persentase rumah tangga ber-PHBS) mempengaruhi secara lokal terhadap suatu wilayah di Jawa Tengah, sedangkan variabel X_3(persentase balita gizi buruk) mempengaruhi secara global untuk seluruh wilayah di Jawa Tengah. Ada beberapa variabel yang tidak bersifat global, maka dilakukan pemodelan secara lokal.
2. Berdasarkan model SGWR, faktor yang berpengaruh signifikan terhadap persentase balita penerita pneumonia di Jawa Tengah secara global adalah X_3 (persentase balita gizi buruk). Faktor yang berpengaruh signifikan secara lokal yaitu X_1 (kepadatan penduduk), X_2 (suhu rata-rata), dan X_5 (persentase rumah tangga ber-PHBS).
3. Model SGWR memiliki AIC lebih kecil dibandingkan dengan AIC GWR yaitu sebesar 378,230. Oleh karena itu, model SGWR lebih baik untuk memodelkan data persentase balita penderita pneumonia di Jawa Tengah.

DAFTAR PUSTAKA

Fotheringham, A.S., Brunsdon, C., dan Charlton, M. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.

Gujarati, D. N. 2003. *Basic Econometrics fourth edition*. New York: McGraw-Hill.

Hidayat, A. A. A. 2008. *Pengantar Ilmu Kesehatan Anak untuk Pendidikan Kebidanan*. Jakarta: Salemba Medika.

Leung, Y., Mei, C. L., dan Zhang, W. X. 2000. *Statistical Test for Spatial Nonstationary Based on The Geographically Weighted Regression Model*. Environment and Planning A Vol. 32, hal. 9-32.

Ma, Y. dan Gopal, S. 2018. *Geographically Weighted Regression Models in Estimating Median Home Prices in Towns of Sustainability Framework*. Sustainability, Vol. 10.

Mar'ah, Z., Djuraidah, A., dan Wigena, A. H. 2017. *Semi-parametric Geographically Weighted Regression Modelling using Linear Model of Coregionalization*. International Journal of Sciences: Basic and Applied Research, hal. 178-186.

Mei, C.L., Wang, N., dan Zhang, W. X. 2006. *Testing The Importance of the Explanatory Variables in A Mixed Geographically Weighted Regression Model*. Environment and Planning A Vol. 3, hal. 587-598.

Prihaningtyas, R. A. 2014. *Deteksi dan Cepat Obati 30+ Penyakit yang Sering Menyerang Anak*. Yogyakarta: Media Pressindo.

Purhadi dan Yasin, H. 2012. *Mixed Geographically Weighted Regression Model* (*Case Study: the Percentage of Poor Households in Mojokerto 2008*). European Journal of Scientific Research, Vol. 69, issue 2, hal. 188-196.