

PEMODELAN *MIXED GEOGRAPHICALLY WEIGHTED REGRESSION* DENGAN *ADAPTIVE BANDWIDTH* UNTUK ANGKA HARAPAN HIDUP (Studi Kasus : Angka Harapan Hidup di Jawa Tengah)

Rizki Faizatun Nisa¹, Sugito^{2*}, Arief Rachman Hakim³

^{1,2,3} Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

*email: sugitostat@gmail.com

ABSTRACT

Life expectancy at birth (AHH) is an estimate of the years a person will take from birth. AHH is used as an indicator of public health and welfare. These two indicators are of concern to the government in relation to human development. It is hoped that the AHH value will continue to increase so that the quality of human development will also increase. Modeling of the factors that influence AHH needs to be done so that efforts to increase AHH become more effective. The AHH value for Central Java (Central Java) in 2020 is 74.37. Factors thought to influence AHH in Central Java are the percentage of poor people (X_1), the percentage of households with proper sanitation (X_2), the percentage of children under five who are fully immunized (X_3) and the open unemployment rate (X_4). The assumption of homoscedasticity in AHH modeling in Central Java using linear regression was not fulfilled, meaning that there was spatial heterogeneity between districts/cities, so the Geographically Weighted Regression (GWR) method was used. The weighting function used is the Bisquare and Tricube kernels with adaptive bandwidth. The GWR method will encounter problems if not all independent variables are local, so the Mixed Geographically Weighted Regression (MGWR) method is used. The results of the GWR analysis for the two weighting functions are that the X_1 variable is not local, so the MGWR method is used. The results of MGWR modeling for the two weighting functions are that local variables and global variables have a significant effect. The best model is the MGWR model with Kernel Tricube weighting because it has the smallest AICc value.

Keyword : AHH, GWR, MGWR, *Adaptive Kernel Bisquare*, *Adaptive Kernel Tricube*, AICc

1. PENDAHULUAN

Angka harapan hidup saat lahir (AHH) merupakan rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. Nilai AHH dapat mencerminkan derajat kesehatan suatu masyarakat (BPS, 2021). AHH juga dapat dijadikan sebagai indikator kesejahteraan masyarakat (BPS, 2021). Kesehatan dan kesejahteraan merupakan 2 faktor penting yang diperhatikan oleh pemerintah terutama kaitannya dengan pembangunan manusia. Semakin tinggi nilai AHH maka akan semakin tinggi pula derajat kesehatan dan kesejahteraan masyarakat di suatu wilayah yang pada akhirnya akan meningkatkan kualitas pembangunan manusia di wilayah tersebut. Pemodelan faktor-faktor yang mempengaruhi AHH menjadi penting untuk dilakukan karena dalam usaha meningkatkan nilai AHH menjadi lebih efektif.

Jawa Tengah merupakan salah satu provinsi di Indonesia yang memiliki nilai AHH cukup tinggi pada tahun 2020 yaitu 74,37 tahun, namun nilai AHH untuk masing-masing kabupaten/kota di Jawa Tengah memiliki *range* yang besar. Terdapat beberapa faktor yang diduga mempengaruhi AHH di Jawa Tengah diantaranya adalah persentase penduduk miskin, persentase rumah tangga dengan sanitasi layak, persentase balita yang diimunisasi lengkap dan tingkat pengangguran terbuka. Menurut Halicioglu (2011), tinggi rendahnya nilai harapan hidup dipengaruhi oleh beberapa hal diantaranya adalah ekonomi, sosial dan lingkungan. Berdasarkan hal tersebut *range* yang besar dari AHH antar kabupaten/kota di Jawa Tengah kemungkinan disebabkan oleh perbedaan karakteristik wilayah antar kabupaten/kota yang menjadikan adanya perbedaan ekonomi, sosial atau lingkungan. Hal ini ditunjukkan dengan tidak terpenuhinya asumsi homoskedastisitas dalam pemodelan faktor-faktor yang mempengaruhi AHH di Jawa Tengah. Pelanggaran asumsi ini

mengindikasikan bahwa diduga terdapat heterogenitas spasial antar Kabupaten/Kota di Jawa Tengah, oleh karena itu dibutuhkan metode pemodelan yang juga memperhatikan faktor lokasi / wilayah.

Geographically Weighted Regression (GWR) adalah pengembangan dari metode regresi yang dalam pemodelannya dipengaruhi oleh faktor lokasi dimana terdapat matriks pembobot dalam estimasi parameternya (Fotheringham, et al, 2002). Besaran bobot ditentukan oleh kedekatan jarak antar lokasi. Semakin dekat jarak, maka nilai bobot pengaruhnya akan semakin besar. Salah satu cara pemberian bobot dalam metode GWR adalah menggunakan fungsi kernel, dimana fungsi kernel memberikan pembobot sesuai bandwidth optimum. Besaran nilai bandwidth bisa berubah-ubah untuk setiap lokasi atau dikenal dengan *Adaptive Bandwidth*. Terdapat beberapa macam fungsi kernel yang dapat digunakan dalam pembobotan GWR, diantaranya adalah fungsi kernel *bisquare* dan fungsi kernel *tricube*. Hasil pemodelan dari metode GWR adalah setiap lokasi memiliki model regresinya tersendiri atau disebut dengan model regresi lokal yang tersusun atas faktor-faktor yang mempengaruhi variabel dependen atau disebut dengan variabel independen yang juga bersifat lokal. Hal ini akan menimbulkan permasalahan apabila terdapat variabel independen yang tidak bersifat lokal tetapi diduga memiliki pengaruh terhadap variabel dependen secara global sehingga digunakan metode *Mixed Geographically Weighted Regression* (MGWR) untuk menangani permasalahan tersebut (Fotheringham, et al, 2002). Berdasarkan latar belakang tersebut maka tujuan penelitian kali ini adalah untuk menganalisis bagaimana model yang terbentuk, faktor-faktor apa saja yang berpengaruh signifikan, dan manakah yang merupakan model terbaik untuk pemodelan faktor-faktor yang mempengaruhi AHH di Jawa Tengah menggunakan *Mixed Geographically Weighted Regression* dengan *Adaptive Kernel Bisquare* dan *Adaptive Kernel Tricube*.

2. TINJAUAN PUSTAKA

2.1. Regresi Linier

Analisis regresi berkenaan dengan studi ketergantungan satu variabel tak bebas (Y) pada satu atau lebih variabel bebas (X). Menurut Draper & Smith (1992), model regresi dapat dituliskan dalam bentuk matriks sebagai berikut :

$$Y = X\beta + \epsilon \quad (2.1)$$

Y : matriks variabel dependen (nx1), X : matriks variabel independen (nx(p+1)), β : vektor dari parameter regresi ((p+1)x1) dan ϵ : matriks dari error (nx1), error berdistribusi normal dengan $E(\epsilon):0$ dan $Var(\epsilon):I\sigma^2$ serta elemen dari ϵ tidak berkorelasi.

2.1.2. Estimasi Parameter Regresi Linier

Estimasi parameter regresi menggunakan metode OLS atau *Ordinary Least Square*.

Berdasarkan metode OLS estimasi parameter regresi linier dapat dirumuskan dengan

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2.2)$$

2.1.3. Uji Signifikansi Parameter Regresi Linier

a. Uji Serentak

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal terdapat } \beta_k \neq 0 ; k = 1, 2, 3, \dots, p$$

$$F_{hitung} = \frac{MSE}{MSR} = \frac{ESS/p}{RSS/(n-p-1)} \quad (2.3)$$

Jika nilai $F_{hitung} > F_{tabel} (F_{\alpha;p;n-p-1})$ maka H_0 ditolak.

b. Uji Parsial

$$H_0 : = 0 \text{ (variabel } X_k \text{ tidak signifikan terhadap variabel dependen)}$$

$$H_1 : \neq 0 \text{ (variabel } X_k \text{ signifikan terhadap variabel dependen)}$$

$$t_{hitung} = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)} \quad (2.4)$$

dengan $se(\hat{\beta}_k)$ merupakan akar dari $var(\hat{\beta}_k)$, $varcov(\hat{\beta}) = (\mathbf{X}^T \mathbf{X})^{-1} \hat{\sigma}^2$ $var(\hat{\beta}_k) = varcov(\hat{\beta})[k, k]$ dengan $\hat{\sigma}^2 = \frac{RSS}{n-p-1}$. Jika $|t_{hitung}| > t_{\alpha/2; n-p-1}$ maka H_0 ditolak.

2.1.4. Uji Asumsi Klasik

1. Asumsi Normalitas

Uji normalitas yang sering digunakan adalah uji Lilliefors, hipotesisnya adalah

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

$$D = \max_x |S_n(X) - F^*(X)| \quad (2.5)$$

Daerah penolakan H_0 jika nilai $D > D_{tabel}$ (Tabel Lilliefors dengan derajat bebas (n, α)).

2. Asumsi Multikolinieritas

$$VIF_k = \frac{1}{1-R_k^2}; k=1,2,3 \dots p \quad (2.6)$$

Ketika nilai $VIF_k > 10$ akan mengakibatkan koefisien regresi k dianggap lemah atau tidak akurat karena terdapat masalah multikolinieritas (Shrestha, 2020).

3. Asumsi Homoskedastisitas

Uji yang digunakan adalah uji Glejser. Prinsipnya adalah meregresikan nilai absolut dari residual dengan variabel independennya. Hipotesisnya adalah

$H_0 : \beta_k = 0$

$H_1 : \beta_k \neq 0$

dengan nilai k adalah $k = 1, 2, 3, \dots, p$

$$t_{hitung} = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)} \quad (2.7)$$

Jika $|t_{hitung}| > t_{\alpha/2; n-p-1}$ maka H_0 ditolak. Apabila tidak terdapat parameter yang signifikan terhadap variabel dependen (harga mutlak residual) maka asumsi homoskedastisitas terpenuhi.

4. Asumsi Non-Autokorelasi

Uji yang digunakan adalah uji Durbin-Watson dengan hipotesisnya adalah

H_0 : Tidak terjadi autokorelasi

H_1 : Terjadi autokorelasi

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (2.8)$$

e_i merupakan *error* ke i . Pengambilan keputusan dilakukan berdasarkan kriteria berikut :

Tabel 1. Daerah Penolakan Durbin Watson

Nilai d	Keputusan	Kesimpulan
$d < dL$	Menolak H_0	Terdapat Autokorelasi
$d > 4-dL$	Menolak H_0	Terdapat Autokorelasi
$dU < d < 4-dU$	Tidak Menolak H_0	Tidak Terdapat Autokorelasi
$dL \leq d \leq dU$	Daerah Ragu-ragu	Pengujian tidak meyakinkan
$4-dU \leq d \leq 4-dL$	Daerah Ragu-ragu	Pengujian tidak meyakinkan

2.2. Geographically Weighted Regression (GWR)

Model GWR dinotasikan seperti persamaan (2.14) (Caraka & Yasin, 2017)

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon_i; \quad i = 1, 2, 3, \dots, n \quad (2.9)$$

dengan y_i variabel dependen pada lokasi ke i , (u_i, v_i) merupakan titik koordinat lokasi ke i dan $\beta_k(u_i, v_i)$ nilai parameter β ke k untuk lokasi ke i yang memiliki titik koordinat (u_i, v_i) , x_{ik} variabel independen ke k pada lokasi ke i dan ε_i merupakan nilai *error* pada lokasi ke- i .

2.2.1 Estimasi Parameter GWR

Estimasi parameter model GWR menggunakan metode *Weighted Least Square* (WLS).

Berdasarkan metode WLS estimasi parameter GWR dirumuskan dengan

$$\hat{\beta}(\mathbf{u}_i, \mathbf{v}_i) = (\mathbf{X}^T \mathbf{W}(\mathbf{u}_i, \mathbf{v}_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{u}_i, \mathbf{v}_i) \mathbf{Y} \quad (2.10)$$

Jika \mathbf{C} merupakan notasi untuk $(\mathbf{X}^T \mathbf{W}(\mathbf{u}_i, \mathbf{v}_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{u}_i, \mathbf{v}_i)$ maka $\hat{\beta}(\mathbf{u}_i, \mathbf{v}_i)$ dapat dinotasikan dengan (Fotheringham *et al.*, 2002)

$\hat{\beta}(\mathbf{u}_i, \mathbf{v}_i) = \mathbf{C} \mathbf{y}$, sehingga varian untuk estimasi parameter dinotasikan dengan

$$\text{Var}(\hat{\beta}(\mathbf{u}_i, \mathbf{v}_i)) = \mathbf{C} \mathbf{C}^T \sigma^2 \quad (2.11)$$

dengan $\sigma^2 = \sum_{i=1}^n (y_i - \hat{y}_i) / (n - 2v_1 + v_2)$, $v_1 = \text{tr}(\mathbf{S})$ dan $v_2 = 2\text{tr}(\mathbf{S}^T \mathbf{S})$. \mathbf{S} adalah *hat matrix* untuk memetakan \hat{y} ke y atau dirumuskan dengan $\hat{y} = \mathbf{S} \mathbf{y}$, setiap barisnya dinotasikan dengan $r_i = (\mathbf{X}_i^T \mathbf{W}(\mathbf{u}_i, \mathbf{v}_i) \mathbf{X})^{-1} \mathbf{X}_i^T \mathbf{W}(\mathbf{u}_i, \mathbf{v}_i)$ (Fotheringham *et al.*, 2002).

2.2.2. Pembobotan GWR

Fungsi pembobot yang digunakan adalah dengan fungsi kernel. Fungsi kernel memberikan pembobot sesuai *bandwidth* optimum. Untuk memperoleh *bandwidth* yang optimum digunakan nilai *Cross Validation* (CV) yang minimum, secara matematis nilai CV dapat didefinisikan sebagai berikut

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2 \quad (2.12)$$

Nilai *bandwidth* dalam fungsi kernel dapat bernilai sama untuk setiap wilayah (*fixed bandwidth*). Namun tak jarang di sebuah titik amatan memiliki penyebaran titik yang lebih jarang dibandingkan titik amatan yang lainnya. Permasalahan ini sangat memungkinkan terdapat adanya masalah pembentukan parameter jika nilai *bandwidth* yang digunakan sama besar karena akan menghasilkan standar *error* yang besar dikarenakan model dikalibrasi oleh titik data yang sangat sedikit, sehingga akan lebih baik jika *bandwidth* yang digunakan disesuaikan dengan padat tidaknya sebaran titik di sekitar amatan. Nilai *bandwidth* yang berbeda untuk setiap amatan inilah yang disebut dengan *adaptive bandwidth*. Fungsi *adaptive kernel tricube* dan *adaptive kernel bisquare* dirumuskan dengan :

1. *Adaptive kernel bisquare* (Fotheringham *et al.*, 2002)

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_i}\right)^2\right]^2 & ; d_{ij} \leq b_i \\ 0 & ; d_{ij} > b_i \end{cases} \quad (2.13)$$

2. *Adaptive kernel tricube* (LeSage, 2001)

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_i}\right)^3\right]^3 & ; d_{ij} \leq b_i \\ 0 & ; d_{ij} > b_i \end{cases} \quad (2.14)$$

2.2.3. Uji Hipotesis GWR

1. Pengujian Pengaruh Lokasi Secara Parsial

$H_0 : \beta_{1k} = \beta_{2k} = \beta_{3k} = \dots = \beta_{nk} \quad k = 1, 2, \dots, p$

$H_1 : \text{tidak semua } \beta_{ik} \text{ bernilai sama, } k = 1, 2, \dots, p$

$$F_3(k) = \frac{v_k^2 / \text{tr}\left(\frac{1}{n} \mathbf{B}_k^T \left[\mathbf{I} - \frac{1}{n} \mathbf{J}\right] \mathbf{B}_k\right)}{RSS_g / \delta_1} \quad (2.15)$$

$$\text{dengan } \mathbf{B}_k = \begin{bmatrix} \mathbf{e}_k^T (\mathbf{X}^T \mathbf{W}(\mathbf{u}_1, \mathbf{v}_1) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{u}_1, \mathbf{v}_1) \\ \mathbf{e}_k^T (\mathbf{X}^T \mathbf{W}(\mathbf{u}_2, \mathbf{v}_2) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{u}_2, \mathbf{v}_2) \\ \vdots \\ \mathbf{e}_k^T (\mathbf{X}^T \mathbf{W}(\mathbf{u}_n, \mathbf{v}_n) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{u}_n, \mathbf{v}_n) \end{bmatrix}$$

$$V_k^2 = \frac{1}{n} \hat{\beta}_k(\mathbf{u}_i, \mathbf{v}_i)^T \left(\mathbf{I} - \frac{1}{n} \mathbf{J}\right) \hat{\beta}_k(\mathbf{u}_i, \mathbf{v}_i) \quad (2.16)$$

\mathbf{J} : matriks berukuran $n \times n$ yang semua elemennya 1 dan adalah vektor kolom berukuran $(p+1)$ yang bernilai 1 untuk elemen ke- k dan nol untuk lainnya. F_3 akan mengikuti distribusi F dengan derajat bebas adalah $df_1 = \frac{\gamma_1^2}{\gamma_2}$ dan $df_2 = \frac{\delta_1^2}{\delta_2}$ dengan nilai $\gamma_i = \text{tr}\left(\frac{1}{n} \mathbf{B}_k^T \left[\mathbf{I} - \frac{1}{n} \mathbf{J}\right] \mathbf{B}_k\right)^i$ dan $\delta_i = \text{tr}((\mathbf{I} - \mathbf{L})^T (\mathbf{I} - \mathbf{L}))^i$

\mathbf{I} : matriks identitas berukuran $n \times n$ serta \mathbf{L} : matriks proyeksi dari model GWR, sehingga matriks proyeksinya adalah:

$$\mathbf{L} = \begin{bmatrix} x_1^T (X^T W(u_1, v_1) X)^{-1} X^T W(u_1, v_1) \\ x_2^T (X^T W(u_2, v_2) X)^{-1} X^T W(u_2, v_2) \\ \vdots \\ x_n^T (X^T W(u_n, v_n) X)^{-1} X^T W(u_n, v_n) \end{bmatrix}$$

H_0 ditolak jika $F_3 > F_{\alpha; df_1; df_2}$ atau jika nilai p -value $< \alpha$.

2. Pengujian Parsial Signifikan Parameter Model

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

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

dengan $i=1,2,\dots,n$ dan $k=1,2,\dots,p$

$$t_{hitung} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma} \sqrt{C_{kk}}} \quad (2.17)$$

dengan $\hat{\beta}_k(u_i, v_i)$ merupakan penaksir parameter dari $\beta_k(u_i, v_i)$ dan nilai C_{kk} adalah elemen diagonal ke- k dari matriks $\mathbf{C}\mathbf{C}^T$ dengan $\mathbf{C} = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i)$ dan $\hat{\sigma} = \frac{RSS_g}{\delta_1}$.

H_0 ditolak apabila nilai $|t_{hitung}| > t_{\frac{\alpha}{2}, df}$ atau p -value $< \alpha$ dengan nilai $= \frac{\delta_1^2}{\delta_2}$.

2.3. Mixed Geographically Weighted Regression (MGWR)

Model *MGWR* dapat dinotasikan sebagai berikut (Fotheringham *et al.*, 2002)

$$y_i = \sum_{k=1}^{pa} a_k x_{ik} + \sum_{k=pa+1}^{pb} b_k(u_i, v_i) x_{ik} + \varepsilon_i \quad ; i=1,2,\dots,n \quad (2.18)$$

dengan y_i adalah variabel independen, (u_i, v_i) menunjukkan lokasi pengamatan ke i , a_k menunjukkan parameter yang bersifat global untuk variabel ke k sedangkan b_k menunjukkan parameter yang bersifat lokal untuk variabel ke k .

2.3.1 Estimasi Parameter MGWR

Model umum *MGWR* dituliskan dalam bentuk matriks sebagai berikut :

$$\mathbf{y} = \mathbf{X}_g \boldsymbol{\beta}_g + \mathbf{X}_l \boldsymbol{\beta}_l(u_i, v_i) + \boldsymbol{\varepsilon} \quad (2.19)$$

dengan \mathbf{X}_g adalah matriks variabel global, \mathbf{X}_l adalah matriks variabel lokal, $\boldsymbol{\beta}_g$ adalah vektor dari parameter global dan $\boldsymbol{\beta}_l(u_i, v_i)$ matriks dari parameter lokal.

Berdasarkan persamaan (2.19) misal $\tilde{\mathbf{y}}$ mendefinisikan $\mathbf{y} - \mathbf{X}_g \boldsymbol{\beta}_g$ maka

$$\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{X}_g \boldsymbol{\beta}_g = \mathbf{X}_l \boldsymbol{\beta}_l(u_i, v_i) + \boldsymbol{\varepsilon} \quad (2.20)$$

Jika diperhatikan persamaan tersebut menjadi seperti pada model *GWR*, sehingga seperti estimasi pada model *GWR* langkah pertama yang harus dilakukan adalah mendefinisikan *error*, berdasarkan model (2.20) maka *error* dapat didefinisikan dengan

$$\boldsymbol{\varepsilon} = \tilde{\mathbf{y}} - \mathbf{X}_l \boldsymbol{\beta}_l(u_i, v_i) \quad (2.21)$$

dan fungsi kuadrat *error* terbobotnya jika dinotasikan dengan R adalah sebagai berikut

$$R = \boldsymbol{\varepsilon}^T W(u_i, v_i) \boldsymbol{\varepsilon} = \tilde{\mathbf{y}}^T W(u_i, v_i) \tilde{\mathbf{y}} - 2 \boldsymbol{\beta}_l^T(u_i, v_i) X_l^T W(u_i, v_i) \tilde{\mathbf{y}} + \boldsymbol{\beta}_l^T(u_i, v_i) X_l^T W(u_i, v_i) X_l \boldsymbol{\beta}_l(u_i, v_i) \quad (2.22)$$

Seperti halnya pada model *GWR* memaksimumkan fungsi dilakukan dengan cara menurunkan fungsi tersebut terhadap parameternya yaitu $\boldsymbol{\beta}_l^T(u_i, v_i)$ dan disamadengankan nol, sehingga hasil estimasi nya menjadi

$$\hat{\boldsymbol{\beta}}_l(u_i, v_i) = (X_l^T W(u_i, v_i) X_l)^{-1} X_l^T W(u_i, v_i) \tilde{\mathbf{y}} \quad (2.23)$$

Jika $x_{il}^T = (1, x_{i1}, x_{i2}, \dots, x_{ipb})$ adalah elemen ke- i dari matriks \mathbf{X}_l , maka nilai prediksi untuk $\tilde{\mathbf{y}}$ pada (u_i, v_i) dapat dinotasikan sebagai berikut (Purhadi & Yasin, 2012):

$$\hat{\tilde{\mathbf{y}}}_i = x_{il}^T \hat{\boldsymbol{\beta}}_l(u_i, v_i) = x_{il}^T (X_l^T W(u_i, v_i) X_l)^{-1} X_l^T W(u_i, v_i)$$

Sehingga secara umum dapat dituliskan dengan :

$$\hat{\tilde{\mathbf{y}}} = (\hat{\tilde{\mathbf{y}}}_1, \hat{\tilde{\mathbf{y}}}_2, \dots, \hat{\tilde{\mathbf{y}}}_n) = \mathbf{S}_l \tilde{\mathbf{y}} \text{ dengan nilai } S_l \text{ adalah}$$

$$S_l = \begin{pmatrix} x_{l1}^T (X_l^T W(u_1, v_1) X_l)^{-1} X_l^T W(u_1, v_1) \\ x_{l2}^T (X_l^T W(u_2, v_2) X_l)^{-1} X_l^T W(u_2, v_2) \\ \vdots \\ x_{ln}^T (X_l^T W(u_n, v_n) X_l)^{-1} X_l^T W(u_n, v_n) \end{pmatrix}$$

Kemudian, substitusikan elemen dari $\widehat{\beta}_l(\mathbf{u}_i, \mathbf{v}_i)$ ke dalam model MGWR maka persamaannya menjadi

$$\widehat{\beta}_g = [X_g^T(I - S_l)^T(I - S_l)X_g]^{-1}X_g^T(I - S_l)^T(I - S_l)y \quad (2.24)$$

$$\text{dengan } S_g = X_g(X_g^T X_g)^{-1}X_g^T \quad (2.25)$$

dengan mensubstitusikan $\widehat{\beta}_g$ kedalam (2.23) maka akan menghasilkan estimasi koefisien lokal pada lokasi (u_i, v_i) yaitu

$$\widehat{\beta}_l(\mathbf{u}_i, \mathbf{v}_i) = [X_l^T W(\mathbf{u}_i, \mathbf{v}_i) X_l]^{-1} X_l^T W(\mathbf{u}_i, \mathbf{v}_i) (y - X_g \widehat{\beta}_g) \quad (2.26)$$

$\widehat{\beta}_g$ dan $\widehat{\beta}_l(\mathbf{u}_i, \mathbf{v}_i)$ merupakan estimator tak bias untuk estimator β_g dan $\beta_l(\mathbf{u}_i, \mathbf{v}_i)$.

Sehingga nilai \widehat{y} untuk seluruh n lokasi adalah sebagai berikut (Purhadi & Yasin, 2012) :

$\widehat{y} = S y$ dengan nilai S diperoleh dari

$$S = S_l + (I - S_l)X_g[X_g^T(I - S_l)^T(I - S_l)X_g]^{-1}X_g^T(I - S_l)^T(I - S_l) \quad (2.27)$$

Estimasi parameter σ^2 dapat dituliskan:

$$\widehat{\sigma}^2 = \frac{y^T(I-S)^T(I-S)y}{tr((I-S)^T(I-S))} \quad (2.28)$$

2.3.2 Uji Signifikansi Parameter MGWR

1. Uji Parameter Secara Serentak

- Variabel Global

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{pa} = 0$$

$$H_1 : \text{terdapat } \beta_k \neq 0, k=1,2,3,\dots,pa$$

$$F_2 = \frac{y^T[(I-S_l)^T(I-S_l)-(I-S)^T(I-S)]y/r_1}{y^T(I-S)^T(I-S)y/u_1} \quad (2.29)$$

$u_i = r[(I - S)^T(I - S)]^i$; $r_i = tr([(I - S_l)^T(I - S_l) - (I - S)^T(I - S)]^i)$ untuk $i=1,2$; $df_1 = \frac{r_1^2}{r_2}$ dan $df_2 = \frac{u_1^2}{u_2}$. Apabila nilai $F_2 > F_{\alpha;df_1;df_2}$ atau $p\text{-value} < \alpha$ maka H_0 ditolak.

- Variabel Lokal

$$H_0 : \beta_{pa+1}(u_i, v_i) = \beta_{pa+2}(u_i, v_i) = \dots = \beta_{pb}(u_i, v_i) = 0$$

$$H_1 : \text{minimal terdapat } \beta_k(u_i, v_i) \neq 0, k=(pa+1),(pa+2),(pa+3)\dots,pb$$

$$F_3 = \frac{y^T[(I-S_g)^T(I-S_g)-(I-S)^T(I-S)]y/t_1}{y^T(I-S)^T(I-S)y/u_1} \quad (2.30)$$

$t_i = tr([(I - S_g)^T(I - S_g) - (I - S)^T(I - S)]^i)$ untuk $i=1,2$ dengan $df_1 = \frac{t_1^2}{t_2}$ dan $df_2 = \frac{u_1^2}{u_2}$.

Apabila nilai $F_3 > F_{\alpha;df_1;df_2}$ maka H_0 ditolak.

2. Uji Parameter Secara Parsial

- Variabel Global

$$H_0 : \beta_k = 0 \text{ (parameter global } \beta_k \text{ tidak berpengaruh signifikan)}$$

$$H_1 : \beta_k \neq 0 \text{ (parameter global } \beta_k \text{ berpengaruh signifikan)}$$

$$T_g = \frac{\widehat{\beta}_k}{\widehat{\sigma}\sqrt{g_{kk}}} \quad (2.31)$$

dengan g_{kk} merupakan elemen diagonal ke k dari matriks GG^T . $G = [X_g^T(I - S_l)^T(I - S_l)X_g]^{-1}X_g^T(I - S_l)^T(I - S_l)$. Apabila nilai $|T_{g_{hitung}}| > t_{\alpha/2;df}$ atau $p\text{-value} < \alpha$ maka H_0 ditolak,

dengan $df = \frac{u_1^2}{u_2}$.

- Variabel Lokal

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

$$H_1 : \beta_k(u_i, v_i) \neq 0, k=(pa+1),(pa+2),(pa+3)\dots,pb$$

$$T_l = \frac{\widehat{\beta}_k(u_i, v_i)}{\widehat{\sigma}\sqrt{m_{kk}}} \quad (2.32)$$

dengan m_{kk} merupakan elemen diagonal ke k dari matriks MM^T . $M =$

$[X_1^T W(u_i, v_i) X_1]^{-1} X_1^T W(u_i, v_i) (I - X_g G)$. Apabila nilai $|T_{hitung}| > t_{\alpha/2; df}$ atau nilai $p\text{-value} < \alpha$ maka H_0 ditolak, dengan $df = \frac{u_1^2}{u_2}$.

2.4 Pemilihan Model Terbaik

$$AIC_c = 2n \ln(\hat{\sigma}) + n \ln(2\pi) + n \left\{ \frac{n+tr(S)}{n-2-tr(S)} \right\} \quad (2.33)$$

dengan subskrip c menandakan koreksi, sehingga rumus tersebut menghasilkan nilai AIC_c yang telah dikoreksi, $\hat{\sigma}$ merupakan nilai estimator standar deviasi dari *error* hasil estimasi maksimum likelihood, yaitu $\hat{\sigma}^2 = \frac{RSS}{n}$ dan S merupakan matriks proyeksi dengan $\hat{y} = Sy$. Model terbaiknya adalah model yang memiliki nilai AIC_c terkecil (Fotheringham *et al.*, 2002).

3. METODE PENELITIAN

Data yang digunakan dalam penelitian kali ini bersumber dari *website* BPS Provinsi Jawa Tengah, *website* SIRUSA BPS, buku Profil Kesehatan Provinsi Jawa Tengah 2020 dan buku Provinsi Jawa Tengah dalam Angka 2021. Terdapat dua jenis variabel yang digunakan dalam penelitian ini, yaitu variabel dependen dan variabel independen. Variabel dependen yang digunakan adalah Angka Harapan Hidup di Provinsi Jawa Tengah untuk setiap Kabupaten dan/Kota tahun 2020 (y). Sedangkan untuk variabel independen yang digunakan adalah Persentase Penduduk Miskin (X_1), Persentase Rumah Tangga dengan Sanitasi Layak (X_2), Persentase Balita yang diimunisasi Lengkap (X_3) dan Tingkat Pengangguran Terbuka (X_4) untuk setiap Kabupaten dan atau Kota di Jawa Tengah tahun 2020.

4. HASIL DAN PEMBAHASAN

Analisis data menggunakan metode regresi linier menghasilkan kesimpulan bahwa berdasarkan uji signifikansi parameter secara simultan variabel independen secara bersama-sama memiliki pengaruh terhadap variabel AHH. Pengujian signifikansi parameter secara parsial ternyata hanya 3 variabel independen yang signifikan yaitu X_1 , X_2 dan X_4 . Pada uji asumsi klasik ternyata terdapat pelanggaran asumsi homoskedastisitas sedangkan asumsi yang lain terpenuhi. Hal ini berarti varian dari residual tidak konstan yang mengindikasikan diduga terdapat heterogenitas spasial dalam pembentukan model AHH di Jawa Tengah, sehingga pemodelan dilakukan dengan metode GWR.

Estimasi parameter model GWR digunakan metode WLS yang prinsipnya adalah meminimumkan jumlah kuadrat *error* terboboti. Berdasarkan hasil estimasi parameter, model GWR yang terbentuk menggunakan fungsi pembobot *adaptive kernel bisquare* dan *adaptive kernel tricube* untuk masing-masing Kabupaten dan atau Kota berbeda-beda begitu juga untuk signifikansi parameter untuk setiap modelnya. Uji asumsi normalitas residual untuk metode GWR dengan kedua fungsi pembobot tersebut memberikan kesimpulan bahwa residual berdistribusi normal. Selanjutnya, pada uji pengaruh lokasi secara parsial dengan kedua fungsi pembobot tersebut dapat diketahui ternyata pada taraf signifikansi α sebesar 5% terdapat variabel independen yang bersifat lokal dan bersifat global, dengan hasil ujinya dijelaskan pada Tabel 2.

Tabel 2. Uji Pengaruh Lokasi Secara Parsial

Fungsi Kernel	Parameter	F_3	$P\text{-value}$	Keputusan
<i>Bisquare</i>	<i>Intercept</i>	7,8234	6,93E-06	H_0 ditolak
	X_1	1,2802	0,2843307	H_0 diterima
	X_2	4,3989	0,0009367	H_0 ditolak
	X_4	8,7486	2,38E-06	H_0 ditolak
<i>Tricube</i>	<i>Intercept</i>	8,0949	5,05E-06	H_0 ditolak
	X_1	1,3479	0,2466546	H_0 diterima
	X_2	4,6225	0,0006472	H_0 ditolak

$$\frac{X_4}{9,005} = 1,64E-06 \quad H_0 \text{ ditolak}$$

Karena terdapat variabel yang bersifat lokal dan global maka digunakan pemodelan AHH menggunakan metode *Mixed Geographically Weighted Regression* (MGWR).

Variabel lokal yang digunakan dalam metode MGWR adalah X_2 dan X_4 sedangkan variabel globalnya adalah variabel X_1 baik untuk fungsi pembobot *adaptive kernel bisquare* maupun *adaptive kernel tricube*. Berdasarkan hasil estimasi parameter MGWR dengan fungsi pembobot *adaptive kernel bisquare* nilai estimasi parameter untuk variabel X_1 adalah -0,2145432, sedangkan estimasi parameter menggunakan fungsi pembobot *adaptive kernel tricube* adalah -0,2170699. Estimasi parameter selanjutnya adalah estimasi parameter untuk variabel lokal, dimana kedua fungsi pembobot memberikan nilai estimasi yang berbeda-beda untuk setiap Kabupaten dan atau Kota.

Terdapat beberapa uji signifikansi parameter di metode MGWR, yaitu :

1. Uji Signifikansi Parameter Model MGWR Secara Serentak

Uji signifikansi parameter secara serentak pada model MGWR dalam penelitian kali ini hanya untuk variabel lokal saja karena variabel global yang digunakan hanya 1. Uji signifikansi parameter variabel lokal secara serentak diterapkan pada model MGWR AHH di Jawa Tengah dengan *adaptive kernel bisquare* dan *tricube*. Hasil uji difokan melalui Tabel 3. Kesimpulan yang dapat diambil adalah pada taraf signifikansi α sebesar 5% uji parameter simultan variabel lokal untuk kedua fungsi pembobot adalah signifikan.

Tabel 3. Uji Signifikansi Parameter Variabel Lokal Model MGWR Secara Serentak

Fungsi Pembobot	F _{Hitung}	F _{Tabel}	Keputusan
<i>Adaptive Kernel Bisquare</i>	2859,15907961848	1,81232662015801	H ₀ ditolak
<i>Adaptive Kernel Tricube</i>	3254,65369564082	1,82331704103359	H ₀ ditolak

2. Uji Signifikansi Parameter Model MGWR Secara Parsial

Tujuan dari uji ini adalah mengetahui signifikansi parameter untuk masing-masing model MGWR AHH di Jawa Tengah dengan fungsi pembobot *adaptive kernel bisquare* dan *adaptive kernel tricube* di setiap wilayah. Berdasarkan hasil uji dapat diketahui bahwa parameter yang signifikan terhadap model di setiap wilayah berbeda-beda. Berdasarkan parameter yang signifikan terhadap model, Kabupaten dan atau Kota di Jawa Tengah dapat diklasifikasikan menjadi 3 baik untuk fungsi pembobot *adaptive kernel bisquare* maupun *adaptive kernel tricube* yang ditunjukkan pada Tabel 4 dan Tabel 5.

Tabel 4. Klasifikasi Kab./Kota Berdasarkan Signifikansi Parameter (MGWR *Bisquare*)

Beta Nol	Beta Nol dan Beta X_2	Beta Nol dan Beta X_4
Kab. Banjarnegara, Kab. Kebumen, Kab. Purworejo, Kab. Wonosobo, Kab. Rembang, Kab. Pati, Kab. Kudus, Kab. Jepara, Kab. Demak, Kab. Kendal dan Kab. Batang	Kab. Magelang, Kab. Boyolali, Kab. Klaten, Kab. Sukoharjo, Kab. Wonogiri, Kab. Karanganyar, Kab. Sragen, Kab. Grobogan, Kab. Blora, Kab. Semarang, Kab. Temanggung, Kota Magelang, Kota Surakarta, Kota Salatiga dan Kota Semarang	Kab. Cilacap, Kab. Banyumas, Kab. Purbalingga, Kab. Pekalongan, Kab. Pemalang, Kab. Tegal, Kab. Brebes, Kota Pekalongan dan Kota Tegal

Tabel 5. Klasifikasi Kab./Kota Berdasarkan Signifikansi Parameter (MGWR *Tricube*)

Beta Nol	Beta Nol dan Beta X_2	Beta Nol dan Beta X_4
Kab. Banjarnegara, Kab. Kebumen, Kab. Purworejo, Kab. Wonosobo, Kab. Rembang, Kab. Pati, Kab. Kudus, Kab. Jepara, Kab.	Kab. Magelang, Kab. Boyolali, Kab. Klaten, Kab. Sukoharjo, Kab. Wonogiri, Kab. Karanganyar, Kab. Sragen, Kab. Grobogan, Kab. Blora,	Kab. Cilacap, Kab. Banyumas, Kab. Purbalingga, Kab. Pekalongan, Kab. Pemalang, Kab. Tegal, Kab. Brebes,

Demak, Kab. Temanggung, Kab. Kendal dan Kab. Batang	Kab. Semarang, Kota Magelang, Kota Surakarta, Kota Salatiga dan Kota Semarang	Kota Pekalongan dan Kota Tegal
--	--	-----------------------------------

Jika diambil contoh di Kabupaten Pemalang hasil uji signifikansi parameter secara parsial untuk variabel global adalah pada taraf signifikansi α sebesar 5% variabel global X_1 berpengaruh terhadap model AHH di Kabupaten Pemalang menggunakan metode MGWR dengan fungsi pembobot *adaptive kernel bisquare* dan *adaptive kernel tricube*. Sedangkan hasil uji signifikansi parameter secara parsial untuk variabel lokal dijelaskan pada Tabel 6.

Tabel 6. Uji Signifikansi Parameter Secara Parsial Variabel Lokal Kab. Pemalang

		Beta Lokal	Thitung	Keputusan
Adaptive Kernel <i>Bisquare</i>	<i>Intercept</i>	79,55949982	47,90861956	ditolak
	X_2	0,015580409	0,914270567	diterima
	X_4	-0,65128741	-5,258676052	ditolak
Adaptive Kernel <i>Tricube</i>	<i>Intercept</i>	79,60703673	48,2923278	ditolak
	X_2	0,015960217	0,94367961	diterima
	X_4	-0,65938716	-5,3371799	ditolak

Pemilihan model terbaik dilakukan dengan membandingkan nilai AICc. Berdasarkan hasil pengujian, dapat diketahui nilai AICc dari model MGWR dengan fungsi pembobot *adaptive kernel bisquare* adalah 99,666390 sedangkan untuk pembobot *adaptive kernel tricube* adalah 98,26476, sehingga model terbaiknya adalah model MGWR dengan fungsi pembobot *adaptive kernel tricube* karena memiliki nilai AICc yang terkecil.

4. KESIMPULAN

Berdasarkan analisis dan pembahasan, kesimpulannya adalah pemodelan AHH di Jawa Tengah menggunakan Regresi Linier diduga terdapat heterogenitas spasial sehingga pemodelan dilanjutkan menggunakan metode GWR dengan fungsi pembobot *adaptive kernel bisquare* dan *tricube*. Berdasarkan pemodelan GWR diketahui terdapat variabel independen yang tidak bersifat lokal sehingga pengujian dilanjutkan menggunakan metode *Mixed Geographically Weighted Regression* (MGWR).

Pemodelan AHH di Jawa Tengah menggunakan MGWR dengan fungsi pembobot *adaptive kernel bisquare* dan *adaptive kernel tricube* memberikan hasil variabel persentase penduduk miskin signifikan mempengaruhi AHH di Jawa Tengah secara global, sedangkan variabel yang mempengaruhi secara lokal adalah persentase rumah tangga yang memiliki sanitasi layak dan tingkat pengangguran terbuka. Model terbaiknya adalah model MGWR dengan fungsi pembobot *adaptive kernel tricube* karena memiliki nilai AICc terkecil.

DAFTAR PUSTAKA

- BPS. (2011). *ANGKA KEMATIAN BAYI DAN ANGKA HARAPAN HIDUP PENDUDUK INDONESIA Hasil Sensus Penduduk 2010*. Jakarta: Badan Pusat Statistik.
- BPS. (2021). *Angka Harapan Hidup*. Retrieved Oktober 14, 2021, from Badan Pusat Statistik Sirusa: <https://sirusa.bps.go.id/sirusa/index.php/indikator/48>
- BPS. (2021). *Angka Harapan Hidup Saat Lahir*. Retrieved September 29, 2021, from SIRUSA Badan Pusat Statistik: <https://sirusa.bps.go.id/sirusa/index.php/indikator/1060>
- BPS. (2021). *Kemiskinan dan Ketimpangan*. Retrieved Juli 2021, 6, from BADAN PUSAT STATISTIK: <https://www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html>

- BPS. (2021). *PROFIL KESEHATAN PROVINSI JAWA TENGAH 2020*. Badan Pusat Statistik Provinsi Jawa Tengah.
- BPS. (2021). *PROVINSI JAWA TENGAH DALAM ANGKA Jawa Tengah Province in Figures 2021*. Badan Pusat Statistik Provinsi Jawa Tengah.
- BPS. (2021). *SIRUSA*. Retrieved Desember 23, 2021, from Garis Kemiskinan: <https://sirusa.bps.go.id/sirusa/index.php/indikator/50>
- BPS. (2021). *SIRUSA*. Retrieved Desember 23, 2021, from Pengeluaran Per Kapita: <https://sirusa.bps.go.id/sirusa/index.php/indikator/197>
- BPS. (2021). *SIRUSA*. Retrieved Desember 23, 2021, from Angkatan Kerja: <https://sirusa.bps.go.id/sirusa/index.php/indikator/47>
- BPS. (2021). *SIRUSA*. Retrieved Oktober 14, 2021, from Tingkat Pengangguran Terbuka (TPT): <https://sirusa.bps.go.id/sirusa/index.php/indikator/44>
- BPS. (2021). *Tingkat Pengangguran Terbuka (TPT)*. Retrieved Oktober 14, 2021, from Badan Pusat Statistik Sirusa: <https://sirusa.bps.go.id/sirusa/index.php/indikator/44>
- Caraka, R. E., & Yasin, H. (2017). *Geographically weighted regression (gwr) Sebuah Pendekatan Regresi Geografis*. Yogyakarta: Mobius.
- Draper, N., & Smith, H. (1966). *Applied Regression Analysis (Second Edition)*. (B. Sumantri, Trans.) John Willey & Sons.
- Draper, N., & Sumantri, H. S. (1992). *Analisis Regresi Terapan Edisi Kedua*. Jakarta: Gramedia Pustaka Utama.
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *GEOGRAPHICALLY WEIGHTED REGRESSION the analysis of spatially varying relationship*. Chichester: John Wiley and sons.
- Gujarati, D. (1978). *BASIC ECONOMETRICS*. (S. Zein, Trans.) New York: McGraw-Hill.
- Gujarati, D. N. (2003). *Basic Econometrics Fourth Edition*. New York: McGraw-Hill.
- Gujarati, D., & Zein, S. (1978). *EKONOMETRIKA DASAR*. New York: McGraw-Hill.
- Halicioglu, F. (2011). Modeling life expectancy in Turkey. *Elsevier*, 2750-2082.
- LeSage, J. P. (2001). A Family of Geographically Weighted Regression Models. *Departement of Economics University of Toledo*.
- Leung, Y., Mei, C.-L., & Zang, W.-X. (2000). Statistical tests for Spatisal nonstationarity based on the geographically weighted regression model. *Environment and Palnning A*, 9-32.
- Purhadi, & Yasin, H. (2012). Mixed Geographically Weighted Regression Model (Case Study: the Percentage of Poor Haouseholds in Mojokerto 2008). *European Journal of Scientific Research*, 188-196.
- Razali, N. M., & Yap, B. w. (2011). Power Comparisons of Shapiro-wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-darling Tests. *Journal of Statistical Modeling and Analytics*, 21-33.
- Shrestha, N. (2020). Detecting Multicollinearity in Regression Analysis. *American Journal of Applied Mathematics and Statistics*, 39-42.
- Widayaka, P. G. (2016). Pendekatan Mixed Geographically Weighted Regression Untuk Pemodelan Pertumbuhan Ekonomi Menurut Kabupaten/Kota Di Jawa Tengah. *Jurnal gaussian*, 727-736.