
Pemodelan Indeks Pembangunan Manusia (IPM) Metode Baru Menurut Provinsi Tahun 2015 Menggunakan *Geographically Weighted Regression (GWR)*

Akbar Maulana¹, Renny Meilawati², and Vita Widiastuti³
^{1,2,3} Program Studi Statistika, Universitas Sebelas Maret, Surakarta

akbar.maulana2298@gmail.com, rennymeilawati@gmail.com, vitawidi28@gmail.com

Abstract. The Human Development Index (HDI) is a parameter of quality of life for an area. The HDI explains how residents can access the results of development in obtaining income, health and education. One method that can be used to find out the factors that influence the human development index in modeling is regression analysis of ordinary least square (OLS). In the Human Development Index data, there is a dependency between measuring data and the location of a region. Therefore, spatial regression analysis can be used in this study. The local form of spatial regression analysis is *geographically weighted regression (GWR)*. GWR shows the existence of spatial heterogeneity (location). This study compares between OLS regression and GWR in the new human development index method by province in 2015. In the GWR model we use fixed Gaussian kernel and kernel fixed bisquare as weighted function. The optimal bandwidth value is obtained by minimizing the cross validation (CV) and Akaike information criterion (AIC) coefficients. The results showed that the GWR model with Gaussian kernel function is better than GWR with bisquare kernel function and OLS model.

Keywords: human development index, ordinary least square, geographically weighted regression, kernel fixed Gaussian, kernel fixed bisquare

1. Pendahuluan

Pembangunan manusia didefinisikan sebagai proses perluasan pilihan bagi penduduk. Tujuan utama pembangunan adalah menciptakan lingkungan yang memungkinkan rakyat untuk menikmati umur panjang, sehat, dan menjalankan kehidupan yang produktif [1]. Indeks pembangunan manusia (IPM) menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya. IPM diperkenalkan oleh UNDP pada tahun 1990 dan dipublikasikan secara berkala dalam laporan tahunan Human Development Report (HDR). IPM merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). Selain itu, IPM juga dapat menentukan peringkat atau level pembangunan suatu wilayah/negara. Bagi Indonesia, IPM merupakan data strategis karena selain sebagai ukuran kinerja Pemerintah, IPM juga digunakan sebagai salah satu pertimbangan penentuan Dana Alokasi Umum (DAU). Salah satu metode yang dapat digunakan untuk mengetahui faktor-faktor yang berpengaruh terhadap indeks pembangunan manusia dan memodelkannya adalah analisis regresi *ordinary least square (OLS)*.

Analisis regresi merupakan analisis statistika yang digunakan untuk mengetahui hubungan antara variabel dependen dengan variabel independen. Ada beberapa metode yang dapat digunakan untuk mengestimasi fungsi regresi, salah satunya adalah OLS. OLS merupakan metode estimasi fungsi regresi yang paling sering digunakan. Kriteria OLS adalah "line of best fit" atau dengan kata lain jumlah kuadrat dari deviasi antara titik-titik observasi dengan garis regresi adalah minimum. Dalam model regresi linear memiliki beberapa asumsi dasar yang harus dipenuhi untuk menghasilkan estimasi yang BLUE, yaitu normalitas, homoskedastisitas, non-multikolinearitas dan non-autokorelasi [2]. Pada data IPM, terdapat ketergantungan antara pengukuran data dengan lokasi. Oleh karena itu, analisis regresi spasial dapat digunakan dalam penelitian ini. Bentuk lokal dari analisis regresi spasial adalah *geographically weighted regression* (GWR).

Geographically weighted regression (GWR) adalah salah satu analisis yang membentuk analisis regresi namun bersifat lokal untuk setiap lokasi. Hasil analisis ini adalah model regresi yang nilai-nilai parameternya berlaku hanya pada tiap lokasi pengamatan, dan berbeda dengan lokasi lainnya. Dalam GWR digunakan unsur matriks pembobot yang besarnya tergantung pada kedekatan antar lokasi. Semakin dekat suatu lokasi, bobot pengaruhnya akan semakin besar.

Penelitian ini membandingkan antara regresi *ordinary least square* (OLS) dan *geographically weighted regression* (GWR) pada data Indeks Pembangunan Manusia metode baru menurut provinsi tahun 2015. Pembobot yang digunakan dalam model GWR adalah kernel *fixed* Gaussian dan kernel *fixed* bisquare. Nilai bandwidth optimal diperoleh dengan meminimumkan nilai koefisien *cross validation* (CV) dan *Akaike Information Criterion* (AIC).

2. Model Regresi *Ordinary Least Square* (OLS)

2.1 Model Regresi Linear. Model regresi linear merupakan model tentang hubungan antara variabel dependen y dan variabel independen x_1, x_2, \dots, x_p . Model regresi linier untuk p variabel independen secara umum ditulis dengan

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$$

Jika diambil sebanyak n pengamatan, maka model untuk pengamatan ke- i adalah

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \varepsilon_i \quad (1)$$

dengan $i = 1, 2, \dots, n$; $\beta_0, \beta_1, \dots, \beta_p$ adalah parameter model dan $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ adalah eror yang diasumsikan identik, independen, dan berdistribusi Normal dengan mean nol dan variansi konstan σ^2 atau ($\varepsilon_i \sim IIDN(0, \sigma^2)$). Pada model ini, hubungan antara variabel prediktor dan variabel respon dianggap konstan pada setiap lokasi pengamatan. Jika persamaan (1) dituliskan dalam notasi matriks maka persamaan (1) menjadi

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Nilai estimasi untuk \mathbf{y} dan $\boldsymbol{\varepsilon}$ adalah: $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ dan $\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$. Estimator untuk parameter $\boldsymbol{\beta}$ dan σ^2 adalah

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

dan

$$s^2 = \frac{1}{n-p-1} \sum_{i=1}^n (y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}})^2 \quad (3)$$

dengan $\hat{\boldsymbol{\beta}}$ dan s^2 merupakan estimator yang takbias untuk $\boldsymbol{\beta}$ dan σ^2 jika matriks kovariansi $\text{cov}(\mathbf{y}) = \sigma^2 \mathbf{I}$ [2].

Untuk menguji kesesuaian model regresi linear OLS digunakan analisis dengan hipotesisnya sebagai berikut

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

$$H_1 : \text{minimal ada satu } \beta_i \neq 0$$

Daerah kritisnya adalah menolak H_0 bila $F_{hitung} > F_{\alpha, p, (n-p-1)}$ atau nilai $p < \alpha$. Adapun pengujian secara parsial untuk mengetahui parameter mana saja yang signifikan terhadap model dilakukan dengan hipotesis:

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0$$

Statistik uji dalam pengujian parsial ini adalah $t = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)}$ dengan $se(\hat{\beta}_k) = s\sqrt{g_{kk}}$

adalah standar error dari koefisien $\hat{\beta}_k$. Sedangkan g_{kk} adalah elemen diagonal ke- k dari matriks $(\mathbf{X}^T \mathbf{X})^{-1}$ dan $s = \sqrt{MSE}$. Di bawah H_0 t akan mengikuti distribusi t dengan derajat bebas $(n-p-1)$ sehingga jika diberikan tingkat signifikansi sebesar α , maka diambil keputusan tolak H_0 jika $|t_{hit}| > t_{\alpha/2; n-p-1}$ [3].

2.2 Asumsi-Asumsi Regresi Linear Berganda. Metode Kuadrat Terkecil dapat dilakukan apabila asumsi regresi linear klasik terpenuhi. Beberapa asumsi yang harus dipenuhi oleh persamaan regresi linear berganda ini sebagai berikut:

1. Normalitas, regresi linear klasik mengasumsikan bahwa tiap ε_i mengikuti distribusi normal, $\varepsilon_i \sim N(0, \sigma^2)$.
2. Non autokorelasi antar sisaan, berarti $cov(\varepsilon_i, \varepsilon_j) = 0$
3. Homoskedastisitas, $var(\varepsilon_i) = \sigma^2$ untuk setiap i , $i = 1, 2, \dots, n$ yang artinya varians dari semua sisaan adalah konstan atau homoskedastik.
4. Tidak terjadi multikolinearitas. Tidak terdapat hubungan linear yang sempurna atau pasti diantara variabel.

Untuk mengetahui apakah model persamaan yang digunakan sudah memenuhi asumsi-asumsi regresi tersebut maka perlu dilakukan pemeriksaan pada masing-masing asumsi.

2.3 Aspek Data Spasial. Analisis spasial dilakukan jika data yang digunakan memenuhi aspek spasial yaitu memiliki sifat error yang saling berkorelasi atau memiliki heterogenitas spasial [5]. Heterogenitas spasial dapat dideteksi dengan melakukan uji Breusch-Pagan.

Hipotesis :

$$H_0 : \sigma_1 = \sigma_2 = \dots = \sigma_n \text{ (tidak terjadi heterogenitas spasial)}$$

$$H_1 : \text{Minimal ada satu } \sigma_i^2 = \sigma_j^2 \text{ (terjadi heterogenitas spasial)}$$

Statistik Uji :

$$BP = (1/2) f^T Z (Z^T Z)^{-1} Z^T I$$

Elemen vektor f adalah $f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right)$ dengan e_i merupakan residu untuk observasi ke- i

dan Z merupakan matriks berukuran $n \times (p+1)$ yang berisi vektor yang sudah distandarisasi untuk setiap observasi, H_0 ditolak jika $BP > \chi^2(\alpha, p)$.

2.4 Geographically Weighted Regression (GWR). Model GWR adalah pengembangan dari model regresi dimana setiap parameter dihitung pada setiap lokasi pengamatan, sehingga setiap lokasi pengamatan mempunyai nilai parameter regresi yang berbeda-beda. Model GWR merupakan pengembangan dari model regresi global dimana ide dasarnya diambil dari regresi non parametrik [10]. Variabel respon Y dalam model GWR diprediksi dengan variabel prediktor yang masing-masing koefisien regresinya bergantung pada lokasi dimana data tersebut diamati. Model GWR dapat ditulis sebagai berikut [6].

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

dengan

y_i : nilai observasi variabel respon untuk lokasi ke- i

(u_i, v_i) : koordinat letak geografis (*longitude, latitude*) dari lokasi pengamatan ke- i

$\beta_k(u_i, v_i)$: koefisien regresi variabel prediktor ke- k pada lokasi pengamatan ke- i

x_{ik} : nilai observasi variabel prediktor ke- k pada lokasi pengamatan ke- i

ε_i : eror pengamatan ke- i dengan $\varepsilon_i \sim IIDN(0, \sigma^2)$.

2.5 Estimasi Parameter Model GWR

Estimasi parameter model GWR dilakukan dengan metode *weighted least squares* (WLS) yaitu dengan memberikan pembobot yang berbeda untuk setiap lokasi dimana data diamati. Pemberian bobot ini sesuai dengan Hukum I Tobler yang menyatakan “segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang dekat lebih mempunyai pengaruh daripada sesuatu yang jauh”. Oleh karena itu pada model GWR diasumsikan bahwa daerah yang dekat dengan lokasi pengamatan ke- i mempunyai pengaruh yang besar terhadap estimasi parameternya daripada daerah yang lebih jauh. Misalkan pembobot untuk setiap lokasi adalah $j = 1, 2, \dots, n$ maka parameter pada lokasi

pengamatan diestimasi dengan menambahkan unsur pembobot pada persamaan (2.1) dan kemudian meminimumkan jumlah kuadrat residu

$$\sum_{j=1}^n w_j(u_i, v_i) \varepsilon_j^2 = \sum_{j=1}^n w_j(u_i, v_i) \left[y_j - \beta_0(u_i, v_i) - \sum_{k=1}^p \beta_k(u_i, v_i) x_{jk} \right]^2$$

atau dalam bentuk matriks jumlah kuadrat residu

$$\begin{aligned} \boldsymbol{\varepsilon}^T \mathbf{W}(u_i, v_i) \boldsymbol{\varepsilon} &= \mathbf{y}^T \mathbf{W}(u_i, v_i) \mathbf{y} - 2\boldsymbol{\beta}^T(u_i, v_i) \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \\ &+ \boldsymbol{\beta}^T(u_i, v_i) \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \boldsymbol{\beta}(u_i, v_i) \end{aligned} \quad (5)$$

dengan

$$\boldsymbol{\beta}(u_i, v_i) = \begin{pmatrix} \beta_0(u_i, v_i) \\ \beta_1(u_i, v_i) \\ \vdots \\ \beta_p(u_i, v_i) \end{pmatrix} \text{ dan } \mathbf{W}(u_i, v_i) = \text{diag}(w_1(u_i, v_i), w_2(u_i, v_i), \dots, w_n(u_i, v_i))$$

Jika persamaan (5) diturunkan terhadap $\boldsymbol{\beta}^T(u_i, v_i)$ dan hasilnya disamakan dengan nol maka diperoleh estimator parameter model GWR

$$\begin{aligned} \frac{\partial \boldsymbol{\varepsilon}^T \mathbf{W}(u_i, v_i) \boldsymbol{\varepsilon}}{\partial \boldsymbol{\beta}^T(u_i, v_i)} &= 0 - 2\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} + 2\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \boldsymbol{\beta}(u_i, v_i) = 0 \\ -2\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} + 2\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \boldsymbol{\beta}(u_i, v_i) &= 0 \\ \left[\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \right]^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \boldsymbol{\beta}(u_i, v_i) &= \left[\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \right]^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \\ \hat{\boldsymbol{\beta}}(u_i, v_i) &= \left[\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \right]^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \end{aligned} \quad (6)$$

Misalkan \mathbf{X}_i adalah elemen baris ke- i dari matriks \mathbf{X} . Nilai estimasi untuk y pada lokasi pengamatan adalah

$$\hat{y}_i = \mathbf{x}_i^T \hat{\boldsymbol{\beta}}(u_i, v_i) = \mathbf{x}_i^T \left(\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \right)^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y}$$

Untuk seluruh pengamatan dapat dituliskan dengan

$$\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)^T = \mathbf{L} \mathbf{y} \text{ dan}$$

$$\hat{\boldsymbol{\varepsilon}} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n)^T = (\mathbf{I} - \mathbf{L}) \mathbf{y}$$

dengan \mathbf{I} adalah matriks identitas berukuran $n \times n$ dan

$$\mathbf{L} = \begin{pmatrix} \mathbf{x}_1^T (\mathbf{X}^T \mathbf{W}(u_1, v_1) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_1, v_1) \\ \mathbf{x}_2^T (\mathbf{X}^T \mathbf{W}(u_2, v_2) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_n^T (\mathbf{X}^T \mathbf{W}(u_n, v_n) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_n, v_n) \end{pmatrix} \quad (7)$$

Estimator pada persamaan (6) merupakan estimator tak bias dan konsisten [7].

3. Metode Penelitian

3.1. Sumber Data

Penelitian ini menggunakan data Indeks Pembangunan Manusia (IPM) Metode Baru sebagai variabel dependen. Variabel independennya adalah Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS) dan Pengeluaran Perkapita yang disesuaikan (PPD) menurut provinsi di Indonesia tahun 2015. Data-data tersebut diambil dari Badan Pusat Statistik (BPS) Nasional [4].

3.2. Langkah Penelitian

Tahapan atau langkah-langkah dalam penelitian ini adalah

1. Melakukan analisis deskriptif data sebagai gambaran awal untuk mengetahui Indeks Pembangunan Manusia di Indonesia tahun 2015.
2. Menganalisis model regresi *Ordinary Least Square* (OLS) dengan langkah-langkah sebagai berikut:
 - a. Melakukan estimasi parameter model regresi OLS
 - b. Melakukan pengujian kesesuaian model regresi OLS
 - c. Melakukan pengujian parsial setiap parameter regresi OLS
 - d. Melakukan pengujian asumsi model regresi klasik, yaitu uji normalitas, non autokorelasi, non multikolinieritas, dan uji heterogenitas spasial
3. Menganalisis model GWR dengan langkah-langkah sebagai berikut [8]:
 - a. Menentukan bandwidth optimum dengan meminimumkan nilai CV dan AIC
 - b. Menghitung matriks pembobot dengan bandwidth optimum
 - c. Mendapatkan estimator parameter model GWR

- d. Melakukan pengujian kesesuaian model GWR
- e. Melakukan pengujian pengaruh lokasi secara parsial pada setiap variabel independen
- f. Melakukan pengujian secara parsial pada parameter GWR

4. Hasil dan Pembahasan

4.1. Deskripsi Data

Deskripsi data berdasarkan variabel dependen dan independen yang digunakan dalam analisis ini ditunjukkan oleh tabel berikut.

Tabel 4.1 Deskripsi Data Pembangunan Manusia di Indonesia
Menurut Provinsi Tahun 2015

Variabel	<i>N</i>	Rata-rata	Deviasi standar	Minimum	Maksimum
IPM	34	68,577	4,168	57,250	78,990
AHH	34	69,321	2,655	64,220	74,680
HLS	34	12,656	0,787	9,950	15,030
RLS	34	8,021	0,961	5,990	10,700
PPD	34	9778	2075	6469	17075

Tabel 4.1 menunjukkan bahwa rata-rata indeks pembangunan manusia di Indonesia adalah 68,577 dimana nilai IPM terendah berada pada provinsi Papua dengan nilai 57,25 sedangkan nilai IPM tertinggi adalah 78,99 berada pada provinsi DKI Jakarta.

4.2. Estimasi Model Regresi *Ordinary Least Square* (OLS)

Model regresi OLS digunakan untuk mengetahui variabel independen mana saja yang berpengaruh secara signifikan terhadap indeks pembangunan manusia dengan tanpa melibatkan faktor lokasi pengamatan. Berdasarkan hasil uji *F* diperoleh nilai $F_{hitung} = 1324,46$ dan nilai $p = 0,000$ dengan $\alpha = 0,05$ sehingga dapat disimpulkan bahwa secara simultan paling tidak ada satu koefisien regresi yang signifikan terhadap model regresi. Selain itu, nilai $R^2_{adj} = 99,4\%$, yang artinya sebesar 99,4% variabilitas IPM dijelaskan oleh adanya variabel AHH, HLS, RLS dan PPD, sedangkan sisanya dijelaskan oleh variabel lain di luar model.

Selain uji F , pada model OLS juga dilakukan uji parsial dengan menggunakan uji t . Berdasarkan hasil uji tersebut diperoleh bahwa semua koefisien regresi dari masing-masing variabel bebas dan *intercept* signifikan terhadap model karena semua nilai $p < 0,05$. Berdasarkan hal tersebut dapat disimpulkan bahwa variabel independen yang berpengaruh terhadap indeks pembangunan manusia adalah angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah dan pengeluaran perkapita yang disesuaikan. Persamaan regresi yang diperoleh adalah

$$\widehat{TPM} = - 0,21 + 0,516 \text{ AHH} + 1,19 \text{ HLS} + 1,09 \text{ RLS} + 0,000941 \text{ PPD} \quad (8)$$

Langkah selanjutnya adalah melakukan uji asumsi klasik pada model OLS yang sudah terbentuk. Uji normalitas dengan uji Shapiro-Wilk menunjukkan bahwa residu berdistribusi normal dengan nilai signifikansi 0,1 lebih dari $\alpha=0,05$. Selanjutnya dilakukan uji multikolinearitas dengan melihat *variance inflation factor* (VIF). Hasil uji VIF menunjukkan bahwa antar variabel independen tidak terdapat multikolinearitas, karena residu VIF kurang dari 5.

Asumsi independensi residu bertujuan untuk mengetahui apakah dalam model regresi linier terdapat korelasi antar residu. Cara yang dilakukan yaitu dengan melihat statistik uji Durbin-Watson. Hasil nilai statistik Durbin-Watson adalah 1,57876. Apabila dibandingkan dengan tabel statistik uji Durbin-Watson untuk $\alpha=0,05$, $n=34$ dan dengan variabel independen sebanyak 4 diperoleh batas bawah (dL)= 1,2078 dan batas atas (dU)= 1,7277 maka nilai statistik Durbin-Watson berada diantara nilai dL dan dU sehingga tidak dapat disimpulkan apakah terjadi autokorelasi atau tidak.

Model regresi yang sah mengasumsikan variansi dari erornya tetap (homoskedastisitas). Statistik uji yang digunakan adalah uji Glejser [8]. Statistik uji ini diperoleh dengan meregresikan nilai absolut error dari model (8) dengan semua variabel independennya. Hasil pengolahan menunjukkan bahwa variabel independennya ternyata signifikan mempengaruhi nilai absolut error, ini berarti bahwa dalam data terdapat kasus heteroskedastisitas.

Untuk mengidentifikasi terjadinya kasus heterogenitas spasial pada regresi digunakan uji Breusch-Pagan. Berdasarkan pengolahan data diperoleh nilai statistik BP adalah 10,748 dengan nilai p sebesar 0,02954. Nilai tabel khi-kuadrat dengan derajat bebas 4 dan $\alpha=0,05$ adalah 9,487. Dengan demikian dapat disimpulkan bahwa terjadi kasus heterogenitas spasial pada data penelitian, hasil ini memperkuat hasil pengujian asumsi heteroskedastisitas pada OLS yang menggunakan uji Glejser. Oleh karena itu dapat

disimpulkan bahwa metode regresi global kurang tepat untuk menggambarkan pemodelan indeks pembangunan manusia di Indonesia, akan lebih baik jika digunakan model yang mengakomodasi faktor lokasi pengamatan.

4.3. Estimasi Model *Geographically Weighted Regression* (GWR)

Terjadinya kasus heterogenitas spasial pada data indeks pembangunan manusia di Indonesia mengindikasikan bahwa parameter model regresi dipengaruhi oleh faktor lokasi pengamatan, dalam hal ini adalah letak geografis provinsi. Oleh karena itu dilakukan pemodelan dengan mengakomodasi faktor lokasi yaitu dengan model GWR. Langkah pertama adalah menentukan letak geografis tiap provinsi di Indonesia. Langkah kedua adalah menentukan matriks pembobot GWR dengan menggunakan kernel *fixed* Gaussian dan *fixed bisquare*. Untuk menentukan fungsi pembobot yang menghasilkan *bandwidth* minimum dan menentukan model terbaik digunakan indikator AIC, CV, MSE, dan R^2 . Tabel 4.2 menunjukkan hasil perbandingannya.

Tabel 4.2 Pemilihan *bandwidth* optimum dan model terbaik

Statistik	<i>Fixed</i> Gaussian	<i>Fixed bisquare</i>
Bandwidth	522,199	1558,369
AIC	-8,4698	-3,6344
CV	0,1020	0,0958
MSE	0,0393	0,0425
R^2	0,999	0,998

Pada Tabel 4.2 menunjukkan bahwa fungsi kernel *fixed* Gaussian memiliki AIC dan MSE yang lebih kecil, serta memiliki nilai R^2 yang lebih besar dibandingkan fungsi kernel *fixed bisquare*. Dengan melihat indikator-indikator tersebut, maka matriks pembobot untuk GWR dalam penelitian ini ditentukan dengan fungsi kernel *fixed* Gaussian dengan nilai *bandwidth* optimum sebesar 522,199. Tabel 4.3 merupakan rangkuman hasil estimasi parameter model GWR.

Tabel 4.3 Nilai Estimasi Parameter GWR

Variabel	Minimum	Maksimum
Konstanta	-1,088	11,616
AHH	0,395	0,495
HLS	0,874	1,249
RLS	0,969	1,212
PPD	0,00088	0,0013

Pengujian kesesuaian model GWR dilakukan dengan menggunakan selisih jumlah kuadrat residu model GWR dan model regresi OLS. Model GWR akan berbeda signifikan dengan model regresi OLS jika dapat menurunkan jumlah kuadrat residu secara signifikan. Uji kesesuaian model menunjukkan bahwa nilai statistik uji F sebesar 4,449 dengan nilai p sebesar 0,0037. Dengan menggunakan tingkat signifikansi sebesar 5% maka dapat disimpulkan bahwa model GWR berbeda signifikan dengan model regresi OLS sehingga model GWR lebih layak untuk menggambarkan indeks pembangunan manusia di Indonesia tahun 2015.

Suatu variabel memiliki heterogenitas spasial jika nilai *difference of criterion* (*DIFF of Criterion*) yang bernilai negatif. Apabila nilai *difference of criterion* bernilai positif artinya variabel bebas tersebut bersifat global. Tabel 4.4 menunjukkan bahwa variabel angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah dan pengeluaran perkapita yang disesuaikan secara signifikan memiliki heterogenitas spasial atau bersifat lokal pada masing-masing provinsi di Indonesia.

Tabel 4.4 Pengujian Variasi Koefisien Lokal

Variabel	F hitung	Derajat bebas	F tabel	<i>DIFF of Criterion</i>
Intercept	229338,225100	0,311	17,868	-230050,783933
AHH	424,421386	1,726	17,868	-795,461721
HLS	337,172795	1,760	17,868	-18,100497
RLS	816,831568	1,438	17,868	-27,767021
PPD	36,025275	1,800	17,868	-26,868022

Metode GWR menghasilkan persamaan regresi yang bersifat lokal dan bervariasi antar wilayah. Variasi dari koefisien regresi tersebut menunjukkan adanya variasi spasial sehingga untuk mengatasi masalah pembangunan manusia diperlukan kebijakan yang berbeda antar wilayah, sesuai dengan karakteristik dari masing-masing provinsi. Dengan demikian penelitian ini akan menghasilkan 34 persamaan regresi sesuai dengan jumlah Provinsi yang ada di Indonesia. Gambaran sebaran pengaruh masing-masing variabel untuk setiap Provinsi di Indonesia dapat dilihat di Tabel 4.5.

Tabel 4.5 Variabel Independen yang Signifikan di Setiap Provinsi

Provinsi	Variabel yang Berpengaruh
Aceh, Banten, Jambi, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, Kep. Bangka Belitung, Kep. Riau, Lampung, Maluku, Maluku Utara, Nusa Tenggara Barat, Nusa Tenggara Timur, Papua, Papua Barat, Sulawesi Barat, Sulawesi Selatan, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara, Sumatera Barat, Sumatera Selatan, Sumatera Utara, Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Timur.	angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah, pengeluaran perkapita yang disesuaikan
Bengkulu, Gorontalo.	angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah

Dalam penelitian ini, terdapat 29 provinsi yang dipengaruhi oleh variabel angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah dan pengeluaran perkapita yang disesuaikan dan terdapat 5 provinsi yang hanya dipengaruhi oleh variabel angka harapan hidup, harapan lama sekolah dan rata-rata lama sekolah, perbedaan lain terdapat pada koefisien regresi masing-masing provinsi. Persamaan GWR yang diperoleh di Provinsi Papua adalah

$$\widehat{IPM} = -0,94608 + 0,482675AHH + 1,245689HLS + 1,004157RLS + 0,001294PPD$$

sedangkan persamaan GWR untuk Provinsi DKI Jakarta adalah

$$\widehat{IPM} = 6,288433 + 0,436794AHH + 1,195353HLS + 1,04087RLS + 0,00088PPD$$

5. Kesimpulan

Rata-rata indeks pembangunan manusia di Indonesia adalah 68,577 dimana nilai IPM terendah berada pada provinsi Papua dengan nilai 57,25 sedangkan nilai IPM tertinggi adalah 78,99 berada pada provinsi DKI Jakarta. Model GWR dengan fungsi pembobot *fixed* Gaussian memberikan hasil yang lebih baik dibandingkan dengan model GWR dengan fungsi pembobot *fixed bisquare*. Angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah, dan pengeluaran perkapita yang disesuaikan memiliki pengaruh yang berbeda pada masing-masing provinsi di Indonesia. Provinsi yang dipengaruhi oleh variabel angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah, dan pengeluaran perkapita yang disesuaikan adalah Provinsi Aceh, Banten, Jambi, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, Kep. Bangka Belitung, Kep. Riau, Lampung, Maluku, Maluku Utara, Nusa Tenggara Barat, Nusa Tenggara Timur, Papua, Papua Barat, Sulawesi Barat, Sulawesi Selatan, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara, Sumatera Barat, Sumatera Selatan, Sumatera Utara, Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Timur, sedangkan provinsi yang dipengaruhi oleh variabel angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah adalah Provinsi Bengkulu dan Gorontalo.

Daftar Pustaka

- [1] United Nations Development Programme. *Human Development Report 2000*. Oxford Oxford University Press. New York. 2000.
- [2] Rencher, A.C. and Schaalje, G.B. *Linier Models in Statistics*. John Wiley & Sons Inc. Singapore. 2000.
- [3] Anselin, L. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers. Dordrecht. 1998.
- [4] Badan Pusat Statistik Nasional. *Indeks Pembangunan Manusia Metode Baru Tahun 2015*. BPS. Jakarta. 2015.
- [5] Charlton, M. and Fotheringham, A.S. *Geographically Weighted Regression White Paper*. National Centre for Geocomputation. Maynooth University. 2009.
- [6] Chasco, C., Garcia, I., & Vicens, J. Modeling Spastial Variations in Household Disposable Income with Geographically Weighted Regression. *Munich Personal RePEc Arkhiv (MPRA)*. Working Papper No. 1682. 2007.
- [7] Fotheringham, A.S., Brunson, C., dan Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Jhon Wiley & Sons, LTD. 2002.
- [8] Nurdin, F.E. *Estimasi dan Pengujian Hipotesis Geographically Weighted Regression Studi Kasus Produktivitas Padi Sawah di Jawa Timur*. Surabaya: Jurusan Statistika FMIPA ITS. 2008.