

PEMODELAN PERSENTASE BALITA GIZI BURUK DI JAWA TENGAH DENGAN PENDEKATAN *GEOGRAPHICALLY WEIGHTED REGRESSION PRINCIPAL COMPONENTS ANALYSIS (GWRPCA)*

Novika Pratnyaningrum¹, Hasbi Yasin², Abdul Hoyyi³

¹Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

^{2,3}Staf Pengajar Jurusan Statistika FSM Universitas Diponegoro

ABSTRACT

Geographically Weighted Regression Principal Components Analysis (GWRPCA) is a combination of method of Principal Components Analysis (PCA) and Geographically Weighted Regression (GWR). PCA is used to eliminate the multicollinearity effect in regression analysis. GWR is a local form of regression and a statistical method used to analyze the spatial data. In GWRPCA predictor variables is a principal components of the PCA result. Estimates of parameters of the GWRPCA model can use Weighted Least Square (WLS). Selection of the optimum bandwidth use Cross Validation (CV) method. Conformance testing PCA regression and GWRPCA models approximated by the F distribution, while the partial identification of the model parameters using the t distribution. In PCA obtained variables that affect the percentage of severe children malnutrition in Central Java in 2012 can be represented or replaced with PC₁ and PC₂ which can explain the total variance of data is 78.43%. Application GWRPCA models at the percentage of severe children malnutrition in Central Java in 2012 showed every regency locations have different model with global coefficient of determination is 0.6313309 and the largest local coefficient of determination is 0.72793026 present in Batang regency, while the smallest local coefficient of determination is 0.03519539 present in Sukoharjo regency.

Keywords : Severe Malnutrition, Multicollinearity, Geographically Weighted Regression Principal Components Analysis, Weighted Least Square, Coefficient of Determination.

1. PENDAHULUAN

Dewasa ini, kesehatan masyarakat masih menjadi masalah bagi negara-negara di dunia termasuk Indonesia. Dalam menilai derajat kesehatan masyarakat, terdapat beberapa indikator yang dapat digunakan. Indikator-indikator tersebut pada umumnya tercermin dalam kondisi angka kematian, angka kesakitan dan status gizi (Dinkes, 2013). Sebagai negara dengan keberagaman penduduk yang tinggi, Indonesia dihadapkan oleh persoalan gizi buruk. Persoalan gizi buruk di masyarakat sering digambarkan dengan masalah gizi pada kelompok balita (bawah lima tahun). Hal tersebut dikarenakan kelompok balita merupakan bagian yang sangat sensitif terhadap perubahan kondisi gizi di dalam masyarakat (Sediaoetama, 2006).

Persoalan gizi pada kelompok balita masih menjadi masalah serius bagi sebagian kabupaten/kota di Jawa Tengah. Provinsi Jawa Tengah selama 6 tahun berturut-turut (2005-2010) masuk ke dalam kategori 10 provinsi dengan kasus tertinggi. Bahkan pada tahun 2006, Jawa Tengah menyumbang angka gizi buruk tertinggi dalam skala nasional, yaitu 10.376 kasus. Pada tahun 2011 jumlah penderita gizi buruk dapat ditekan menjadi 3.178 kasus dan menurun kembali pada tahun 2012 yaitu berjumlah 1.131 (Dinkes, 2013). Meskipun angka penderita gizi buruk di Jawa Tengah mengalami penurunan akan tetapi permasalahan ini harus segera diselesaikan mengingat dampak jangka

panjang dari gizi buruk. Balita yang digolongkan gizi buruk berisiko memiliki kecerdasan yang kurang dibandingkan dengan balita yang lebih sehat. Ini semua disebabkan oleh kenyataan bahwa masalah gizi merupakan faktor dasar (*underlying factor*) dari berbagai masalah kesehatan, terutama pada bayi dan anak-anak (FKM UI, 2008). Dengan demikian, gizi harus mendapatkan perhatian serius dari semua pihak, dan gizi harus diposisikan sebagai pusat dari pembangunan suatu bangsa.

Penyebab mendasar dari gizi buruk adalah ketidakcukupan pasokan zat gizi ke dalam sel. Meskipun banyak disebabkan oleh kekurangan zat gizi yang esensial, tetapi faktor penyebab gizi buruk sebenarnya sangat kompleks. Beberapa penelitian telah dilakukan untuk mengetahui faktor-faktor yang mempengaruhi gizi buruk, salah satunya adalah Ramadani (2013) yang menggunakan metode *Spatial Durbin Model*. Dalam penelitian tersebut didapatkan hasil bahwa bayi yang lahir dengan berat badan rendah (BBLR), tempat tinggal dengan kategori rumah sehat, dan akses terhadap air bersih mempengaruhi gizi buruk balita di Jawa Tengah. Selain itu, faktor-faktor yang secara global mempengaruhi gizi buruk balita di Jawa Tengah adalah pemberian ASI eksklusif, bayi yang lahir dengan berat badan rendah (BBLR), tempat tinggal dengan kategori rumah sehat, akses terhadap air bersih, kepemilikan fasilitas Buang Air Besar (BAB), dan pemberian imunisasi secara lengkap.

Berdasarkan penyebab-penyebab balita mengalami gizi buruk tersebut ingin dilakukan penyederhanaan (reduksi) variabel-variabel penyebab balita gizi buruk dengan mempertahankan sebanyak mungkin informasinya dan menghilangkan korelasi antar variabel-variabel tersebut. Metode yang sesuai dengan tujuan tersebut adalah Analisis Komponen Utama atau *Principal Components Analysis* (PCA). Menurut Johnson dan Wichern (2007) PCA merupakan suatu teknik statistik yang menerangkan struktur varian kovarian dari sekumpulan variabel melalui kombinasi linier yang lebih sedikit dari variabel tersebut, tujuannya adalah reduksi data dan interpretasi.

Pemodelan persentase balita gizi buruk berdasarkan karakteristik daerah akan dipengaruhi oleh letak geografis antar daerah. Hal ini dikarenakan perbedaan letak geografis akan mempengaruhi taraf kesehatan, sehingga diperlukan suatu metode pemodelan statistik yang memperhatikan letak geografis atau faktor lokasi pengamatan. Salah satu metode untuk menganalisisnya adalah dengan menggunakan model *Geographically Weighted Regression* (GWR) (Fotheringham *et al.*, 2002). Model GWR merupakan pengembangan dari model regresi linier. Pada model regresi linier hanya dihasilkan estimator parameter yang berlaku secara global, sedangkan dalam model GWR dihasilkan estimator parameter model yang bersifat lokal untuk setiap lokasi pengamatan (Purhadi dan Yasin, 2012). Oleh karena itu, untuk melakukan pemodelan balita gizi buruk di Jawa Tengah dengan terlebih dahulu melakukan analisis komponen utama pada variabel-variabel yang mempengaruhi persentase balita gizi buruk digunakan metode *Geographically Weighted Regression Principal Components Analysis* (GWRPCA).

2. TINJAUAN PUSTAKA

2.1 Gizi Buruk

Gizi merupakan salah satu penentu kualitas sumber daya manusia. Status gizi adalah keadaan tubuh sebagai akibat konsumsi makanan dan penggunaan zat-zat gizi. Salah satu klasifikasi status gizi adalah gizi buruk. Gizi buruk merupakan kondisi terparah dari PCM (*Protein Calori Malnutrition*), termasuk marasmus (kekurangan

karbohidrat), kwasiorkor (kekurangan protein) dan marasmik-kwasiorkor (kekurangan karbohidrat dan protein).

2.2 Uji Asumsi *Principal Components Analysis*

Dalam penggunaan *Principal Components Analysis* (PCA), terdapat beberapa asumsi yang harus dipenuhi. Asumsi-asumsi tersebut antara lain pengujian kecukupan data, multikolinieritas dan normal multivariat. Pengujian kecukupan data menggunakan *Kaiser-Meyer-Olkin* (KMO) dan *Measure of Sampling Adequacy* (MSA). Dalam pengujian normal multivariat menggunakan uji *Kolmogorov Smirnov*. Sedangkan untuk mendeteksi ada tidaknya multikolinieritas menggunakan *Bartlett Test of sphericity*.

2.3 *Principal Components Analysis*

Principal Components Analysis (PCA) merupakan teknik statistika yang menerangkan struktur varian kovarian dari sekumpulan variabel melalui kombinasi linier yang lebih sedikit dari variabel tersebut, tujuannya adalah reduksi data dan interpretasi (Johnson dan Wichern, 2007). Komponen utama ke- i dari variabel yang distandarkan $\mathbf{Z}^T = [Z_1, Z_2, \dots, Z_p]$ dengan $\text{Cov}(\mathbf{Z}) = \boldsymbol{\rho}$, pasangan nilai eigen dan vektor eigen $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$ dan $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ yaitu:

$$\text{PC}_1 = \mathbf{e}_1^T \mathbf{Z} = e_{11}Z_1 + e_{12}Z_2 + \dots + e_{1p}Z_p$$

$$\text{PC}_2 = \mathbf{e}_2^T \mathbf{Z} = e_{21}Z_1 + e_{22}Z_2 + \dots + e_{2p}Z_p$$

⋮

$$\text{PC}_p = \mathbf{e}_p^T \mathbf{Z} = e_{p1}Z_1 + e_{p2}Z_2 + \dots + e_{pp}Z_p$$

Proporsi dari varian populasi yang distandarkan yang dijelaskan oleh komponen utama

$$k\text{-k} = \frac{\lambda_k}{p}; \quad k = 1, 2, \dots, p$$

dengan λ_k adalah nilai eigen dari $\boldsymbol{\rho}$.

Pemilihan komponen utama pada dasarnya sesuai kebutuhan peneliti karena tidak ada aturan khusus dalam pemilihan jumlah komponen utama. Menurut Simamora (2005) jumlah komponen utama dapat ditentukan dengan melihat persentase varian kumulatif yang dipilih mampu menerangkan total varian data sekitar 70% sampai 80%. Jumlah komponen utama juga dapat diketahui dengan melihat pola penurunan nilai eigen pada *scree plot*. Penurunan yang tajam pada *scree plot* menunjukkan perubahan nilai eigen yang besar. *Scree plot* adalah plot antara nilai eigen λ_k dengan k . Selain itu, pemilihan komponen utama juga dapat ditentukan berdasarkan nilai eigen yang lebih dari satu. Hal tersebut dikarenakan nilai eigen yang mendekati nol dianggap tidak memberikan pengaruh yang penting.

2.4 *Geographically Weighted Regression Principal Components Analysis* (GWRPCA)

Geographically Weighted Regression Principal Components Analysis (GWRPCA) merupakan gabungan dari metode Analisis Komponen Utama atau *Principal Components Analysis* (PCA) dan *Geographically Weighted Regression* (GWR). Setelah didapatkan komponen utama, maka dilakukan uji heterogenitas spasial untuk mengetahui adanya perbedaan karakteristik satu wilayah dengan wilayah lainnya (efek wilayah yang random). Menurut Purhadi dan Yasin (2012) pengujian heterogenitas spasial dilakukan dengan menggunakan statistik uji *Breusch-Pagan Test*.

Model *Geographically Weighted Regression* (GWR) merupakan pengembangan dari model regresi dimana setiap parameter dihitung pada setiap lokasi pengamatan, sehingga setiap lokasi pengamatan mempunyai nilai parameter regresi yang berbeda-

beda atau bersifat lokal (Fotheringham *et al.*, 2002). Variabel respon y dalam model GWR ditaksir dengan variabel prediktor yang masing-masing koefisien regresinya bergantung pada lokasi dimana data tersebut diamati. Model GWR dapat ditulis sebagai berikut:

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

Estimasi parameter pada model GWR adalah dengan metode *Weighted Least Square* (WLS) yaitu dengan memberikan pembobot yang berbeda untuk setiap lokasi dimana data tersebut diambil (Fotheringham *et al.*, 2002). Pada model GWR diasumsikan bahwa daerah yang dekat dengan lokasi pengamatan ke- i mempunyai pengaruh yang besar terhadap estimasi parameternya dari pada daerah yang lebih jauh. Estimator parameter dari model GWR untuk setiap lokasi (Fotheringham *et al.*, 2002) adalah:

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

Peran pembobot dalam GWR sangat penting karena nilai pembobot mewakili letak data observasi antara satu dengan lainnya. Matriks pembobot berdimensi $n \times n$ adalah sebagai berikut:

$$W(u_i, v_i) = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{in} \end{bmatrix}$$

dengan w_{in} adalah bobot untuk data pada titik ke- n dalam pengujian model di sekitar titik i .

Matriks pembobot lokasi ke- i diperoleh dari pembobot fungsi kernel *adaptive bisquare* dengan formula sebagai berikut:

$$w_j(u_i, v_i) = \left(1 - \left(\frac{d_{ij}}{h_i}\right)^2\right)^2, \text{ jika } d_{ij} \leq h_i \text{ dan } w_j(u_i, v_i) = 0, \text{ untuk yang lain}$$

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

u_i = koordinat latitude pada lokasi ke- i

v_i = koordinat longitude pada lokasi ke- i

h_i = *bandwidth* pada lokasi ke- i

Fungsi kernel *adaptive* menyesuaikan ukuran variasi pada densitas data. Fungsi tersebut mempunyai *bandwidth* yang lebih besar ketika data menyebar dan lebih kecil ketika data memusat (Fotheringham *et al.*, 2002). *Bandwidth* adalah ukuran jarak fungsi pembobot dan sejauh mana pengaruh suatu lokasi terhadap lokasi lain. Terdapat beberapa metode yang digunakan untuk memilih *bandwidth* optimum, salah satu diantaranya adalah metode *Cross Validation* (CV) yang secara matematis didefinisikan sebagai berikut (Fotheringham *et al.*, 2002):

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

Dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir y_i dimana pengamatan di lokasi (u_i, v_i) dihilangkan dari proses penaksiran. Untuk mendapatkan nilai *bandwidth* (h) yang optimal maka diperoleh dari h yang menghasilkan nilai CV yang minimum.

2.4.1 Pengujian Hipotesis Model GWRPCA

Pengujian hipotesis pada model GWRPCA terdiri dari pengujian kesesuaian model GWRPCA dan pengujian parameter model. Pengujian kesesuaian model dapat dilakukan dengan hipotesis sebagai berikut:

$H_0 : \beta_k(u_i, v_i) = \beta_k$ untuk setiap $k = 0, 1, 2, \dots, p$ dan $i = 1, 2, \dots, n$
 (tidak ada perbedaan yang signifikan antara model regresi PCA dengan GWRPCA)

H_1 : Minimal ada satu $\beta_k(u_i, v_i) \neq \beta_k, k = 0, 1, 2, \dots, p$
 (ada perbedaan yang signifikan antara model regresi PCA dan GWRPCA).

Penentuan statistik uji berdasarkan *Residual Sum of Square* (RSS). Statistik uji:

$$F_1 = \frac{RSS(H_1)/df_1}{RSS(H_0)/df_2}$$

Dibawah H_0 , statistik uji F_1 akan mengikuti distribusi F dengan derajat bebas $df_1 = \frac{\delta_1^2}{\delta_2}$ dan $df_2 = (n - p - 1)$. Jika diambil taraf signifikansi α maka tolak H_0 jika $F_1 < F_{1-\alpha, df_1, df_2}$.

Jika disimpulkan bahwa model GWRPCA berbeda nyata dengan model regresi PCA, maka langkah selanjutnya adalah melakukan uji parsial untuk mengetahui apakah ada perbedaan pengaruh yang signifikan dari variabel prediktor x_k antara satu lokasi dengan lokasi lainnya (Mei *et al.*, 2004). Pengujian ini dapat dilakukan dengan hipotesis sebagai berikut:

$H_0: \beta_k(u_1, v_1) = \beta_k(u_2, v_2) = \dots = \beta_k(u_n, v_n)$ untuk suatu $k (k = 0, 1, 2, \dots, p)$
 (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 $i = 1, 2, \dots, n$ yang berbeda.
 (ada perbedaan pengaruh yang signifikan dari variabel prediktor x_k antara satu lokasi dengan lokasi lainnya)

Statistik uji:

$$F_2 = \frac{v_k^2 / \text{tr}(\frac{1}{n} \mathbf{B}_k^T [\mathbf{I} - \frac{1}{n} \mathbf{J}] \mathbf{B}_k)}{RSS(H_1)/\delta_1}$$

Di bawah H_0 , statistik uji F_2 akan mengikuti distribusi F dengan derajat bebas $df_1 = \left(\frac{\gamma_1^2}{\gamma_2} \right)$ dan $df_2 = \left(\frac{\delta_1^2}{\delta_2} \right)$. Jika diberikan taraf signifikansi sebesar α maka akan menolak H_0 jika $F_2 \geq F_{\alpha, df_1, df_2}$.

Pengujian signifikansi parameter model pada setiap lokasi dilakukan dengan menguji parameter secara parsial (Purhadi dan Yasin, 2012). Tujuan dari pengujian ini untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Dengan hipotesis sebagai berikut:

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

H_1 : Minimal ada satu $\beta_k(u_i, v_i) \neq 0$ dengan $k = 1, 2, \dots, p, i = 1, 2, \dots, n$

Statistik uji yang digunakan adalah:

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

T akan mengikuti distribusi t dengan derajat bebas $df = \frac{\delta_1^2}{\delta_2}$ dan $\hat{\sigma}^2 = \frac{RSS(H_1)}{\delta_1}$. Jika diberikan taraf signifikansi sebesar α , maka diambil keputusan menolak

H_0 atau dengan kata lain parameter $\beta_k(u_i, v_i)$ signifikan terhadap model jika $|T_{hit}| > t_{\alpha/2; df}$.

2.4.2 Koefisien Determinasi

Koefisien determinasi merupakan suatu nilai atau ukuran yang dapat digunakan untuk mengetahui seberapa jauh kecocokan dari suatu model regresi. Nilai koefisien determinasi menunjukkan proporsi atau persentase variasi total dalam variabel respon Y yang dijelaskan oleh variabel prediktor x_k ($k = 1, 2, \dots, p$). Dalam GWRPCA terdapat dua koefisien determinasi yaitu koefisien determinasi global dan koefisien determinasi lokal. Menurut Fotheringham *et al.* (2002) koefisien determinasi global mengukur proporsi varian data yang dijelaskan model, sedangkan koefisien determinasi lokal menunjukkan bagaimana model lokal dapat meniru data yang tercatat dalam daerah sekitar titik regresi dengan baik.

3. METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder tentang faktor-faktor yang diduga mempengaruhi persentase balita gizi buruk di Provinsi Jawa Tengah tahun 2012. Data ini bersumber dari Profil Kesehatan Provinsi Jawa Tengah Tahun 2012 yang dipublikasikan oleh Dinas Kesehatan Provinsi Jawa Tengah.

3.2 Variabel Penelitian

Penelitian ini menggunakan persentase balita gizi buruk sebagai variabel terikat dan lima variabel bebas yang diduga mempengaruhi balita gizi buruk. Variabel penelitian akan disajikan pada Tabel 1.

Tabel 1 Variabel Penelitian

Kode	Variabel
Y	Persentase Balita Gizi Buruk
u	Garis Lintang Selatan
v	Garis Bujur Timur
X_1	Banyaknya Bayi (0-6 Bulan) yang Diberi ASI Eksklusif
X_2	Banyaknya Bayi yang Mendapat Imunisasi BCG
X_3	Banyaknya Balita yang Mendapat Vitamin A
X_4	Banyaknya Balita yang Menderita Pneumonia
X_5	Banyaknya Rumah Tangga yang Memiliki Akses Air Bersih

3.3 Langkah Analisis Data

Langkah-langkah yang dilakukan untuk menganalisis data dalam penelitian ini adalah sebagai berikut :

1. Menentukan variabel bebas yang diperkirakan mempengaruhi persentase balita gizi buruk di Jawa Tengah.
2. Melakukan uji asumsi *Principal Components Analysis* (PCA) yang terdiri dari:
 - a. Uji kecukupan data dengan *Kaiser-Meyer-Olkin* (KMO) dan *Measure of Sampling Adequacy* (MSA).
 - b. Uji normal multivariat.
 - c. Uji multikolinieritas dengan *Bartlett Test of Sphericity*.

3. Melakukan *Principal Components Analysis* (PCA) yaitu menentukan komponen utama yang digunakan untuk mewakili (menggantikan) data asli dengan kriteria variansi yang dijelaskan adalah 70%-80%.
4. Melakukan uji Heterogenitas Spasial dengan uji *Breusch-Pagan*.
5. Melakukan analisis *Geographically Weighted Regression Principal Components Analysis* (GWRPCA) yang terdiri dari:
 - a. Menghitung jarak *Euclidean*
 - b. Menentukan bandwidth optimum dengan menggunakan metode *Cross Validation* (CV)
 - c. Menghitung matriks pembobot dengan *bandwidth* optimum
 - d. Mendapatkan estimator parameter model GWRPCA
 - e. Melakukan pengujian kesesuaian model GWRPCA
 - f. Melakukan identifikasi pengaruh faktor geografis secara parsial pada model GWRPCA
 - g. Melakukan identifikasi variabel yang signifikan mempengaruhi model GWRPCA.

4. HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

Tabel 2 Deskripsi Data Penelitian

Variabel	Minimal	Maksimal	Rataan	Standar Deviasi
Y	0,00	2,51	0,55057	0,65475
X ₁	455	9547	2973,97143	2432,99702
X ₂	2067	32270	11672,54286	7784,30638
X ₃	6581	115724	39833,08571	26827,00779
X ₄	331	6141	1278,31429	1789,62526
X ₅	14814	297766	78688,91429	79348,90592

Pada Tabel 2 variabel respon (Y) yaitu persentase balita gizi buruk, mempunyai nilai minimal 0,00% yang terdapat pada Kota Surakarta. Sedangkan nilai maksimal persentase balita gizi buruk adalah 2,51% yang terdapat pada Kabupaten Jepara.

4.2 Uji Asumsi *Principal Components Analysis*

4.2.1 Kecukupan Data

Pada uji KMO didapatkan nilai KMO (0,589) lebih besar dari 0,5 sehingga ukuran data cukup untuk dianalisis komponen utama. Sedangkan pada uji MSA didapatkan nilai MSA setiap variabel lebih besar dari 0,5 maka setiap variabel layak untuk dianalisis lebih lanjut menggunakan analisis komponen utama.

4.2.2 Uji Normal Multivariat

Dengan taraf signifikansi (α) sebesar 5% dan menolak H_0 jika nilai $D > W_{(0,95;35)}(0,224)$ atau $p\text{-value} < 0,05$ maka H_0 diterima karena nilai D (0,155) lebih kecil dari $W_{(0,95;35)}(0,224)$ dan $p\text{-value}$ (0,3347) lebih besar dari 0,05 sehingga data pengamatan berdistribusi normal multivariat.

4.2.3 Uji Multikolinieritas

Dengan taraf signifikansi (α) sebesar 5% dan menolak H_0 jika nilai $\chi^2_{hitung} \geq \chi^2_{0,05;10}(18,307)$ atau $p\text{-value} < \alpha$ maka H_0 ditolak karena χ^2_{hitung} (151,738) lebih besar

dari $\chi^2_{0,05;10}$ (18,307) dan p -value (0,000) lebih kecil dari 0,05 sehingga terdapat korelasi antar variabel.

4.3 Pembentukan Komponen Utama

Variabel $X_1, X_2, X_3, X_4,$ dan X_5 distandarkan sehingga terbentuk variabel $Z_1, Z_2, Z_3, Z_4,$ dan Z_5 . Kemudian variabel $Z_1, Z_2, Z_3, Z_4,$ dan Z_5 dikombinasi linier sehingga menghasilkan variabel-variabel baru (komponen utama) yang tidak berkorelasi. Pembentukan komponen utama adalah sebagai berikut:

$$PC_1 = -0,4065526Z_1 - 0,5333176Z_2 - 0,5255215Z_3 - 0,3515096Z_4 - 0,3880147Z_5$$

$$PC_2 = -0,5482232Z_1 + 0,0319204Z_2 + 0,01112886Z_3 + 0,80728577Z_4 - 0,21586628Z_5$$

$$PC_3 = 0,214Z_1 + 0,3066489Z_2 + 0,2497527Z_3 - 0,1073483Z_4 - 0,8867189Z_5$$

$$PC_4 = -0,6951926Z_1 + 0,3067409Z_2 + 0,4451881Z_3 - 0,4584936Z_4 + 0,1191990Z_5$$

$$PC_5 = 0,07125716Z_1 - 0,72555487Z_2 + 0,68053977Z_3 + 0,05468661Z_4 - 0,04865737Z_5$$

Komponen utama pertama (PC_1) dan komponen utama kedua (PC_2) mampu menerangkan total varian data 78,43%. Selain itu pada *scree plot* terjadi penurunan yang tajam pada komponen utama kedua (PC_2) yang menunjukkan perubahan nilai eigen yang besar. Berdasarkan dua alasan tersebut maka komponen utama pertama (PC_1) dan komponen utama kedua (PC_2) dipilih untuk mewakili atau menggantikan variabel-variabel yang berpengaruh terhadap persentase balita gizi buruk.

4.4 Pemodelan Menggunakan GWRPCA

Setelah didapatkan variabel PC_1 dan PC_2 , maka dilakukan uji heterogenitas spasial. Dengan taraf signifikansi (α) sebesar 10% dan menolak H_0 jika nilai $BP > \chi^2_{(2)}$ (4,605) atau p -value $< 0,1$ maka H_0 ditolak karena nilai BP (5,120) lebih besar dari nilai $\chi^2_{(2)}$ (4,605) dan p -value (0,077) lebih kecil dari 0,1 sehingga menunjukkan terdapat heterogenitas spasial.

Langkah selanjutnya setelah asumsi heterogenitas spasial terpenuhi adalah menentukan lokasi setiap sampel yaitu letak geografis kabupaten dan kota di Jawa Tengah. Kemudian menghitung jarak *Euclidean* dengan menggunakan letak geografis tersebut. Langkah selanjutnya adalah menentukan *bandwidth* optimum dengan menggunakan metode *Cross Validation* (CV). *Bandwidth* optimum ini digunakan untuk memperoleh matriks pembobot dari pembobot fungsi kernel adaptive *bisquare*.

Tahapan berikutnya setelah mendapatkan matriks pembobot untuk setiap lokasi pengamatan adalah mengestimasi parameter model GWRPCA. Contoh model GWRPCA pada Kabupaten Purbalingga adalah sebagai berikut:

$$\hat{y} = 0,8455201 + 0,2261713PC_1 + 0,1087978PC_2$$

Jika model tersebut dibawa ke dalam model variabel yang distandarkan, maka akan menjadi model GWRPCA sebagai berikut:

$$\hat{y} = 0,8455201 - 0,1515961Z_1 - 0,1171483Z_2 - 0,1176471Z_3 + 0,0083296Z_4 - 0,1112436Z_5$$

Kemudian jika model variabel yang distandarkan dibawa ke dalam model variabel yang belum distandarkan didapatkan model GWRPCA sebagai berikut:

$$\hat{y} = 0,2055008 - 0,0000623X_1 - 0,0000150X_2 - 0,0000044X_3 + 0,0000047X_4 - 0,0000014X_5$$

Persamaan pada model GWRPCA dengan variabel yang distandarkan menjelaskan bahwa banyaknya bayi (0-6 bulan) yang diberi ASI eksklusif (Z_1) memiliki pengaruh paling besar terhadap persentase balita gizi buruk di Purbalingga, sedangkan banyaknya balita yang menderita pneumonia (Z_4) memiliki pengaruh paling kecil terhadap persentase balita gizi buruk di Purbalingga. Selain itu, banyaknya bayi (0-6 bulan) yang diberi ASI eksklusif (Z_1), banyaknya bayi yang mendapat imunisasi BCG (Z_2), banyaknya balita yang mendapat vitamin A (Z_3), dan banyaknya rumah tangga yang

memiliki akses air bersih (Z_5) memiliki pengaruh negatif atau berbanding terbalik terhadap persentase balita gizi buruk, sedangkan banyaknya balita yang menderita pneumonia (Z_4) memiliki pengaruh positif atau berbanding lurus terhadap persentase balita gizi buruk.

4.4.1 Pengujian Hipotesis Model GWRPCA

Pada pengujian kesesuaian model GWRPCA dengan taraf signifikansi (α) sebesar 5% dan menolak H_0 jika $p\text{-value} < 0,05$ maka H_0 ditolak karena $p\text{-value}$ (0,01384) lebih kecil dari 0,05 sehingga menunjukkan ada perbedaan yang signifikan antara model regresi PCA dan GWRPCA.

Tabel 3 Uji Pengaruh Faktor Geografis secara Parsial

Variabel	F statistic	p-value
<i>Intercept</i>	4,0638	0,0004936
PC ₁	4,1680	0,0004423
PC ₂	3,0665	0,0050120

Berdasarkan Tabel 3 dengan taraf signifikansi (α) 5 % dan menolak H_0 jika $p\text{-value} < 0,05$ maka H_0 ditolak untuk *intercept*, PC₁ dan PC₂ karena mempunyai $p\text{-value}$ berturut-turut 0,0004936, 0,0004423 dan 0,0050120 ketiganya lebih kecil dari 0,05 sehingga ada perbedaan yang signifikan dari *intercept* serta variabel prediktor PC₁ dan PC₂ antara satu lokasi dengan lokasi lainnya.

4.4.2 Koefien Determinasi (R^2)

Dalam GWRPCA terdapat dua koefisien determinasi yaitu koefisien determinasi global dan koefisien determinasi lokal. Koefisien determinasi global pada pemodelan persentase balita gizi buruk menggunakan metode GWRPCA sebesar 0,6313309. Hal ini menunjukkan bahwa sebesar 63,13309% variasi total dalam variabel respon (Y) yaitu persentase balita gizi buruk dijelaskan oleh variabel prediktor yaitu banyaknya bayi (0-6 bulan) yang diberi ASI eksklusif (X_1), banyaknya bayi yang mendapat imunisasi BCG (X_2), banyaknya balita yang mendapat vitamin A (X_3), banyaknya balita yang menderita pneumonia (X_4), dan banyaknya rumah tangga yang memiliki akses air bersih (X_5), sedangkan sisanya sebesar 36,86691% dijelaskan oleh faktor lain.

Menurut Fotheringham *et al.* (2002) koefisien determinasi lokal menunjukkan bagaimana model lokal dapat meniru data yang tercatat dalam daerah sekitar titik regresi dengan baik. Koefisien determinasi lokal terbesar yaitu 0,72793026 terdapat pada Kabupaten Batang, sedangkan koefisien determinasi lokal terkecil yaitu 0,03519539 terdapat pada Kabupaten Sukoharjo.

5. KESIMPULAN

1. Variabel-variabel yang berpengaruh terhadap persentase balita gizi buruk di Jawa Tengah yaitu banyaknya bayi (0-6 bulan) yang diberi ASI eksklusif, banyaknya bayi yang mendapat imunisasi BCG, banyaknya balita yang mendapat vitamin A, banyaknya balita yang menderita pneumonia, dan banyaknya rumah tangga yang memiliki akses air bersih dapat diganti atau diwakili dengan variabel baru (komponen utama) hasil *Principal Components Analysis* (PCA) yaitu PC₁ dan PC₂ yang mampu menerangkan total varian data 78,43%.
2. Pemodelan persentase balita gizi buruk menggunakan metode GWRPCA berbeda-beda di setiap lokasi kabupaten dan kota di Jawa Tengah dengan koefisien determinasi global sebesar 0,6313309 dan koefisien determinasi lokal

terbesar yaitu 0,72793026 terdapat pada Kabupaten Batang, sedangkan koefisien determinasi lokal terkecil yaitu 0,03519539 terdapat pada Kabupaten Sukoharjo.

6. DAFTAR PUSTAKA

- [Dinkes] Dinas Kesehatan. 2013. *Profil Kesehatan Provinsi Jawa Tengah Tahun 2012*. Semarang: Dinas Kesehatan Provinsi Jawa Tengah.
- [FKM UI] Fakultas Kesehatan Masyarakat Universitas Indonesia. 2008. *Gizi dan Kesehatan Masyarakat*. Jakarta: PT RajaGrafindo Persada.
- Fotheringham, A.S., Brunson, C., and Charlton, M. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: John Wiley and Sons.
- Johnson, R.A. and Wichern, D.W. 2007. *Applied Multivariate Statistical Analysis*. Sixth edition. New Jersey: Pearson Prentice Hall.
- Mei C.L., He S.Y., Fang K.T. 2004. "A note on the mixed geographically weighted regression model", *Journal of Regional Science*, 44, 143-157.
- 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.
- Ramadani, I.R., Rahmawati, R., dan Hoyyi, A. 2013. Analisis Faktor-Faktor yang Mempengaruhi Gizi Buruk di Jawa Tengah dengan Metode Spatial Durbin Watson. *Jurnal Gaussian*, Vol.2, Nomor 4, hal 333-342.
- Sediaoetama, A. D. 2006. *Ilmu Gizi II*. Jakarta: Dian Rakyat.
- Simamora, B. 2005. *Analisis Multivariat Pemasaran*. Jakarta: PT Gramedia Pustaka Utama.