

Perluasan *Geographically Weighted Regression* Menggunakan Fungsi Polinomial

Toha Saifudin¹, Fatmawati², Nur Chamidah³
^{1,2,3} Program Studi Statistika, Universitas Airlangga Surabaya
Email: tohasaifudin@gmail.com

Info Artikel

Riwayat Artikel:

Diterima: 15 Mei 2017
Direvisi: 1 Juni 2017
Diterbitkan: 31 Juli 2017

Kata Kunci:

Geographically weighted regression
Weighted least square
Fungsi polinomial
Tingkat kesesuaian model

ABSTRAK

Geographically weighted regression (GWR) merupakan metode regresi pada data spasial dengan koefisien regresi bervariasi antar pengamatan. Dalam GWR, variabel-variabel bebas dan variabel tak bebas dihubungkan menggunakan fungsi linier. Sementara itu, dalam kondisi riil ada banyak kemungkinan kasus data spasial yang menunjukkan bahwa hubungan antara variabel tak bebas dengan variabel bebas cenderung tidak linier. Pemaksaan dalam menggunakan hubungan linier terhadap kasus tersebut bisa jadi merupakan salah satu faktor penyebab rendahnya kesesuaian model GWR. Oleh karena itu diperlukan perluasan fungsi pada model GWR. Tujuan paper ini adalah membuat model perluasan GWR menggunakan fungsi polinomial. Estimasi parameter model perluasan GWR diuraikan menggunakan prosedur *Weighted Least Square* (WLS). Hasil-hasil numerik berdasarkan studi kasus menunjukkan bahwa perluasan GWR dengan fungsi polinomial menghasilkan tingkat kesesuaian model yang lebih baik daripada GWR klasik.

Copyright © 2017 SIMANIS.
All rights reserved.

Korespondensi:

Toha Saifudin,
Program Studi Statistika, FST
Universitas Airlangga Surabaya,
Kampus C, Jl. Mulyorejo, Surabaya, Jawa Timur, Indonesia 60115
Email: tohasaifudin@gmail.com

1. PENDAHULUAN

Metode GWR telah diusulkan oleh Brunson dkk tahun 1996 [1]. Metode tersebut kemudian menjadi populer untuk analisis data spasial dalam sejumlah bidang, seperti: geografi, ekonometri, epidemiologi, dan ilmu lingkungan. Dalam model GWR, keberagaman model spasial dieksplorasi dengan cara mengakomodasi koefisien regresi yang beragam secara lokal. Dalam perkembangan teori, GWR telah dipelajari secara luas dalam beberapa penelitian. Beberapa contoh perkembangan teori tersebut adalah: GWR dibandingkan dengan metode *expansion* [5], usulan untuk uji kesesuaian model GWR [9], usulan untuk inferensi GWR berdasarkan *maximum likelihood* [11], beberapa perluasan GWR untuk model *generalized linear* [7], pembahasan tentang multikolinieritas diantara estimator GWR [13], dan evaluasi model GWR dibandingkan regresi linier global melalui suatu studi simulasi [8]. Di sisi lain terdapat beberapa penelitian yang mengembangkan penerapan GWR dalam berbagai bidang, misalnya: analisis perilaku sekolah menggunakan GWR [6], dan analisis kepuasan pelanggan kendaraan bermotor di Amerika menggunakan GWR [10].

Upaya perluasan model GWR telah dilakukan oleh beberapa peneliti. Salah satu contoh, model GWR semiparametrik [2], [7]. Perluasan tersebut dapat dipandang sebagai upaya untuk meningkatkan ketepatan atau kesesuaian model, karena salah satu faktor yang dapat berpengaruh pada ketepatan atau kesesuaian model adalah bentuk atau fungsi regresi yang digunakan.

Dalam kasus nyata data spasial, ada kemungkinan bahwa hubungan antara variabel tak bebas dengan variabel bebas cenderung tidak linier. Selama ini, semua kasus tersebut dimodelkan dengan pendekatan linier yaitu GWR. Pemaksaan dalam menggunakan hubungan linier tersebut merupakan salah satu faktor penyebab rendahnya kesesuaian model. Sebagai contoh, pengaruh faktor *convenience* terhadap harga perumahan di

Taipei, Taiwan menunjukkan pengaruh yang cenderung tidak linier [4]. Contoh lain, pada pemodelan tingkat kerawanan kejadian demam berdarah dengue di Surabaya berdasarkan model regresi logistik spasial, diperoleh hasil yang kurang memuaskan, yang ditengarai karena pendekatan fungsi linier yang digunakan kurang sesuai untuk kasus tersebut [3].

Termotivasi oleh hal-hal di atas, tujuan paper ini adalah membuat model perluasan GWR menggunakan pendekatan fungsi polinomial. Selanjutnya, estimasi parameter model perluasan GWR akan diuraikan menggunakan prosedur *Weighted Least Square* (WLS). Potensi perluasan model GWR menggunakan pendekatan fungsi polinomial ini akan dikaji melalui studi kasus dengan berdasar pada hasil-hasil numerik ukuran kesesuaian model.

2. MODEL GWR KLASIK

Menurut beberapa referensi, seperti [7] dan [9], model GWR dibentuk dengan membuat koefisien regresi linier bervariasi secara lokal dalam bentuk

$$y_i = \sum_{j=1}^p \beta_j(u_i, v_i) x_{ij} + \varepsilon_i, i = 1, 2, \dots, n, \quad (1)$$

dengan $(y_i, x_{i1}, x_{i2}, \dots, x_{ip})$ adalah nilai observasi variabel tak bebas y dan variabel-variabel bebas $x_1, x_2, \dots$, dan x_p di lokasi (u_i, v_i) dalam area geografis yang diamati, $\beta_j(u_i, v_i), j = 1, 2, \dots, p$ adalah parameter atau koefisien regresi di (u_i, v_i) , dan $\varepsilon_i, i = 1, 2, \dots, n$, adalah galat model yang diasumsikan berdistribusi Normal dengan rata-rata 0 dan variansi σ^2 yang biasanya dinyatakan dengan $\varepsilon_i \sim N(0, \sigma^2)$. Untuk membedakan dengan model perluasannya, dalam paper ini penulis menyebut model (1) sebagai model GWR klasik. Model tersebut dapat mengakomodasi koefisien intersep dengan memilih $x_{i1} = 1$, untuk $i = 1, 2, \dots, n$.

Untuk suatu lokasi (u_0, v_0) yang diberikan dalam area geografis yang diamati, misalkan d_{0i} merupakan jarak antara lokasi (u_0, v_0) dengan lokasi (u_i, v_i) . Parameter regresi dalam model (1) di lokasi (u_0, v_0) dapat diestimasi secara lokal dengan metode *weighted least square* (WLS), yaitu dengan meminimumkan jumlah kuadrat galat terboboti dalam bentuk

$$\sum_{i=1}^n (y_i - \sum_{j=1}^p \beta_j(u_0, v_0) x_{ij})^2 K_h(d_{0i}) \quad (2)$$

terhadap $\beta_j(u_i, v_i)$, dengan $K_h(\cdot) = K(\frac{\cdot}{h})$ untuk $K(\cdot)$ adalah fungsi kernel yang diberikan, dan h adalah *bandwidth*. Penyelesaian masalah WLS menghasilkan estimator $\hat{\beta}_j(u_0, v_0)$ untuk $j = 1, 2, \dots, p$ di lokasi (u_0, v_0) . Dalam bentuk matriks, estimator parameter regresi dapat dinyatakan dalam bentuk

$$\begin{aligned} \hat{\boldsymbol{\beta}}(u_0, v_0) &= [\hat{\beta}_1(u_0, v_0), \hat{\beta}_2(u_0, v_0), \dots, \hat{\beta}_p(u_0, v_0)]^T \\ &= [\mathbf{X}^T \mathbf{W}(u_0, v_0) \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{W}(u_0, v_0) \mathbf{y}, \end{aligned} \quad (3)$$

dengan

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, \mathbf{y} = (y_1, y_2, \dots, y_n)^T, \quad (4)$$

dan

$$\mathbf{W}(u_0, v_0) = \text{diag}[K_h(d_{01}), K_h(d_{02}), \dots, K_h(d_{0n})] \quad (5)$$

adalah matrik pembobot. Dengan menerapkan estimasi tersebut pada setiap lokasi $(u_i, v_i), i = 1, 2, \dots, n$, dapat diperoleh vektor nilai estimasi untuk variabel tak bebas y di n lokasi $(u_i, v_i), i = 1, 2, \dots, n$ sebagai berikut:

$$\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)^T = \mathbf{L} \mathbf{y}, \quad (6)$$

dengan

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

dinamakan *hat matrix*, dan $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ adalah baris ke- i dari matriks $\mathbf{X}$ dalam persamaan (4). Berdasarkan *hat matrix* $\mathbf{L}$, vektor residual dapat dinyatakan sebagai

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{L}) \mathbf{y}, \quad (8)$$

dan jumlah kuadrat galat (JKG) adalah

$$\text{JKG} = \hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} = \mathbf{y}^T (\mathbf{I} - \mathbf{L})^T (\mathbf{I} - \mathbf{L}) \mathbf{y}, \quad (9)$$

dengan $\mathbf{I}$ adalah matriks identitas orde n . Untuk menentukan nilai *bandwidth* optimal umumnya digunakan prosedur *cross validation* (CV). Dalam prosedur CV, skor

$$\text{CV}(h) = \sum_{i=1}^n (y_i - \hat{y}_{(i)}(h))^2 \quad (10)$$

digunakan sebagai fungsi obyektif, dengan $\hat{y}_{(i)}(h)$ adalah nilai estimasi dari y_i untuk suatu *bandwidth* h dengan observasi pada lokasi (u_i, v_i) tidak ikut digunakan dalam proses estimasi tersebut. Pilih h_0 sebagai nilai *bandwidth* optimal, sedemikian hingga

$$CV(h_0) = \min CV(h), h > 0. \quad (11)$$

3. MODEL GWR DIPERLUAS MENGGUNAKAN FUNGSI POLINOMIAL

Regresi polinomial merupakan sebuah perluasan dari regresi linier dengan menggunakan fungsi polinomial [12]. Dalam regresi polinomial, fungsi regresi dikembangkan sebagai fungsi polinomial orde d dari variabel bebas yang dapat dinyatakan sebagai

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^d + \varepsilon. \quad (12)$$

Dalam paper ini, GWR klasik diperluas menggunakan fungsi polinomial dengan cara mengganti hubungan linier pada persamaan (1) dengan fungsi polinomial, sehingga model menjadi

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p f_{ik}(x_{ik}) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (13)$$

dengan $\varepsilon_i \sim N(0, \sigma^2)$ dan $f_{ik}(x_{ik})$ adalah fungsi polinomial dari variabel bebas ke- k . Jika fungsi polinomial dinyatakan sebagai $f_{ik}(x_{ik}) = \sum_{j=1}^{d_k} \beta_{k,j}(u_i, v_i) x_{ik}^j$, maka persamaan (13) dapat dinyatakan sebagai

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \sum_{j=1}^{d_k} \beta_{k,j}(u_i, v_i) x_{ik}^j + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (14)$$

Model (14) dinamakan perluasan GWR menggunakan fungsi polinomial. Untuk tujuan kesederhanaan penulisan, selanjutnya penamaan model tersebut disingkat menjadi GWRPol. Dalam format matriks, model (14) dapat dinyatakan sebagai

$$y_i = \mathbf{x}_i^{*T} \boldsymbol{\beta}_{pol}(u_i, v_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (15)$$

dengan

$$\mathbf{x}_i^{*T} = (1 \ x_{i1} \ x_{i1}^2 \ \dots \ x_{i1}^{d_1} \ \dots \ x_{ip} \ x_{ip}^2 \ \dots \ x_{ip}^{d_p}), \quad (16)$$

dan

$$\boldsymbol{\beta}_{pol}^T(u_i, v_i) = (\beta_0(u_i, v_i) \ \beta_{1,1}(u_i, v_i) \ \beta_{1,2}(u_i, v_i) \ \dots \ \beta_{1,d_1}(u_i, v_i) \ \dots \ \beta_{p,1}(u_i, v_i) \ \beta_{p,2}(u_i, v_i) \ \dots \ \beta_{p,d_p}(u_i, v_i)). \quad (17)$$

Untuk suatu lokasi (u_0, v_0) , $\boldsymbol{\beta}_{pol}^T(u_0, v_0)$ dapat diestimasi dengan meminimumkan fungsi jumlah kuadrat galat

$$\sum_{i=1}^n (y_i - \mathbf{x}_i^{*T} \boldsymbol{\beta}_{pol}(u_0, v_0))^2 K_h(d_{0i}), \quad (18)$$

terhadap masing-masing elemen vektor $\boldsymbol{\beta}_{pol}(u_0, v_0)$, dengan $K_h(d_{0i})$ adalah sama seperti yang tertulis di persamaan (2).

Misalkan matriks observasi ditulis sebagai

$$\mathbf{X}_{pol} = \begin{bmatrix} 1 & x_{11} & x_{11}^2 & \dots & x_{11}^{d_1} & \dots & x_{1p} & x_{1p}^2 & \dots & x_{1p}^{d_p} \\ 1 & x_{21} & x_{21}^2 & \dots & x_{21}^{d_1} & \dots & x_{2p} & x_{2p}^2 & \dots & x_{2p}^{d_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n1}^2 & \dots & x_{n1}^{d_1} & \dots & x_{np} & x_{np}^2 & \dots & x_{np}^{d_p} \end{bmatrix}, \quad (19)$$

dan vektor parameter di suatu lokasi (u_0, v_0) ditulis sebagai

$$\boldsymbol{\beta}_{pol}^T(u_0, v_0) = (\beta_0(u_0, v_0) \ \beta_{1,1}(u_0, v_0) \ \beta_{1,2}(u_0, v_0) \ \dots \ \beta_{1,d_1}(u_0, v_0) \ \dots \ \beta_{p,1}(u_0, v_0) \ \beta_{p,2}(u_0, v_0) \ \dots \ \beta_{p,d_p}(u_0, v_0)). \quad (20)$$

Misalkan

$$\hat{\boldsymbol{\beta}}_{pol}^T(u_0, v_0) = (\hat{\beta}_0(u_0, v_0) \ \hat{\beta}_{1,1}(u_0, v_0) \ \hat{\beta}_{1,2}(u_0, v_0) \ \dots \ \hat{\beta}_{1,d_1}(u_0, v_0) \ \dots \ \hat{\beta}_{p,1}(u_0, v_0) \ \hat{\beta}_{p,2}(u_0, v_0) \ \dots \ \hat{\beta}_{p,d_p}(u_0, v_0)), \quad (21)$$

menotasikan estimator untuk parameter pada persamaan (20), maka penyelesaian masalah kuadrat terkecil pada persamaan (18) di lokasi (u_0, v_0) dapat dinyatakan sebagai

$$\hat{\boldsymbol{\beta}}_{pol}^T(u_0, v_0) = [\mathbf{X}_{pol}^T \mathbf{W}(u_0, v_0) \mathbf{X}_{pol}]^{-1} \mathbf{X}_{pol}^T \mathbf{W}(u_0, v_0) \mathbf{y}, \quad (22)$$

dengan $\mathbf{y}$ dan $\mathbf{W}(u_0, v_0)$ masing-masing seperti pada persamaan (4) dan (5).

Dengan menerapkan prosedur estimasi tersebut pada setiap lokasi (u_i, v_i) , $i = 1, 2, \dots, n$, dapat diperoleh vektor nilai estimasi model GWRPol untuk variabel tak bebas y di n lokasi sebagai

$$\hat{\mathbf{y}}_{pol} = (\hat{y}_1^*, \hat{y}_2^*, \dots, \hat{y}_n^*)^T = \mathbf{C} \mathbf{y}, \quad (23)$$

dengan

$$\mathbf{C} = \begin{bmatrix} \mathbf{x}_1^{*T} [\mathbf{X}_{pol}^T \mathbf{W}(u_1, v_1) \mathbf{X}_{pol}]^{-1} \mathbf{X}_{pol}^T \mathbf{W}(u_1, v_1) \\ \mathbf{x}_2^{*T} [\mathbf{X}_{pol}^T \mathbf{W}(u_2, v_2) \mathbf{X}_{pol}]^{-1} \mathbf{X}_{pol}^T \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_n^{*T} [\mathbf{X}_{pol}^T \mathbf{W}(u_n, v_n) \mathbf{X}_{pol}]^{-1} \mathbf{X}_{pol}^T \mathbf{W}(u_n, v_n) \end{bmatrix} \quad (24)$$

dinamakan *hat matrix* dari model GWRPol, dan $\mathbf{x}_i^{*T}$ adalah vektor yang tertulis di persamaan (16) atau baris ke- i dari matriks $\mathbf{X}_{pol}$ dalam persamaan (19). Berdasarkan matriks $\mathbf{C}$, vektor residual dapat ditulis sebagai

$$\hat{\boldsymbol{\varepsilon}}_{pol} = \mathbf{y} - \hat{\mathbf{y}}_{pol} = (\mathbf{I} - \mathbf{C})\mathbf{y}, \quad (25)$$

dan JKG model GWRPol adalah

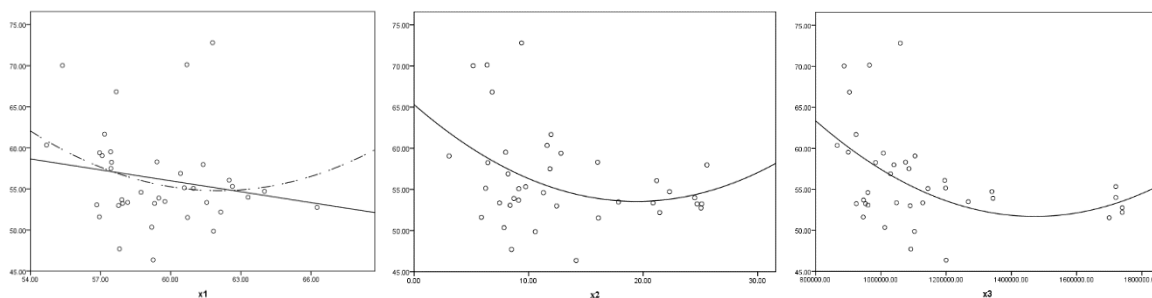
$$\text{JKG}_{pol} = \hat{\boldsymbol{\varepsilon}}_{pol}^T \hat{\boldsymbol{\varepsilon}}_{pol} = \mathbf{y}^T (\mathbf{I} - \mathbf{C})^T (\mathbf{I} - \mathbf{C}) \mathbf{y}, \quad (26)$$

dengan $\mathbf{I}$ adalah matriks identitas orde n .

4. STUDI KASUS

Sebagai contoh kasus, dalam paper ini digunakan data tingkat partisipasi angkatan kerja wanita 38 kabupaten/kota di Jawa Timur, Indonesia pada tahun 2014. Data diperoleh dari BPS propinsi Jawa Timur. Variabel-variabel yang terlibat dalam model antara lain: Tingkat Partisipasi Angkatan Kerja (TPAK) wanita dalam % (Y), penduduk wanita yang berusia 15 – 64 tahun dalam % (X_1), penduduk wanita berdasarkan tingkat pendidikan terakhir tamat SMA dalam % (X_2), dan Upah Minimum Kabupaten/ Kota (UMK) dalam Rupiah (X_3).

Diagram pencar dan kemungkinan pola garis regresi yang mendekati untuk ketiga variabel bebas X terhadap variabel tak bebas Y dapat dilihat pada Gambar 1. Secara visual terlihat bahwa variabel X_2 dan X_3 cenderung menunjukkan pola hubungan yang tidak linier terhadap Y , sedangkan variabel X_1 dapat mengambil pola linier maupun tidak linier. Pola tersebut sebagai indikator awal untuk penerapan model GWRPol. Selanjutnya, dalam paper ini dilakukan uji Ramsey RESET untuk menentukan apakah fungsi linier cocok digunakan untuk memodelkan sampel. Hipotesis nol menyatakan bahwa data cocok dimodelkan menggunakan fungsi linier, sedangkan hipotesis alternatif menyatakan sebaliknya. Dengan software R, diperoleh nilai statistik uji F-RESET = 4,8518, dengan derajat bebas pembilang 2 dan derajat bebas penyebut 32, dan p -value sebesar 0,01444. Dengan mengambil tingkat signifikansi $\alpha = 0,05$ maka hipotesis nol ditolak dan disimpulkan bahwa data tidak cocok dimodelkan menggunakan fungsi linier.

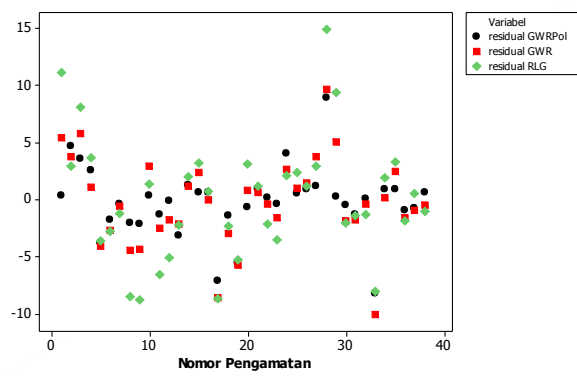


Gambar 1. Diagram pencar dengan pendekatan garis regresi antara x_1 (kiri), x_2 (tengah), dan x_3 (kanan) terhadap y

Estimasi parameter regresi spasial dalam paper ini menggunakan fungsi pembobot kernel Gaussian. Berdasarkan pembobot tersebut, diperoleh bandwidth optimal untuk pemodelan GWR klasik sebesar 0,37 dengan nilai CV sebesar 4084,158. Selanjutnya untuk pemodelan GWRPol, eksperimen dilakukan dengan orde polinomial bagi tiap-tiap variabel bebas dibatasi maksimum 2. Eksperimen menghasilkan orde optimal sebesar 1 untuk variabel bebas X_1 , dan sebesar 2 untuk masing-masing variabel bebas X_2 dan X_3 . Pada kondisi demikian, diperoleh bandwidth optimal sebesar 0,36 dan nilai CV sebesar 1335,769.

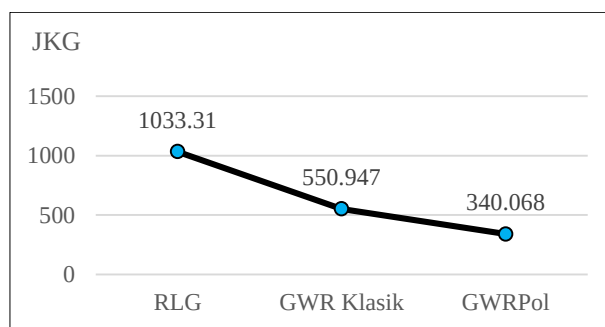
Evaluasi terhadap hasil estimasi model RLG, GWR Klasik, dan GWRPol dilakukan dengan mengamati beberapa ukuran kesesuaian model. Ukuran tersebut diantaranya adalah simpangan baku residual, jumlah kuadrat galat, dan koefisien determinasi model. Simpangan baku residual menjadi salah satu kriteria kesesuaian model dalam paper ini. Semakin kecil nilai simpangan baku residual berarti semakin sesuai model yang digunakan. Residual model atas 38 kabupaten/kota di Jawa Timur yang dihasilkan oleh RLG, GWR Klasik, dan GWRPol dapat dilihat pada Gambar 2. Dari gambar tersebut terlihat bahwa sebaran residual GWRPol cenderung lebih sempit daripada residual model yang lain. Hasil perhitungan simpangan baku residual untuk

RLG, GWR Klasik, dan GWRPol masing-masing adalah 5,51; 3,85; dan 3,02. Hasil tersebut menunjukkan bahwa simpangan baku residual model GWRPol adalah yang paling rendah. Hal ini mengindikasikan bahwa estimasi tingkat partisipasi angkatan kerja wanita kabupaten/kota di Jawa Timur menggunakan GWRPol cenderung lebih baik daripada menggunakan GWR Klasik ataupun RLG.



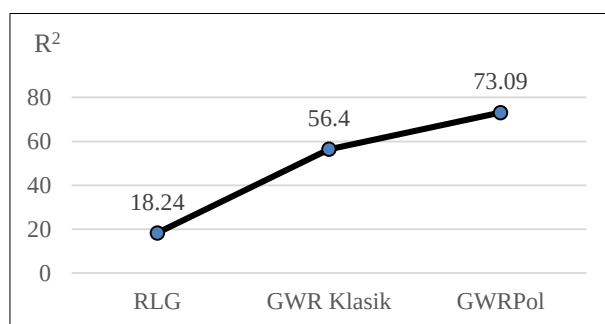
Gambar 2. Residual model RLG, GWR Klasik, dan GWRPol untuk data sampel

Jumlah Kuadrat Galat (JKG) merupakan kriteria yang umumnya dipakai untuk kesesuaian model. Semakin rendah nilai JKG berarti semakin sesuai model yang digunakan. JKG ketiga model atas data tingkat partisipasi angkatan kerja wanita 38 kabupaten/kota di Jawa Timur disajikan dalam Gambar 3. Hasil tersebut menunjukkan bahwa JKG model GWRPol adalah yang paling rendah. Hal ini mengindikasikan bahwa model GWRPol cenderung lebih sesuai daripada GWR Klasik maupun RLG dalam memodelkan hubungan regresi tingkat partisipasi angkatan kerja wanita kabupaten/kota di Jawa Timur.



Gambar 3. JKG model RLG, GWR Klasik, dan GWRPol berdasarkan data sampel

Selain ukuran di atas, kesesuaian model biasanya diukur dengan menggunakan koefisien determinasi yang dinotasikan dengan R^2 . Koefisien determinasi menyatakan persentase besarnya keragaman respon yang dapat dijelaskan oleh model. Semakin besar nilai R^2 berarti semakin baik model yang digunakan dalam menjelaskan keragaman respon. Berdasarkan data sampel, nilai R^2 model RLG, GWR Klasik, dan GWRPol masing-masing adalah 18,24%; 56,40%; dan 73,09%. Hal tersebut mengindikasikan bahwa model GWRPol lebih dapat menjelaskan variansi tingkat partisipasi angkatan kerja wanita kabupaten/kota di Jawa Timur daripada model GWR Klasik ataupun RLG. Secara visual, *slope* kenaikan nilai R^2 dari RLG ke GWR Klasik dan dari GWR Klasik ke GWRPol dapat dilihat pada Gambar 4.



Gambar 4. Nilai R^2 model RLG, GWR Klasik, dan GWRPol berdasarkan data sampel

Berdasarkan fakta-fakta tersebut, pemodelan GWRPol memiliki potensi untuk menjadi model yang paling layak digunakan diantara tiga model yang ada dalam memodelkan data tingkat partisipasi angkatan kerja wanita kabupaten/kota di Jawa Timur. Meskipun demikian, dari sudut pandang statistika masih terdapat pertanyaan besar yaitu apakah dapat disimpulkan secara statistika bahwa GWRPol signifikan lebih baik dari GWR Klasik maupun RLG dalam memodelkan data tersebut?.

Terdapat kelebihan dan kekurangan pada masing-masing model. Biasanya model dengan orde lebih tinggi atau jumlah parameter lebih banyak, dapat menggambarkan data lebih baik daripada model orde rendah. Namun model tersebut biasanya memiliki kelemahan pada rumitnya aplikasi dan interpretasi. Di sisi lain, kelebihan model dengan orde lebih rendah atau jumlah parameter lebih sedikit adalah biasanya mudah diaplikasikan dan diinterpretasikan. Dalam pemilihan model biasanya digunakan prinsip parsimoni, yaitu memilih model sederhana untuk kesederhanaan interpretasi. Jika model GWRPol tidak menunjukkan performa yang signifikan lebih baik dari GWR Klasik, maka GWR Klasik lebih baik dipilih karena alasan prinsip parsimoni. Di sisi lain, jika GWRPol menunjukkan performa yang secara statistika signifikan lebih baik daripada GWR, maka GWRPol lebih baik dipilih dan digunakan dalam pemodelan. Dengan demikian, masih diperlukan penelitian lanjutan untuk mengkaji uji signifikansi kesesuaian model GWRPol.

5. KESIMPULAN

Berdasarkan data tingkat partisipasi angkatan kerja wanita 38 kabupaten/kota di Jawa Timur tahun 2014, identifikasi awal menunjukkan ada hubungan tidak linier diantara beberapa variabel bebas terhadap variabel tak bebas. Berdasarkan kriteria simpangan baku residual, JKG, dan R^2 dapat disimpulkan secara empiris bahwa GWRPol lebih sesuai digunakan daripada GWR Klasik ataupun RLG untuk kasus tersebut. Hal ini menunjukkan bahwa GWRPol memiliki potensi sebagai model yang lebih baik daripada GWR klasik jika terdapat indikasi hubungan tidak linier dalam data sampel. Namun untuk dapat menyimpulkan secara statistika bahwa GWRPol signifikan lebih baik, diperlukan penelitian lanjutan untuk mengkaji uji kesesuaian model GWRPol.

REFERENSI

- [1] Brunson, C., Fotheringham A.S., & Charlton, M. Geographically Weighted Regression: A Method For Exploring Spatial Nonstationarity. *Geographical Analysis*. 1996; 28(4):281–298.
- [2] Brunson, C., Fotheringham A.S., & Charlton, M. Some Notes on Parametric Significance Tests for Geographically Weighted Regression. *Journal of Regional Science*. 1999; 38(3): 497–524.
- [3] Chamidah, N., Saifuddin, T., dan Rifada, M. The Vulnerability Modeling of Dengue Hemorrhagic Fever Disease in Surabaya Based on Spatial Logistic Regression Approach, *Applied Mathematical Sciences*. 2014; 8(28): 1369 – 1379.
- [4] Chiang, Y-H, Peng, T-C, & Chang, C-O. The nonlinear effect of convenience stores on residential property prices: A case study of Taipei, Taiwan. *Habitat International*. 2015; 46: 82–90.
- [5] Fotheringham, A.S., Charlton, M.E., & Brunson, C. Geographically Weighted Regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*. 1998; 30: 1905–1927.
- [6] Fotheringham, A.S., Charlton, M.E., & Brunson, C. Spatial Variation in School Performance: a Local Analysis Using Geographically Weighted Regression. *Geographical & Environmental Modelling*. 2001; 5(1): 43–66.
- [7] Fotheringham A.S., Brunson C., & Charlton M. **Geographically Weighted Regression: The Analysis Of Spatially Varying Relationships**. John Wiley and Sons, USA. 2002.
- [8] Harris, P., Fotheringham, A.S., Crespo, R., & Charlton, M. The Use of Geographically Weighted Regression for Spatial Prediction: An Evaluation of Models Using Simulated Data Sets. *Math GeoSci*. 2010; 42: 657–680.
- [9] Leung, Y., Mei, C.L., & Zhang, W.X. Statistical Tests for Spatial Nonstationarity based on The Geographically Weighted Regression Model. *Environment and Planning A*. 2000; 32: 9–32.
- [10] Mittal, V., Kamakura, W.A., & Govind, R. Geographic Pattern in Customer Service and Satisfaction: An Empirical Investigation. *Journal of Marketing*. 2004; 68: 48–62.
- [11] Paez, A., Uchida, T., & Miyamoto, K. A general Framework for Estimation and Inference of Geographically Weighted Regression Models: 2. Spatial Association and Model Specification Tests. *Environment and Planning A*. 2002; 34: 883–904.
- [12] Rawlings, J.O., Pantula, S.G., & Dickey, D.A. **Applied Regression Analysis: A Research Tool**, 2nd Edition. Springer-Verlag New York, Inc., USA. 1998.
- [13] Wheeler, D., & Tiefelsdorf, M. Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression. *Journal of Geographical Systems*. 2005; 7: 161–187.