

PEMODELAN REGRESI BERGANDA DAN *GEOGRAPHICALLY WEIGHTED REGRESSION* PADA TINGKAT PENGANGGURAN TERBUKA DI JAWA TENGAH

Tiani Wahyu Utami¹, Abdul Rohman², Alan Prahutama³

^{1,2}Jurusan Statistika Universitas Muhammadiyah Semarang

³Departemen Statistika, Universitas Diponegoro

e-mail: tiani.utami88@gmail.com

DOI: 10.14710/medstat.9.2.133-147

Abstract

The problems in employment was the growing number of Open Unemployment Rate (OUR). The open unemployment rate is a number that indicates the number of unemployed to the 100 residents are included in the labor force. The purpose of this study is mapping the data OUR in Central Java and the suspect and identify linkages between factors that cause OUR in the District / City of Central Java in 2014. Factors that allegedly include population density (X_1), Inflation (X_2), the GDP value (X_3), UMR Value (X_4), the percentage of GDP growth rate (X_5), Hope of the old school (X_6), the percentage of the labor force by age (X_7) and the percentage of employment (X_8). Geographically Weighted Regression (GWR) is a method for modeling the response of the predictor variables, by including elements of the area (spatial) into the point-based model. This research resulted in the conclusion that the OLS regression models have poor performance because the residual variance is not homogeneous. There were no significant differences between GWR models with OLS model or in other words generally predictor variables did not affect the response variable (rate of unemployment in Central Java) spatially. However, GWR model could captured modelling in each region.

Keywords: *multiple linear regression, geographiically weighted regression, open unemployment rate in Central Java.*

1. PENDAHULUAN

Permasalahan yang menghambat pada program pembangunan ketenagakerjaan adalah semakin meningkatnya jumlah pengangguran. Indikator yang biasa digunakan untuk mengukur angka pengangguran mengikuti konsep ketenagakerjaan adalah Tingkat Pengangguran Terbuka (TPT), yaitu perbandingan antara banyaknya pencari kerja dengan angkatan kerja. Tingginya angka pengangguran dipengaruhi antara lain oleh presentase penduduk usia kerja berumur 15 tahun keatas berdasarkan pendidikan tertinggi yang ditamatkan, kepadatan penduduk, angka melek huruf, PDRB, tingkat inflasi, tingkat partisipasi angkatan kerja, upah minimum kabupaten/kota, pertumbuhan ekonomi daerah.

Untuk mengetahui hubungan antara TPT dengan variabel yang mempengaruhinya digunakan metode regresi klasik. Dalam pemodelan regresi klasik harus memenuhi beberapa asumsi klasik diantaranya data berdistribusi normal, non-autokorelasi,

homoskedastisitas, tidak ada multikolinearitas, dan penduga parameter yang dihasilkan berlaku global (Supranto, 2001). Berbagai penelitian yang berkaitan dengan angka pengangguran telah dilakukan, diantaranya analisis tingkat pengangguran dan faktor yang mempengaruhi di Kota Semarang (Mustika dan Agustina, 2010), pemodelan TPT dengan pendekatan regresi spline multivariabel (Sari dan Budiantara, 2012), analisis PDRB, inflasi, upah minimum dan angka melek huruf terhadap TPT di Provinsi Jawa Tengah (Hajji dan Nugroho, 2013). Penelitian-penelitian tersebut tidak menekankan aspek humaniora. Aspek humaniora, seperti karakteristik ragam budaya yang direpresentasikan karakteristik daerah (kabupaten/kota) masih terbatas untuk dikaji. Oleh karena itu dalam penelitian ini akan dikembangkan pemodelan tingkat pengangguran terbuka yang mengakomodasi adanya aspek perilaku masyarakat yang direpresentasikan dalam spasial (lokasi).

Geographically Weighted Regression (GWR) merupakan salah satu metode statistika yang digunakan untuk memodelkan variabel respon dengan variabel prediktor yang berbasis wilayah atau area (Anselin dan Bera, 1998). Keunggulan model GWR dibandingkan dengan model regresi klasik adalah GWR mampu memberikan model secara lokal. Salah satu penelitian yang dilakukan dengan menggunakan metode GWR adalah pemodelan GWR pada jumlah kematian bayi di Jawa Timur dan Jawa Tengah oleh Aulele (2010). Selain itu, metode GWR juga diterapkan oleh Intan (2010) untuk memodelkan jumlah penderita tuberculosis (TB) di Jawa Timur. Dengan menggunakan metode GWR diharapkan dapat lebih menjelaskan dan mengidentifikasi faktor-faktor yang mempengaruhi tingkat pengangguran terbuka. Berdasarkan penjelasan tersebut akan dilakukan pemodelan TPT di Jawa Tengah berdasarkan faktor-faktor yang mempengaruhi antara lain, Kepadatan penduduk (X_1), Tingkat Inflasi (X_2), Nilai PDRB (X_3), Nilai UMR (X_4), prosentase Laju pertumbuhan PDRB (X_5), Harapan lama sekolah (X_6), prosentase angkatan kerja menurut usia (X_7) dan prosentase kesempatan kerja (X_8) menggunakan GWR dan regresi berganda. Selain itu akan dibandingkan pemodelan TPT menggunakan GWR dan regresi berganda.

2. TINJAUAN PUSTAKA

Beberapa perkembangan model GWR antara lain Chiou *et. al* (2015) melakukan penelitian tentang faktor-faktor penggunaan transportasi publik di beberapa wilayah di Taiwan. Variabel prediktor yang digunakan meliputi beberapa aspek antara lain sosial ekonomi, transportasi umum, dan transportasi pribadi. Pada penelitian tersebut membandingkan model regresi GWR dengan regresi klasik. Hasil yang didapat adalah model GWR memberikan akurasi yang bagus dibandingkan dengan model regresi klasik. Saefudin *et. al* (2011) membandingkan antara model regresi berganda dengan model regresi GWR pada kasus tingkat kemiskinan di Indonesia. Berdasarkan penelitian tersebut model yang dihasilkan GWR lebih bagus dibandingkan dengan regresi berganda.

2.1. Regresi Linier Berganda

Metode regresi linier yang merupakan metode yang memodelkan hubungan antara variabel respon y dan variabel prediktor x_1, x_2, \dots, x_p . Model regresi linier untuk p variabel prediktor secara umum ditulis sebagai:

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

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 nilai residual disetiap pengamatan yang diasumsikan identik, independen, dan berdistribusi Normal dengan mean nol dan varians konstan σ^2 ($\varepsilon_i \sim IIDN(0, \sigma^2)$). Pada model ini, hubungan antara variabel prediktor dan variabel respon dianggap konstan pada setiap lokasi pengamatan. Jika dituliskan dalam notasi matriks maka persamaan (1) menjadi:

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

dengan:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Sedangkan menurut Drapper dan Smith (1992) nilai estimasi untuk \mathbf{y} dan $\boldsymbol{\varepsilon}$ adalah:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad \text{dan} \quad \hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}.$$

2.1.1. Estimasi Parameter $\boldsymbol{\beta}$ Model Regresi Linier

Estimator dari parameter model didapat dengan meminimumkan jumlah kuadrat residual atau yang dikenal dengan *Ordinary Least Square* (OLS), yaitu:

$$\hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} = \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}}$$

Untuk menentukan nilai $\hat{\boldsymbol{\beta}}$ yang meminimumkan $\hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}}$ adalah dengan menurunkan $\hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}}$ terhadap $\hat{\boldsymbol{\beta}}$ dan kemudian disamadengankan nol, sehingga diperoleh:

$$\frac{\partial \hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}}}{\partial \hat{\boldsymbol{\beta}}} = 0 - 2\mathbf{y}^T \mathbf{X} + 2\mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} = 0$$

dan estimasi untuk $\hat{\boldsymbol{\beta}}$ adalah:

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

$\hat{\boldsymbol{\beta}}$ merupakan penaksir yang tak bias untuk $\boldsymbol{\beta}$ (Drapper dan Smith, 1992).

2.1.2. Pengujian Parameter Model Regresi Linier

Untuk menguji pengaruh variabel prediktor secara bersama-sama digunakan analisis varian yang dibuat dengan cara menguraikan bentuk jumlah kuadrat total atau *Sum Square Total* (SST) menjadi dua komponen: Jumlah kuadrat regresi atau *Sum Square*

Regression (SSR) dan Jumlah Kuadrat Error atau *Sum Square Error* (SSE) (Drapper dan Smith, 1992).

Hipotesisnya adalah:

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

$$H_1 : \text{minimal ada satu } \beta_k \neq 0, k = 1, 2, \dots, p$$

Tabel 1. Analisis Varians Model Regresi

Sumber Variasi	Jumlah Kuadrat	Derajat Bebas	Rata-rata Kuadrat	F Hitung
Regresi	$\mathbf{y}^T (\mathbf{H} - \frac{1}{n} \mathbf{J}) \mathbf{y}$	p	$MSR = SSR / p$	$F = \frac{MSR}{MSE}$
Error	$\mathbf{y}^T (\mathbf{I} - \mathbf{H}) \mathbf{y}$	$n - (p + 1)$	$MSE = \frac{SSE}{n - (p + 1)}$	
Total	$\mathbf{y}^T (\mathbf{I} - \frac{1}{n} \mathbf{J}) \mathbf{y}$	$n - 1$		

\mathbf{J} adalah matriks berukuran $n \times n$ dengan semua elemen bernilai 1 dan $\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$.

Kriteria ujinya adalah tolak H_0 bila $F_{hitung} > F_{\alpha, p, (n-p-1)}$ atau p -value $< \alpha$.

Adapun pengujian secara parsial untuk mengetahui parameter mana saja yang signifikan terhadap model dilakukan dengan hipotesis:

$$H_0 : \beta_k = 0, k = 1, 2, \dots, p$$

$$H_1 : \beta_k \neq 0, k = 1, 2, \dots, p$$

Statistik uji dalam pengujian parsial ini adalah $t = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)}$ (Gujarati, 2004), 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 $s = \sqrt{MSE}$. Kriteria pengambilan keputusan tolak H_0 jika $|t_{hit}| > t_{\alpha/2; n-p-1}$.

2.2. Model GWR

Model *Geographically Weighted Regression* (GWR) adalah model regresi yang dikembangkan oleh Brunson *et al.* (2002) untuk variabel respon yang bersifat kontinu yang mempertimbangkan aspek lokasi. Model GWR merupakan model regresi linier lokal (*locally linier regression*) yang menghasilkan estimator parameter model yang mempunyai sifat kelokalan pada masing-masing titik atau lokasi. Menurut Brunson *et al.* (2002), model GWR dapat ditulis:

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

dengan:

- y_i : nilai observasi variabel respon pada lokasi ke- i
- x_{ik} : nilai observasi variabel prediktor k pada lokasi ke- i
- $\beta_0(u_i, v_i)$: nilai intersep model regresi GWR
- $\beta_k(u_i, v_i)$: parameter regresi untuk setiap lokasi ke- i
- (u_i, v_i) : titik koordinat (lintang, bujur) pada lokasi ke- i
- ε_i : residual model GWR pada lokasi ke- i yang diasumsikan IIDN (identik, independen, dan berdistribusi normal) dengan rata-rata nol dan varians konstan σ^2

Estimasi model GWR menggunakan WLS (*Weighted Least Square*), yaitu dengan memberikan bobot yang berbeda disetiap lokasi. Misalkan pembobot untuk setiap lokasi (u_i, v_i) adalah $w_j(u_i, v_i)$, $j = 1, 2, \dots, n$ maka parameter pada lokasi pengamatan (u_i, v_i) diestimasi dengan menambahkan unsur pembobot $w_j(u_i, v_i)$ dan kemudian meminimumkan jumlah kuadrat residual berikut ini:

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

Atau dalam bentuk matriks jumlah kuadrat residualnya adalah:

$$\begin{aligned} \boldsymbol{\varepsilon}^T \mathbf{W}(u_i, v_i) \boldsymbol{\varepsilon} &= \mathbf{y}^T \mathbf{W}(u_i, v_i) \mathbf{y} - 2\boldsymbol{\beta}^T(u_i, v_i) \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \\ &+ \boldsymbol{\beta}^T(u_i, v_i) \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X} \boldsymbol{\beta}(u_i, v_i) \end{aligned} \quad (4)$$

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

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = [\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \quad (5)$$

2.3. Pengujian Kesesuaian Model (*Goodness of Fit*)

Pengujian kesesuaian model (*goodness of fit*) dilakukan dengan menguji kesesuaian dari koefisien parameter secara serentak, yaitu dengan cara mengkombinasikan uji regresi linier dengan model untuk data spasial. Bentuk hipotesis pengujian menurut Brundson *et al.* (2002) adalah sebagai berikut:

$H_0 : \beta_k(u_i, v_i) = \beta_k ; k = 1, 2, \dots, p$ (tidak ada perbedaan yang signifikan antara model OLS dan GWR)

$H_1 : \text{paling sedikit ada satu } \beta_k(u_i, v_i) \text{ yang berhubungan dengan lokasi } (u_i, v_i) \text{ (ada perbedaan yang signifikan antara model OLS dan GWR)}$

Statistik uji:

$$F^* = \frac{SSE(H_1)/(\delta_1^2/\delta_2)}{SSE(H_0)/(n-p-1)}$$

dengan $\delta_1 = \text{tr}((\mathbf{I}-\mathbf{L})^T(\mathbf{I}-\mathbf{L}))$ dan $\delta_2 = \text{tr}((\mathbf{I}-\mathbf{L})^T(\mathbf{I}-\mathbf{L}))^2$

$$SSE(H_1) = \hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} = \mathbf{y}^T (\mathbf{I}-\mathbf{L})^T (\mathbf{I}-\mathbf{L}) \mathbf{y} \text{ dan } SSE(H_0) = \hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} = \mathbf{y}^T (\mathbf{I}-\mathbf{H}) \mathbf{y}$$

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

Kriteria pengujian tolak H_0 jika nilai $F^* > F_{\alpha; (\delta_1^2/\delta_2), n-p-1}$.

2.4. Pengujian Parameter Model GWR

Jika pada pengujian kesesuaian model GWR diperoleh keputusan tolak H_0 maka dilakukan pengujian parameter secara parsial. Pengujian ini untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Bentuk hipotesisnya adalah sebagai berikut:

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

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

Statistik uji yang digunakan adalah sebagai berikut:

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

dengan c_{kk} adalah elemen diagonal ke- k dari matrik $\mathbf{C}_i \mathbf{C}_i^T$.

$$\mathbf{C}_i = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i)$$

Kriteria ujinya adalah tolak H_0 jika nilai $|T_{hit}| > t_{\alpha/2; df}$. dimana

$$df = \left(\frac{\delta_1^2}{\delta_2} \right) \cdot \delta_i = \text{tr} \left(\left[(\mathbf{I}-\mathbf{L})^T (\mathbf{I}-\mathbf{L}) \right]^i \right), \quad i = 1, 2$$

2.5. Pemilihan Pembobot (*Weight*)

Fungsi pembobot adalah untuk memberikan hasil penaksiran parameter yang berbeda pada lokasi yang berbeda. Pada analisis spasial, penaksiran parameter di suatu titik (u_i, v_i) akan lebih dipengaruhi oleh titik-titik yang dekat dengan lokasi (u_i, v_i) daripada titik-titik yang lebih jauh. Pemilihan pembobot spasial yang digunakan dalam menaksir parameter sangat penting. Salah satu pembobot yang digunakan dalam GWR adalah menggunakan fungsi kernel. Pembobot yang terbentuk dengan menggunakan fungsi kernel ini adalah fungsi kernel gaussian, fungsi kernel bisquare, adaptif bisquare kernel, dan fungsi adaptif gaussian kernel. Fungsi kernel gaussian dapat dituliskan sebagai berikut (Leung *et.al.*, 2000):

$$w_j(u_i, v_i) = \exp\left(-\left(d_{ij}/h\right)^2\right) \quad (6)$$

dimana h adalah bandwidth

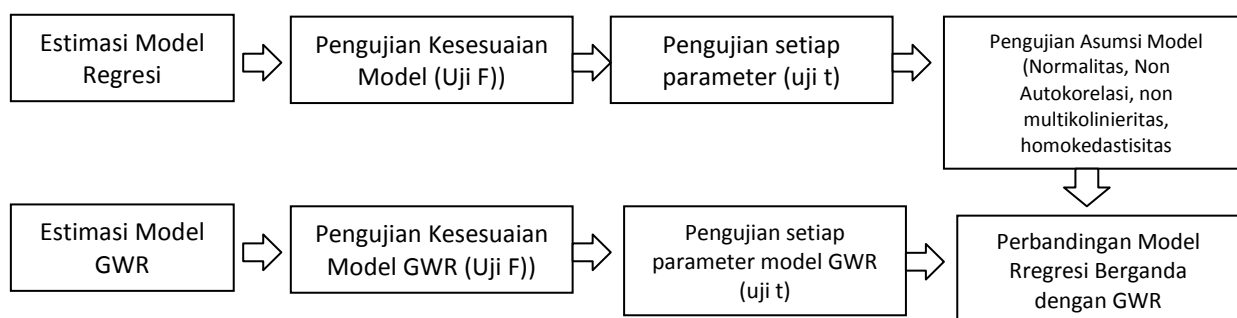
Pemilihan *bandwidth* optimum menjadi sangat penting karena akan mempengaruhi ketepatan model terhadap data, yaitu mengatur varians dan bias dari model. Salah satu metode yang digunakan untuk menentukan *bandwidth* optimum adalah metode *Cross Validation* (CV) (Brunsdon *et al.*, 2002) dan secara matematis dapat dituliskan sebagai berikut:

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(h))^2 \quad (7)$$

dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir y_i dimana pengamatan di lokasi i dihilangkan dari proses penaksiran.

3. METODE PENELITIAN

Pada penelitian ini akan dikaji mengenai faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di setiap Kabupaten/Kota Jawa Tengah tahun 2014. Peneliti mengambil beberapa faktor antara lain Kepadatan penduduk (X_1), Tingkat Inflasi (X_2), Nilai PDRB (X_3), Nilai UMR (X_4), prosentase Laju pertumbuhan PDRB (X_5), Harapan lama sekolah (X_6), prosentase angkatan kerja menurut usia (X_7) dan prosentase kesempatan kerja (X_8). Data tersebut diambil dari Badan Pusat Statistika Provinsi Jawa Tengah. Variabel-variabel prediktor yang digunakan mengacu pada penelitian Mustika dan Agustina (2010). Diasumsikan bahwa semakin tinggi kepadatan penduduk akan meningkatkan TPT. Semakin tinggi nilai inflasi akan semakin tinggi TPT, hal ini dikarenakan nilai investasi pada perusahaan tersebut akan berkurang, sehingga sedikit perusahaan yang membuka lapangan pekerjaan. Nilai PDRB menunjukkan kemajuan perekonomian di wilayah tersebut, semakin tinggi nilai PDRB semakin sejahtera kehidupan di wilayah tersebut, sehingga TPT menurun. Diagram alir alur penelitian ini disajikan pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

4. HASIL DAN PEMBAHASAN

4.1. Pemodelan Tingkat pengangguran dengan Regresi OLS

Setelah dilakukan pengolahan pada data, diperoleh fungsi regresi OLS sebagai berikut:

$$\hat{Y} = 0.0181X_1 + 0.0193X_2 + 0.126X_3 - 0.0025X_4 + 0.0094X_5 + 0.0757X_6 - 0.0665X_7 - 0.7652X_8.$$

Berdasarkan model yang diperoleh, terlihat bahwa kepadatan penduduk, tingkat inflasi, nilai laju pertumbuhan PDRB serta harapan lama sekolah mempunyai hubungan positif terhadap tingkat pengangguran terbuka. Akan tetapi untuk variabel nilai UMR, prosentase angkatan kerja menurut usia serta prosentase kesempatan kerja mempunyai hubungan yang negatif. Misalnya semakin tinggi nilai UMR disuatu wilayah, akan menunjukkan wilayah tersebut sejahtera sehingga tingkat pengangguran terbuka semakin menurun. Model tersebut hanya menghasilkan nilai *R-square* sebesar 57% dengan nilai MSE sebesar 0.5623.

Pengujian secara serentak model regresi OLS

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_6 = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0, k = 1, 2, \dots, 6$$

Statistik uji yang digunakan adalah uji F. H_0 ditolak apabila F hitung $>$ F tabel. Diperoleh nilai F hitung = 4.31 dengan *p-value* = 0.000 sehingga jika $\alpha = 5\%$ maka keputusan yang diambil adalah H_0 ditolak. Dengan kata lain pada tingkat kepercayaan 95% variabel prediktor yang digunakan berpengaruh signifikan secara serentak terhadap variabel respon.

Pengujian secara parsial untuk model regresi OLS:

Hipotesis

$$H_0 : \beta_k = 0 \text{ (variabel prediktor ke-}k\text{ tidak berpengaruh signifikan terhadap respon)}$$

$$H_0 : \beta_k \neq 0 \text{ (variabel prediktor ke-}k\text{ berpengaruh signifikan terhadap respon)}$$

Statistik uji yang digunakan adalah uji t. H_0 ditolak apabila t-hitung $>$ t-tabel atau bila *p-value* $<$ α . Hasil analisis uji ini disajikan pada Tabel 2.

Tabel 2. Hasil Analisis Uji Parsial Model Regresi OLS

Prediktor	Koefisien	SE Koef.	T-hitung	P-value	Keputusan	Kesimpulan
Konstanta	0.0000	0.1268	0.00	1.000	Terima H ₀	Tidak Signifikan
X ₁	0.0181	0.1490	0.12	0.904	Terima H ₀	Tidak Signifikan
X ₂	0.0193	0.1415	0.14	0.892	Terima H ₀	Tidak Signifikan
X ₃	0.1260	0.1453	0.87	0.394	Terima H ₀	Tidak Signifikan
X ₄	-0.0025	0.1909	-0.01	0.990	Terima H ₀	Tidak Signifikan
X ₅	0.0094	0.1680	0.56	0.581	Terima H ₀	Tidak Signifikan
X ₆	0.0757	0.1739	0.44	0.667	Terima H ₀	Tidak Signifikan
X ₇	-0.0665	0.1497	-0.44	0.661	Terima H ₀	Tidak Signifikan
X ₈	-0.7652	0.1520	-5.02	0.000	Tolak H ₀	Signifikan

Dari hasil tersebut terlihat bahwa hanya satu variabel prediktor yang berpengaruh signifikan terhadap respon Y (tingkat pengangguran terbuka di Jawa Tengah) yaitu X₈ (Prosentase kesempatan kerja).

4.2. Pengujian Asumsi model Regresi OLS

Residual pada OLS diasumsikan mengikuti sebaran normal, variansi residual homogen, dan independen. Pengujian sebaran normal residual OLS dilakukan dengan uji Anderson Darling. Hipotesis yang digunakan adalah:

H₀ : Residual berdistribusi normal

H₁ : Residual tidak berdistribusi normal

Statistik uji yang digunakan adalah Kolmogorov Smirnov (KS). H₀ ditolak apabila $KS > KS$ tabel. Diperoleh nilai KS sebesar 0.112, kurang dari KS tabel (3.24). Jadi keputusan yang diambil adalah gagal menolak H₀ dan kesimpulannya adalah residual OLS berdistribusi normal. Uji homogenitas variansi residual digunakan hipotesis sebagai berikut

H₀ : $\sigma_i^2 = \sigma^2$, $i = 1, 2, \dots, 3$

H₁ : minimal ada satu $\sigma_i^2 \neq \sigma^2$

Uji statistik yang digunakan adalah uji Glejser. Uji ini diperoleh dengan meregresikan nilai absolut residual dari model awal dengan semua variabel prediktor yang digunakan. Jika terdapat variabel prediktor yang signifikan mempengaruhi nilai absolut residual maka variansi residual cenderung tidak homogen, sebaliknya jika tidak signifikan menunjukkan bahwa asumsi homogenitas terpenuhi. Hasil regresi absolut residual terhadap semua variabel prediktor disajikan pada Tabel 3. Dari tabel tersebut terlihat bahwa semua variabel prediktor tidak signifikan pada tingkat kepercayaan 5%. Hal ini mengindikasikan bahwa variansi dari residual OLS memenuhi asumsi homokedastisitas.

Kemudian untuk menguji multikolinieritas digunakan nilai VIF (*Variance in Factor*). Jika nilai VIF > 10 , maka terdapat multikolinieritas antar variabel prediktor. Diperoleh nilai VIF untuk masing-masing variabel prediktor adalah X₁ = 1.343; X₂ = 1.21; X₃ = 1.277; X₄ = 2.204; X₅ = 1.706; X₆ = 1.828; X₇ = 1.354 dan X₈ = 1.394. Hal ini menunjukkan bahwa semua variabel prediktor tidak terjadi multikolinieritas.

Tabel 3. Hasil Analisis Uji Regresi Absolut Residual dengan Variabel Prediktor

Prediktor	P-value	Keputusan	Kesimpulan
X ₁	0.86	Terima H ₀	Tidak Signifikan
X ₂	0.14	Terima H ₀	Tidak Signifikan
X ₃	0.25	Terima H ₀	Tidak Signifikan
X ₄	0.52	Terima H ₀	Tidak Signifikan
X ₅	0.30	Terima H ₀	Tidak Signifikan
X ₆	0.36	Terima H ₀	Tidak Signifikan
X ₇	0.24	Terima H ₀	Tidak Signifikan
X ₈	0.52	Terima H ₀	Tidak Signifikan

Selain itu asumsi yang harus dipenuhi dalam model regresi linier berganda adalah tidak terjadi autokorelasi residual. Pengujian autokorelasi dapat menggunakan nilai Durbin-Watson (d) yang dibandingkan dengan nilai tabel Durbin-Watson (d_L dan d_U). Jika nilai $d < d_L$ maka terdapat autokorelasi positif, jika $d < d_U$ maka tidak terdapat autokorelasi dan jika $d_L \leq d \leq d_U$ maka hasil pengujian tidak dapat ditentukan (Gujarati, 2004). Dari hasil pengolahan diperoleh nilai Durbin-Watson sebesar 1.273 dan nilai tabel Durbin-Watson ($d_L = 0.9$ dan $d_U = 1.847$). Jadi nilai Durbin-Watson terletak diantara tabel Durbin-Watson, sehingga tidak dapat diputuskan hasil pengujian autokorelasi.

Pada pengujian asumsi residual model regresi didapatkan bahwa residual berdistribusi normal, homokedastisitas, tidak terjadi multikolinieritas antar variabel prediktor dan residual model tidak dapat diputuskan asumsi residualnya. Walaupun asumsi-asumsi tersebut terpenuhi, akan tetapi hanya satu variabel prediktor saja yang signifikan terhadap model.

4.3. Pemodelan Tingkat Pengangguran dengan GWR

Estimasi parameter model GWR diperoleh dengan memasukkan pembobot spasial dalam perhitungannya dengan menggunakan metode *Weighted Least Squares* (WLS). Hasil estimasi parameter model GWR disajikan pada Tabel 4.

Tabel 4. Hasil Estimasi Parameter Model GWR

Parameter	Rata-Rata	Minimum	Maximum
Konstanta	-0.0202813	-0.162441	0.122474
X ₁	0.0567809	-0.005446	0.109282
X ₂	0.0448375	-0.085636	0.168866
X ₃	0.1006696	0.041824	0.146994
X ₄	0.0289949	0.011314	0.062641
X ₅	0.0723071	0.040947	0.109724
X ₆	0.1158609	0.021501	0.193926
X ₇	-0.0568415	-0.128183	0.009951
X ₈	-0.7807653	-0.795913	-0.735791

Estimasi parameter yang didapatkan tidak berlaku secara global. Hal ini dikarenakan terdapat pembobot untuk tiap daerah sehingga tiap kabupaten/kota memiliki estimasi parameter yang berbeda-beda. Nilai minimum estimasi parameter X₁ yang

dihasilkan model GWR sebesar -0.005446 dan nilai maksimumnya 0.109282. Nilai ini menunjukkan besarnya pengaruh X_1 (kepadatan penduduk di tiap kabupaten/kota) terhadap tingkat pengangguran terbuka di Jawa Tengah berkisar antara -0.005446 sampai dengan 0.109282. Nilai minimum dan maksimum pada Tabel 4 menunjukkan nilai minimum dan maksimum yang didapatkan untuk parameter $\hat{\beta}$ pada tiap kabupaten/kota. Sementara hasil estimasi parameter model GWR untuk tiap kecamatan secara lengkap ditunjukkan pada Lampiran 1.

Terdapat beberapa parameter yang nilainya berkisar antara negatif sampai positif yaitu parameter β_1 , β_2 , dan β_7 . Hal ini mengindikasikan bahwa variabel prediktor X_1 (Kepadatan penduduk di tiap kabupaten/kota), X_2 (Besarnya inflasi tiap tahun di tiap kabupaten/kota), dan X_7 (Angkatan kerja menurut usia di tiap kabupaten/kota) memiliki koefisien regresi negatif pada beberapa kabupaten/kota (bertambahnya satu satuan dari variabel prediktor akan mengurangi jumlah tingkat pengangguran terbuka di Jawa Tengah dengan asumsi variabel lainnya konstan) dan pada kabupaten/kota lainnya memiliki koefisien regresi yang positif (bertambahnya satu satuan dari variabel prediktor akan meningkatkan jumlah tingkat pengangguran terbuka di Jawa Tengah dengan asumsi variabel lainnya konstan). Model tersebut menghasilkan nilai *R-square* 64.83% dengan nilai MSE sebesar 0.5131.

4.4. Pengujian Kesesuaian Model GWR dengan OLS

Setelah diperoleh estimasi parameter untuk model regresi OLS dan GWR kemudian dilakukan uji kesesuaian model (*goodness of fit*) untuk melihat apakah faktor geografis berpengaruh terhadap tingkat pengangguran terbuka di Jawa tengah. Bentuk hipotesisnya adalah sebagai berikut.

$$H_0 : \beta_k(u_1, v_1) = \beta_k(u_2, v_2) = \dots = \beta_k(u_{38}, v_{38}) = \beta_k, k = 1, 2, \dots, 6$$

(Tidak terdapat perbedaan yang signifikan antara model GWR dengan model OLS)

$$H_1 : \text{paling sedikit ada satu } \beta_k(u_i, v_i) \neq \beta_k$$

(Terdapat perbedaan yang signifikan antara model GWR dengan model OLS, Model GWR lebih baik daripada model OLS)

Pengujian ini dilakukan dengan menggunakan uji Leung dan statistik uji F. Hasil pengujian parameter dengan menggunakan software GWR 4 disajikan pada Tabel 5.

Tabel 5. Uji Kesesuaian Model dengan *Software R*

Model	SS	df	MS	F
Global	14.620	9	-	-
GWR Improvement	4.012	3.626	1.107	-
GWR Residual	10.608	22.374	0.474	2.334

Terlihat dari model tersebut diperoleh F-hitung sebesar 2.334. dengan nilai F-tabel ($F_{0.05;3.626;22.374}$) sebesar 3.049. F tabel yang dihasilkan lebih besar sehingga keputusan yang diambil adalah gagal menolak H_0 dan kesimpulannya tidak terdapat perbedaan yang signifikan antara model GWR dengan model OLS. Dalam kasus ini, menunjukkan bahwa

estimasi secara global (regresi linier berganda) dengan estimasi model GWR menghasilkan nilai yang sama.

Mengacu pada Lampiran 1 tentang nilai estimasi parameter model GWR di setiap lokasi, Tabel 6 menunjukkan nilai *p-value* variabel prediktor model GWR. Berdasarkan tabel tersebut terlihat bahwa hanya variabel X_8 yang signifikan di setiap lokasi.

Tabel 6. Nilai *p-value* Setiap Variabel Prediktor Model GWR

Lokasi	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	0.453	0.444	0.171	0.462	0.311	0.382	0.256	0.00
2	0.448	0.448	0.176	0.466	0.306	0.370	0.263	0.00
3	0.461	0.437	0.173	0.473	0.297	0.379	0.260	0.00
4	0.396	0.482	0.197	0.449	0.325	0.318	0.282	0.00
5	0.360	0.468	0.200	0.431	0.358	0.297	0.292	0.00
6	0.304	0.378	0.223	0.419	0.380	0.243	0.331	0.00
7	0.343	0.447	0.221	0.435	0.344	0.264	0.313	0.00
8	0.302	0.328	0.245	0.433	0.356	0.216	0.367	0.00
9	0.314	0.290	0.249	0.449	0.345	0.211	0.393	0.00
10	0.311	0.285	0.248	0.449	0.348	0.211	0.394	0.00
11	0.293	0.291	0.248	0.435	0.363	0.207	0.384	0.00
12	0.302	0.252	0.255	0.449	0.356	0.201	0.412	0.00
13	0.321	0.274	0.251	0.456	0.341	0.210	0.405	0.00
14	0.332	0.280	0.255	0.461	0.329	0.212	0.408	0.00
15	0.317	0.328	0.259	0.443	0.332	0.211	0.380	0.00
16	0.378	0.299	0.260	0.481	0.295	0.226	0.417	0.00
17	0.348	0.282	0.256	0.470	0.318	0.217	0.413	0.00
18	0.371	0.305	0.261	0.476	0.297	0.223	0.412	0.00
19	0.355	0.275	0.259	0.474	0.316	0.217	0.421	0.00
20	0.376	0.318	0.271	0.472	0.285	0.221	0.411	0.00
21	0.344	0.305	0.271	0.458	0.307	0.209	0.407	0.00
22	0.320	0.311	0.257	0.448	0.331	0.211	0.389	0.00
23	0.321	0.297	0.253	0.451	0.335	0.211	0.394	0.00
24	0.374	0.480	0.242	0.456	0.295	0.259	0.319	0.00
25	0.389	0.490	0.224	0.458	0.299	0.284	0.302	0.00
26	0.423	0.450	0.201	0.466	0.295	0.325	0.279	0.00
27	0.360	0.458	0.185	0.417	0.385	0.318	0.281	0.00
28	0.462	0.428	0.179	0.478	0.286	0.370	0.263	0.00
29	0.487	0.398	0.176	0.489	0.269	0.388	0.255	0.00
30	0.304	0.337	0.246	0.433	0.353	0.217	0.364	0.00
31	0.307	0.270	0.250	0.448	0.353	0.207	0.401	0.00
32	0.320	0.305	0.254	0.449	0.334	0.211	0.390	0.00
33	0.331	0.330	0.272	0.448	0.311	0.208	0.389	0.00
34	0.433	0.435	0.207	0.474	0.280	0.323	0.280	0.00
35	0.474	0.414	0.179	0.485	0.276	0.375	0.261	0.00

Jika dibandingkan antara model regresi berganda dengan model GWR, secara statistik uji tidak ada perbedaan. Akan tetapi pada model GWR bisa dilihat nilai parameter di setiap lokasi berbeda-beda, sedangkan pada model regresi berganda tidak membedakan lokasi.

5. KESIMPULAN

Pada pemodelan regresi berganda dan GWR hanya ada satu variabel prediktor yang signifikan yaitu prosentase kesempatan kerja (X_8). Pemodelan regresi berganda memenuhi asumsi normalitas, homokedastisitas, dan non-multikolinieritas, sedangkan non-autokorelasi tidak dapat ditentukan. Secara ukuran kebaikan model, GWR menghasilkan nilai yang R^2 lebih besar dibandingkan regresi berganda dan GWR menghasilkan MSE lebih kecil dibandingkan regresi berganda. Akan tetapi secara statistik uji perbandingan model GWR dan regresi berganda tidak signifikan. Artinya model GWR dan regresi berganda menghasilkan nilai yang sama. Akan tetapi secara pemodelan, model GWR menghasilkan model tiap wilayah sementara model regresi berganda tidak menunjukkan model disetiap wilayah.

DAFTAR PUSTAKA

- Anselin, L. dan Bera, A. 1998. *Spatial Dependence In Linear Regression Models With An Introduction To Spatial Econometrics*. dalam Ullah, A. & Giles, D. *Handbook of Applied Economics Statistic*. Selected Reading, pp. 237-289. Marcel Dekker. New York.
- Aulele, S. N., 2010, *Model Geographically Weighted Poisson Regression (Studi Kasus : Jumlah Kematian Bayi di Provinsi Jawa Timur dan Jawa Tengah Tahun 2007)*, Tesis, Jurusan Statistika FMIPA ITS, Surabaya, Tidak dipublikasikan.
- Brunsdon C., Fotheringham, A. S., dan Charlton, M. E., 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*. John Wiley and Sons Ltd., England.
- Chiou, Y.C., Jou, R.C., dan Yang, C.H., 2015, Factor Affecting Public Transportation Usage Rate: Geographically Weighted Regression, *Transportation Reseach A*, pp.161-177.
- Draper, N. dan Smith, H., 1992, *Analisis Regresi Terapan*. PT. Gramedia Pustaka Utama. Jakarta.
- Gujarati, D.N. 2004. *Basic Econometrics 4th Edition*. The McGraw Hill.
- Hajji, M.S dan Nugroho, SBM., 2013, Analisis PDRB, Inflasi, Upah Minimum Provinsi dan Angka Melek Huruf terhadap Tingkat Pengangguran Terbuka di Provinsi Jawa Tengah tahun 1990-2011, *Diponegoro Journal of Economics*, Vol 2 No 3, UNDIP, Semarang.
- Intan, 2010, *Pemodelan Jumlah Penderita Tuberculosis (TB) di Propinsi Jawa Timur Tahun 2010 dengan Menggunakan Metode Geographically Weighted Regression*, Tugas Akhir Jurusan Statistika FMIPA ITS, Surabaya, Tidak dipublikasikan.
- Leung Y, Mei CL, dan Zhang WX., 2000, *Statistic Tests for Spatial Non-Stationarity Based on the Geographically Weighted Regression Model*, *Environment and Planning A*, 32, pp. 9-32.
- Mustika C.D., dan Agustina, 2010, *Analisis Tingkat Pengaganguran dan Faktor-faktor yang Mempengaruhinya di Kota Semarang*. Skripsi. Fakultas Ekonomi, UNDIP, Semarang, Tidak dipublikasikan.

- Saefuddin A, Nur Andi S., dan Noer Azam A. 2011. On Comparisson between Ordinary Linear Regression and Geographically Weighted Regression: With Application to Indonesian Poverty Data. *European Journal of Scientific Research. Euro Journals Publishing, Inc.*
- Sari, R.S dan Budiantara, I.N., 2012, Pemodelan Pengangguran Terbuka di Jawa Timur dengan Menggunakan Pendekatan Regresi Spline Multivariabel, *Jurnal Sains dan Seni ITS*, Vol 1, No 1, ITS, Surabaya.
- Supranto, J., 2001, *Statistik Teori dan Aplikasi*, Cetakan Kedua, Erlangga, Jakarta.

