

# Perbandingan Performa *Bandwidth CV*, *AICc*, dan *BIC* pada Model *Geographically Weighted Regression* (Aplikasi pada Data Pengangguran di Pulau Jawa)

Carisa Putri Salsabila Purnamasari<sup>1\*</sup> dan Yekti Widyaningsih<sup>2</sup>

<sup>1,2</sup>Matematika, Universitas Indonesia, Depok, Indonesia

\*Corresponding author: carisa.putri@sci.ui.ac.id

Received: 24 January 2023

Revised: 4 August 2023

Accepted: 4 October 2023

**ABSTRAK** – Pengangguran merupakan fenomena sosial yang menjadi salah satu masalah utama yang dihadapi setiap daerah di Indonesia. Salah satu cara yang dapat dilakukan untuk mengurangi angka pengangguran adalah dengan melakukan analisis terhadap faktor-faktor yang mempengaruhi tingkat pengangguran terbuka. Dibandingkan dengan metode analisis regresi linier, metode *Geographically Weighted Regression* (GWR) lebih diunggulkan karena dapat menangani masalah heterogenitas spasial yang biasanya terjadi pada data fenomena sosial. Heterogenitas spasial mengakibatkan hasil analisis regresi linier menjadi tidak akurat di beberapa lokasi. GWR menangani masalah tersebut dengan membangun model regresi di setiap lokasi pengamatan sehingga memungkinkan parameter regresi menjadi berbeda di setiap lokasi pengamatan. Pendugaan parameter pada model GWR menggunakan pembobot berdasarkan lokasi setiap pengamatan. Penentuan pembobot bergantung pada nilai *bandwidth*. Nilai *bandwidth* yang sangat kecil akan mengakibatkan variansi yang besar. Hal tersebut disebabkan karena jika nilai *bandwidth* sangat kecil maka jumlah pengamatan yang berada pada radius  $h$  menjadi sedikit, sehingga menyebabkan model yang diperoleh sangat kasar (*undersmoothing*) karena menggunakan sedikit pengamatan, dan sebaliknya. Oleh karena itu, pemilihan *bandwidth* optimum sangat penting dalam menentukan pembobot karena dapat mempengaruhi ketepatan model yang terbentuk. Penelitian ini bertujuan untuk mengetahui perbandingan performa model GWR yang menggunakan metode *bandwidth Cross Validation* (CV), *Akaike Information Criterion Corrected* (AICc), dan metode *Bayesian Information Criterion* (BIC) dalam pembentukan fungsi pembobot *Fixed Gaussian Kernel* yang diterapkan pada data pengangguran di kabupaten/kota di Pulau Jawa. Hasil penelitian menunjukkan bahwa model GWR *bandwidth CV* lebih baik dalam menjelaskan data pengangguran kabupaten/kota di Pulau Jawa tahun 2020 karena memiliki nilai RMSE paling kecil, yaitu 1,0904 serta nilai  $R^2$  dan *Adjusted-R<sup>2</sup>* paling besar, yaitu 0,8539011 dan 0,7937159.

**Kata kunci**– *Bandwidth*, CV, AICc, BIC, GWR

**ABSTRACT** – Unemployment is a social phenomenon, a problem faced by every region in Indonesia. One way that can be carried out to reduce the unemployment rate is analyzing the factors that affect the open unemployment rate. Rather than using linear regression analysis, Geographically Weighted Regression (GWR) was preferable since it gave a better representative model by effectively resolve spatial heterogeneity problem which is generally exist in spatial data of social phenomenon. Spatial heterogeneity show that linear regression analysis will give a misleading interpretation results in some locations. GWR solve this problem by generating a single model in each observation location so the regression parameters can be different at each observation location. Parameter estimation in the GWR model uses weights based on the location of each observation so that the estimate model applies only to this location. The weighting determination depends on the bandwidth value. Bandwidth is a circle with radius  $h$  from the center point of the observation location which is used as the basis for determining the weight of each observation location. Smaller bandwidth value will result a large variance. It can happen because when the bandwidth is very small, there will be a small number observations in the radius  $h$ , which can makes the estimate model is very rough (*undersmoothing*) because it uses few observations, and vice versa. Therefore, choosing the optimum bandwidth is very important in determining the weights where it can affect the accuracy of the model formed. This study aims to compare the performance of the GWR model using the *Cross Validation* (CV), *Akaike Information Criterion Corrected* (AICc), and *Bayesian Information Criterion* (BIC) bandwidth methods in the formation of *Fixed Gaussian Kernel* weighted function which is applied to unemployment data in districts/cities in Java. The results show that the GWR model with CV bandwidth is better at explaining district/city unemployment data on Java Island in 2020 which it has the smallest RMSE value, 1.0904, and the largest  $R^2$  and *Adjusted-R<sup>2</sup>* values, namely 0.8539011 and 0.7937159, respectively.

**Keywords**– *Bandwidth*, CV, AICc, BIC, GWR

## I. PENDAHULUAN

Pengangguran merupakan fenomena sosial yang menjadi salah satu masalah utama yang dihadapi setiap daerah di Indonesia. Menurut Badan Pusat Statistik (BPS), tingkat pengangguran terbuka (TPT) merupakan indikator utama yang digunakan untuk mengukur angka pengangguran dalam angkatan kerja. Pada Agustus 2021, angka tingkat pengangguran terbuka untuk Indonesia adalah sebesar 6,49%. BPS melaporkan, ada delapan provinsi yang memiliki tingkat pengangguran terbuka di atas rata-rata nasional. Tiga provinsi diantaranya terletak di Pulau Jawa, yaitu Provinsi Jawa Barat dengan tingkat pengangguran terbuka tertinggi kedua sebesar 9,82%; kemudian Provinsi Banten dengan tingkat pengangguran terbuka tertinggi ketiga sebesar 8,98%; dan Provinsi DKI Jakarta dengan tingkat pengangguran terbuka tertinggi keempat sebesar 8,5% [1]. Tingkat pengangguran yang tinggi ini berpengaruh terhadap kesejahteraan masyarakat, diantaranya seperti kemiskinan [2]. Tingkat pengangguran yang tinggi ini juga dapat mengganggu tatanan kehidupan sosial karena dapat menimbulkan berbagai kejahatan, seperti pencurian, anak jalanan, pelacuran, dan lain-lain

[3]. Analisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka merupakan salah satu cara yang dapat dilakukan untuk mengurangi angka pengangguran. Berdasarkan studi literatur, faktor-faktor yang diduga berpengaruh terhadap tingkat pengangguran terbuka adalah kepadatan penduduk, indeks pembangunan manusia, tingkat partisipasi angkatan kerja, upah minimum kabupaten/kota, rata-rata upah sebulan pekerja formal, dan rata-rata pendapatan bersih sebulan pekerja informal, sehingga dalam penelitian ini faktor-faktor tersebut digunakan sebagai variabel independen.

Analisis regresi linier merupakan salah satu metode statistik yang dapat digunakan untuk menganalisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka. Analisis regresi linier adalah metode statistik yang digunakan untuk mengetahui hubungan antara satu variabel dependen dan satu atau lebih variabel independen. Analisis regresi linier adalah pemodelan yang tidak melibatkan unsur spasial walaupun data yang dianalisis adalah data yang mengandung unsur spasial. Pada data yang mengandung unsur spasial, disebut juga data spasial (data yang memuat informasi atribut dan lokasi), analisis regresi akan memberikan pendugaan parameter yang baik jika hubungan antara variabel dependen dan variabel independen bersifat stasioner, yang artinya hubungan antar variabel bersifat tetap di masing-masing lokasi pengamatan. Menurut Fotheringham *et. al.* (2002), hubungan antar variabel pada data fenomena sosial akan cenderung bersifat non-stasioner, atau dapat dikatakan hubungan antar variabel cenderung berbeda pada setiap lokasi [4]. Kondisi hubungan antar variabel yang berbeda-beda di setiap lokasi sering disebut juga dengan heterogen spasial. Faktor-faktor yang menyebabkan terjadinya heterogenitas spasial, antara lain perbedaan kondisi geografis, sosial, budaya, ekonomi, maupun politik yang berlaku pada setiap lokasi. Jika heterogenitas spasial terjadi, maka pemodelan hubungan antara variabel dependen dan variabel independen dilakukan dengan metode *Geographically Weighted Regression*.

*Geographically Weighted Regression* (GWR) adalah metode statistik yang memodelkan hubungan antara variabel dependen dan variabel independen yang mengandung heterogenitas spasial. GWR merupakan pengembangan dari analisis regresi linier yang bersifat lokal untuk setiap lokasi pengamatan. Pendugaan parameter pada model GWR membutuhkan matriks pembobot, yaitu pemberian bobot pada data sesuai dengan kedekatan lokasi pengamatan ke-*i*. Pemberian bobot ini sesuai dengan teori Tobler (1970) yang dikenal dengan hukum geografi pertama bahwa "*Everything is related to everything else, but near things are more related than distant things.*" [5]. Teori tersebut mengandung makna suatu hal pasti memiliki kaitan dengan hal lain, tapi hal-hal yang jaraknya berdekatan lebih erat kaitannya dibandingkan hal-hal yang jaraknya berjauhan. Untuk membentuk matriks pembobot tersebut digunakan suatu fungsi pembobot. Penggunaan fungsi pembobot masih bersifat *trial and error* dan berdasarkan kebiasaan, sehingga pemilihan fungsi pembobot pada penelitian biasanya didasarkan pada penelitian-penelitian sebelumnya. Berdasarkan hasil penelitian yang dilakukan oleh Guo *et. al.* (2008) mengenai model GWR dengan *bandwidth* metode CV dan AIC pada pembobot *Fixed* dan *Adaptive Gaussian Kernel* serta *Fixed* dan *Adaptive Bisquare Kernel*, penggunaan pembobot *Fixed Kernel* menghasilkan model yang lebih baik dibandingkan dengan penggunaan pembobot *Adaptive Kernel* [6]. Rosa (2015) dalam penelitiannya mengenai model GWR dengan *bandwidth* metode CV pada pembobot *Fixed Gaussian Kernel* dan *Fixed Bisquare Kernel* menjelaskan bahwa penggunaan pembobot *Fixed Gaussian Kernel* menghasilkan model yang lebih baik dibandingkan dengan penggunaan pembobot *Fixed Bisquare Kernel* [7]. Sehingga dalam penelitian ini digunakan fungsi pembobot *Fixed Gaussian Kernel*. Fungsi pembobot tersebut tergantung pada nilai *bandwidth*. *Bandwidth* merupakan lingkaran dengan radius *h* dari titik pusat lokasi pengamatan yang digunakan sebagai dasar penentuan pembobot setiap lokasi pengamatan. Oleh karena itu, pemilihan *bandwidth* optimum sangat penting dalam menentukan fungsi pembobot karena dapat mempengaruhi ketepatan model yang terbentuk. Nilai *bandwidth* yang kecil menyebabkan nilai variansi menjadi semakin besar, sehingga jika *bandwidth* sangat kecil maka semakin sedikit pengamatan dalam radius *h* dan model yang diperoleh akan kasar (*undersmoothing*) karena menggunakan sedikit pengamatan. Sementara jika nilai *bandwidth* sangat besar, bias yang dihasilkan akan semakin besar dan model akan terlampaui halus (*oversmoothing*) karena menggunakan banyak pengamatan. Model yang *undersmoothing* akan menghasilkan parameter dengan banyak variasi lokal sehingga sulit untuk melihat pola yang terbentuk. Sedangkan, model yang *oversmoothing* akan menghasilkan parameter yang nilainya hampir sama di seluruh lokasi pengamatan. Sehingga untuk menghindari kedua hal tersebut diperlukan metode untuk menghasilkan nilai *bandwidth* optimum [8].

Menurut Fotheringham *et. al.* (2002), terdapat beberapa metode dalam pemilihan *bandwidth* optimum, diantaranya metode *Cross Validation* (CV), metode *Akaike Information Criterion Corrected* (AICc), dan metode *Bayesian Information Criterion* (BIC) [4]. Metode CV melibatkan konsep dua kelompok data, yaitu data *training* dan data *testing*. Kelompok data *testing* diambil dengan cara menghilangkan nilai titik data ke-*i* untuk dibandingkan dengan kelompok data *training* (data riil atau data keseluruhan) yang dilakukan berulang sampai diperoleh nilai CV minimum. Sedangkan, pemilihan *bandwidth* optimum dengan metode AICc ditentukan menggunakan iterasi sedemikian sehingga diperoleh nilai AICc minimum. Adapun metode BIC didasarkan pada metode *maximum likelihood estimation* dan teorema Bayesian. Metode BIC dilakukan secara iterasi hingga nilai BIC yang dihasilkan minimum.

Berdasarkan hasil penelitian yang dilakukan oleh Ariyanto (2017) mengenai perbandingan *bandwidth* dengan metode CV dan AIC dalam pembentukan fungsi pembobot *Adaptive Gaussian Kernel*, menghasilkan kesimpulan bahwa model GWR dengan *bandwidth* CV memberikan model yang lebih baik dibandingkan model GWR dengan *bandwidth* AIC [9]. Amin (2021) dalam penelitiannya melakukan perbandingan *bandwidth* dengan metode CV dan AICc dalam pembentukan fungsi pembobot *Fixed Gaussian Kernel* dan menghasilkan kesimpulan bahwa model GWR dengan *bandwidth* CV memberikan model yang lebih baik dibandingkan model GWR dengan *bandwidth* AICc [10]. Penelitian – penelitian sebelumnya menggunakan metode CV, AIC ataupun AICc dalam menentukan nilai *bandwidth*, sehingga dalam penelitian ini penulis tertarik untuk melakukan pemodelan GWR dengan penentuan *bandwidth* menggunakan metode

CV, AICc, dan BIC dalam pembentukan fungsi pembobot *Fixed Gaussian Kernel* pada data pengangguran di Pulau Jawa, kemudian dibandingkan performa dari ketiga model tersebut. Penelitian ini bertujuan untuk mengetahui perbandingan performa model GWR yang menggunakan metode *bandwidth* CV, AICc, dan BIC yang diterapkan pada data pengangguran di kabupaten/kota di Pulau Jawa. Terdapat beberapa batasan masalah dalam penelitian ini, yaitu fungsi pembobot yang digunakan adalah fungsi pembobot *Fixed Gaussian Kernel*, penilaian performa model yang dihasilkan menggunakan nilai RMSE, nilai koefisien determinasi ( $R^2$ ), dan nilai koefisien determinasi disesuaikan (*Adjusted-R<sup>2</sup>*), dan data yang digunakan merupakan data pengangguran kabupaten/kota di Pulau Jawa tahun 2020. Adapun manfaat yang dapat diperoleh dari penelitian ini, yaitu hasil penelitian ini diharapkan dapat memberikan gambaran tentang model GWR yang menggunakan *bandwidth* CV, AICc, dan BIC dalam pembentukan fungsi pembobot *Fixed Gaussian Kernel*. Selain itu, dapat menghasilkan pemetaan angka pengangguran di Pulau Jawa yang diharapkan dapat memberikan informasi bagi pemerintah dan pihak terkait dalam membuat kebijakan guna mengatasi permasalahan pengangguran di Pulau Jawa, serta dapat menambah kepustakaan bagi pengguna ilmu statistika.

## II. TINJAUAN PUSTAKA

### A. Heterogenitas Spasial

Heterogenitas spasial atau keragaman spasial menunjukkan adanya keragaman antar lokasi pengamatan. Adanya heterogenitas spasial pada data dapat diuji dengan menggunakan statistik uji *Breusch-Pagan* (BP) dengan hipotesis sebagai berikut [11]:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \text{ (Tidak terdapat heterogenitas spasial)}$$

$$H_1: \text{Minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n \text{ (Terdapat heterogenitas spasial)}$$

Statistik Uji *Breusch-Pagan*:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f}$$

Keterangan:

$\mathbf{f}$  : vektor berukuran  $(n \times 1)$ , yaitu  $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$  dengan  $f_i = \frac{e_i^2}{\sigma^2} - 1$ .

$\mathbf{Z}$  : matriks berukuran  $n \times (p + 1)$  yang berisi vektor yang sudah dinormalkan untuk setiap lokasi pengamatan, atau dengan kata lain, matriks  $\mathbf{X}$  yang telah distandardisasi.

$e_i$  : residual dari model regresi global untuk pengamatan ke- $i$ .

$\sigma^2$  : variansi residual.

$n$  : banyaknya lokasi pengamatan.

$p$  : banyaknya variabel independen.

Aturan keputusannya adalah  $H_0$  ditolak jika  $BP > \chi^2_{(\alpha, p)}$  atau jika  $p - \text{value} < \alpha$ .

### B. Geographically Weighted Regression

Model *Geographically Weighted Regression* atau GWR merupakan pengembangan dari model regresi linier dengan mempertimbangkan faktor lokasi setiap pengamatan dalam melakukan pendugaan parameter, sehingga setiap lokasi pengamatan memiliki parameter regresi yang berbeda-beda [12]. Secara matematis, model GWR dapat dituliskan sebagai berikut [4]:

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

Keterangan:

$y_i$  : nilai variabel dependen pada lokasi pengamatan ke- $i$ .

$x_{ik}$  : nilai variabel independen ke- $k$  pada lokasi pengamatan ke- $i$ .

$(u_i, v_i)$  : titik koordinat lokasi pengamatan (*longitude, latitude*) ke- $i$ .

$\beta_0(u_i, v_i)$  : konstanta/intersep model GWR pada lokasi pengamatan ke- $i$ .

$\beta_k(u_i, v_i)$  : parameter model GWR ke- $k$  pada lokasi pengamatan ke- $i$ .

$\varepsilon_i$  : *error* pada lokasi pengamatan ke- $i$ .

$p$  : banyaknya variabel independen.

$n$  : banyaknya lokasi pengamatan.

Pendugaan parameter model GWR dilakukan menggunakan metode *Weighted Least Squares* (WLS) dengan pemberian bobot yang berbeda pada setiap lokasi pengamatan sesuai dengan Hukum Pertama Tobler. Rumus penduga matriks parameter  $\beta(u_i, v_i)$  menggunakan metode WLS adalah

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

dengan

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

dan  $W(u_i, v_i)$  adalah matriks pembobot spasial.

**C. Matriks Pembobot Spasial**

Matriks pembobot spasial  $W(u_i, v_i)$  merupakan matriks diagonal berukuran  $(n \times n)$  dengan elemen diagonal  $w_j(u_i, v_i)$  dengan  $j = 1, 2, \dots, n$ , yaitu bobot untuk lokasi pengamatan ke- $j$  pada saat pembentukan model GWR di lokasi  $(u_i, v_i)$ .

$$W(u_i, v_i) = \begin{bmatrix} w_1(u_i, v_i) & 0 & \dots & 0 & 0 \\ 0 & w_2(u_i, v_i) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & w_{n-1}(u_i, v_i) & 0 \\ 0 & 0 & \dots & 0 & w_n(u_i, v_i) \end{bmatrix}$$

**D. Fungsi Pembobot Spasial Fixed Gaussian Kernel**

Secara matematis, fungsi pembobot *Fixed Gaussian Kernel* dapat ditulis sebagai berikut:

$$w_j(u_i, v_i) = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{h} \right)^2 \right]$$

dengan  $h$  adalah *bandwidth* optimum dan  $d_{ij}$  adalah jarak *Euclidean* antara lokasi pengamatan  $(u_i, v_i)$  ke lokasi pengamatan  $(u_j, v_j)$  yang diperoleh dari persamaan berikut:

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

**E. Bandwidth**

Definisi *bandwidth* secara teoritis adalah lingkaran dengan radius  $h$  dari titik pusat lokasi pengamatan yang digunakan sebagai dasar penentuan pembobot untuk setiap lokasi pengamatan. Titik-titik lokasi pengamatan yang berada dalam lingkaran dengan radius  $h$  dari titik lokasi pengamatan ke- $i$  akan berpengaruh dalam menduga parameter di titik lokasi pengamatan ke- $i$ .

**F. Bandwidth Cross Validation (CV)**

Proses *Cross Validation* (CV) apabila diterapkan ke dalam konsep *bandwidth*, maka kelompok data *testing* diambil dengan cara menghilangkan nilai titik data ke- $i$  dari proses pendugaan  $y_i$  untuk dibandingkan dengan kelompok data *training* (data riil atau data keseluruhan) yang dilakukan dengan cara iterasi sampai diperoleh nilai CV minimum. Nilai *bandwidth* dikatakan optimum ketika nilai CV yang dihasilkan minimum. Secara matematis, CV dapat dituliskan sebagai berikut [4]:

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

Keterangan:

$\hat{y}_{\neq i}(h)$  : nilai penduga  $y_i$  dengan pengamatan di lokasi ke- $i$  dihilangkan dari proses pendugaan.

$y_i$  : variabel dependen pada pengamatan ke- $i$ .

$n$  : banyaknya pengamatan.

$h$  : bandwidth.

**G. Bandwidth Akaike Information Criterion Corrected (AICc)**

Metode pemilihan *bandwidth* dengan AICc dilakukan secara iterasi, yaitu dengan cara memasukkan berbagai nilai *bandwidth* ke dalam rumus AICc sampai diperoleh nilai AICc minimum. Nilai *bandwidth* dikatakan optimum ketika nilai AICc yang dihasilkan minimum. Secara matematis, AICc dapat dituliskan sebagai berikut [4]:

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

dengan

$$\hat{\sigma} = \sqrt{\frac{\mathbf{y}^T (\mathbf{I} - \mathbf{L})^T (\mathbf{I} - \mathbf{L}) \mathbf{y}}{n}}$$

$$\pi = 3,14$$

$I$  : matriks identitas berukuran  $n \times n$

$n$  : banyaknya pengamatan

$$L_{(n \times n)} = \begin{pmatrix} x_1^T [X^T W(u_1, v_1) X]^{-1} X^T W(u_1, v_1) \\ x_2^T [X^T W(u_2, v_2) X]^{-1} X^T W(u_2, v_2) \\ \vdots \\ x_n^T [X^T W(u_n, v_n) X]^{-1} X^T W(u_n, v_n) \end{pmatrix}$$

$x_i^T = (1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ip})$  adalah elemen baris ke-  $i$  dari  $X$

$$W(u_i, v_i) = \begin{pmatrix} w_1(u_i, v_i) & 0 & \dots & 0 \\ 0 & w_2(u_i, v_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n(u_i, v_i) \end{pmatrix}$$

**H. Bandwidth Bayesian Information Criterion (BIC)**

Metode pemilihan *bandwidth* menggunakan BIC didasarkan pada metode *maximum likelihood estimation* dan teorema Bayesian. Secara matematis, BIC dapat dituliskan sebagai berikut [4]:

$$BIC = k \ln(n) - 2 \ln(\hat{L})$$

dimana  $\hat{L}$  adalah nilai maksimum dari fungsi *likelihood* dari model,  $n$  adalah banyaknya pengamatan, dan  $k$  adalah banyaknya parameter dalam model. Nilai *bandwidth* dikatakan optimum ketika nilai BIC yang dihasilkan minimum.

**I. Uji Signifikansi Parameter Model GWR**

Uji signifikansi parameter model GWR dilakukan secara parsial. Hipotesis yang diuji sebagai berikut:

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

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

Statistik uji yang digunakan:

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

dengan

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

$$\hat{\sigma} = \sqrt{\frac{y^T (I-L)^T (I-L) y}{\delta_1}}$$

$$L_{(n \times n)} = \begin{pmatrix} x_1^T [X^T W(u_1, v_1) X]^{-1} X^T W(u_1, v_1) \\ x_2^T [X^T W(u_2, v_2) X]^{-1} X^T W(u_2, v_2) \\ \vdots \\ x_n^T [X^T W(u_n, v_n) X]^{-1} X^T W(u_n, v_n) \end{pmatrix}$$

$$\delta_i = \text{tr}([(I - L)^T (I - L)]^i); i = 1, 2$$

$$df = \frac{\delta_1^2}{\delta_2}$$

$c_{kk}$  : elemen diagonal ke- $k$  dari matriks  $CC^T$

Aturan keputusannya adalah  $H_0$  ditolak jika  $|t_k(u_i, v_i)| > t_{(\frac{\alpha}{2}, df)}$  atau jika  $p - value < \alpha$  [13].

**J. Penilaian Performa Model**

**RMSE**

Metode yang dapat digunakan untuk menghitung kesalahan dugaan adalah *Root Mean Squared Error* (RMSE). Metode RMSE menunjukkan besarnya perbedaan hasil dugaan dan hasil aktual. Semakin kecil nilai RMSE, maka semakin akurat dugaan yang dilakukan, artinya performa model tersebut semakin bagus. Perhitungan nilai RMSE dituliskan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

dengan  $y_i$  adalah nilai variabel dependen pada pengamatan ke- $i$ ,  $\hat{y}_i$  adalah nilai dugaan  $y_i$ , dan  $n$  adalah banyaknya pengamatan.

**Koefisien Determinasi ( $R^2$ )**

Koefisien determinasi ( $R^2$ ) menunjukkan ketepatan suatu model. Semakin besar nilai  $R^2$  suatu model, maka model tersebut semakin tepat menggambarkan hubungan antara variabel-variabelnya, artinya model tersebut memiliki performa yang semakin baik. Secara sistematis nilai  $R^2$  dituliskan sebagai berikut [14]:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

**Koefisien Determinasi Disesuaikan (Adjusted- $R^2$ )**

Koefisien determinasi disesuaikan atau *Adjusted- $R^2$*  ( $R_a^2$ ) adalah bentuk modifikasi  $R^2$  yang telah disesuaikan dengan banyaknya variabel independen dalam model. Setiap penambahan variabel independen akan menyebabkan terjadinya kenaikan nilai  $R^2$  yang berlebihan sehingga untuk menghindari hal tersebut digunakan nilai koefisien determinasi disesuaikan atau *Adjusted- $R^2$*  ( $R_a^2$ ) yang secara sistematis dituliskan sebagai berikut [15]:

$$R_a^2 = 1 - \left[ (1 - R^2) \left( \frac{n - 1}{n - p - 1} \right) \right]$$

dengan  $R^2$  adalah koefisien determinasi,  $n$  adalah banyaknya pengamatan, dan  $p$  adalah banyaknya variabel independen. Semakin besar nilai  $R_a^2$  suatu model, maka model tersebut memiliki performa yang semakin baik.

**III. METODOLOGI**

**A. Data Penelitian**

Penelitian ini menggunakan data yang diperoleh dari sumber lain yang telah ada, yang disebut dengan data sekunder. Penelitian ini menggunakan data pengangguran kabupaten/kota di Pulau Jawa tahun 2020 yang bersumber dari Badan Pusat Statistik (2021) yang tersaji pada buku Provinsi Jawa Timur Dalam Angka 2021, buku Provinsi Jawa Tengah Dalam Angka 2021, buku Provinsi Jawa Barat Dalam Angka 2021, buku Provinsi Banten Dalam Angka 2021, buku Provinsi DKI Jakarta Dalam Angka 2021, dan buku Provinsi DI Yogyakarta Dalam Angka 2021. Adapun area studi yang digunakan dalam penelitian ini adalah seluruh kabupaten/kota di Pulau Jawa yang berjumlah 118 kabupaten/kota.

**B. Variabel Penelitian**

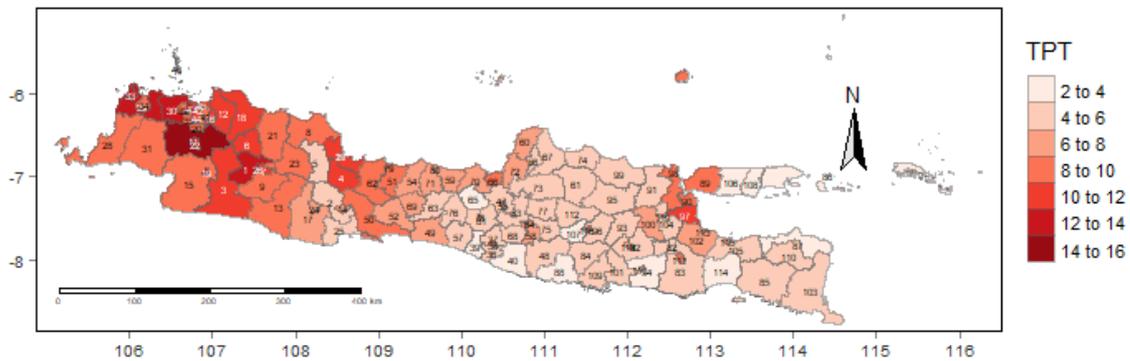
Variabel dalam penelitian ini terdiri dari variabel dependen dan variabel independen. Berikut adalah variabel penelitian yang digunakan:

1. Variabel dependen
  - TPT : Tingkat Pengangguran Terbuka (Persen)
2. Variabel independen
  - KP : Kepadatan Penduduk (Jiwa per Km<sup>2</sup>)
  - IPM : Indeks Pembangunan Manusia
  - TPAK : Tingkat Partisipasi Angkatan Kerja (Persen)
  - UMK : Upah Minimum Kabupaten/Kota (Rupiah)
  - RUSPF : Rata-rata Upah Sebulan Pekerja Formal (Rupiah)
  - RPBSPi : Rata-rata Pendapatan Bersih Sebulan Pekerja Informal (Rupiah).

**IV. HASIL DAN PEMBAHASAN**

**A. Deskripsi Data Tingkat Pengangguran Terbuka di Pulau Jawa**

Dalam laporan Badan Pusat Statistik, tingkat pengangguran terbuka di setiap kabupaten/kota di Pulau Jawa yang tercatat selama tahun 2020 beragam dari yang terendah di Kabupaten Gunung Kidul sebesar 2,16% sampai tertinggi di Kabupaten Bogor sebesar 14,29%. Data sebaran tingkat pengangguran terbuka di Pulau Jawa tahun 2020 dipetakan pada Gambar 1.



Gambar 1 Peta Sebaran Angka Tingkat Pengangguran Terbuka di Pulau Jawa Tahun 2020

Berdasarkan Gambar 1 dapat dilihat bahwa kabupaten/kota yang memiliki angka tingkat pengangguran terbuka yang tinggi cenderung berada di wilayah barat Pulau Jawa. Lima kabupaten/kota dengan tingkat pengangguran terbuka terendah adalah Kabupaten Gunung Kidul, Kabupaten Pacitan, Kabupaten Sumenep, Kabupaten Sampang, dan Kabupaten Lumajang, sedangkan lima kabupaten/kota dengan tingkat pengangguran terbuka tertinggi adalah Kabupaten Bogor, Kota Cimahi, Kabupaten Tangerang, Kota Cilegon, dan Kota Bogor.

**B. Uji Heterogenitas Spasial**

Hipotesis yang diuji dalam uji *Breusch-Pagan* adalah

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{118}^2 = \sigma^2 \text{ (Tidak terdapat heterogenitas spasial)}$$

$$H_1: \text{Minimal terdapat satu } \sigma_i^2 \neq \sigma^2; i = 1, \dots, 118 \text{ (Terdapat heterogenitas spasial)}$$

```
studentized Breusch-Pagan test
data: model
BP = 19.003, df = 6, p-value = 0.004158
```

Gambar 2. Output Software R Uji *Breusch-Pagan*

Hasil uji *Breusch-Pagan* pada Gambar 2 menunjukkan bahwa pada tingkat signifikansi  $\alpha = 0,05$ ;  $p - value = 0,004158 < \alpha$ , sehingga  $H_0$  ditolak yang berarti terdapat heterogenitas spasial.

**C. Pemodelan Geographically Weighted Regression (GWR) dengan Bandwidth Cross Validation (CV)**

Nilai *bandwidth* optimum yang diperoleh dengan metode CV adalah 0,9031467. Sehingga fungsi pembobot spasial *Fixed Gaussian Kernel* yang digunakan untuk membentuk matriks pembobot spasial adalah

$$w_j(u_i, v_i) = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{0,9031467} \right)^2 \right]$$

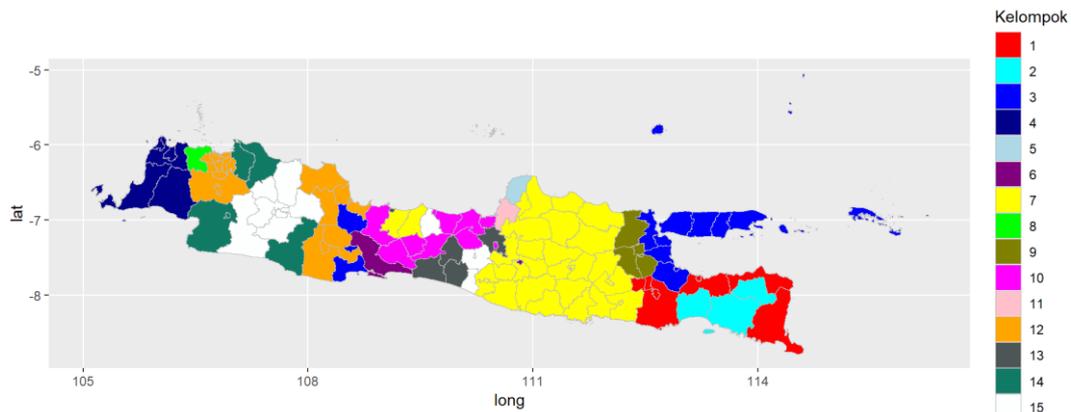
Berdasarkan pendugaan parameter model GWR dengan *bandwidth* CV, maka model GWR yang dihasilkan berbeda-beda untuk setiap kabupaten/kota, yaitu terdapat 118 model GWR *bandwidth* CV yang dihasilkan dari 118 kabupaten/kota yang ada di Pulau Jawa. Pemodelan GWR dengan *bandwidth* CV menghasilkan lima belas kelompok kabupaten/kota berdasarkan variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka yang dijelaskan pada Tabel 1.

Tabel 1 Kelompok Variabel Independen yang Signifikan Mempengaruhi Tingkat Pengangguran Terbuka Berdasarkan Model GWR dengan *Bandwidth* CV

Kelompok	Variabel Independen Signifikan	Jumlah Kab/Kota	Kode Wilayah
1	-	7	82, 83, 87, 103, 105, 111, 115
2	IPM	3	85, 110, 114
3	TPAK	12	4, 14, 25, 86, 89, 90, 97, 98, 102, 106, 108, 113
4	UMK	5	28, 29, 31, 33, 34
5	KP, IPM	1	60
6	KP, TPAK	3	50, 64, 78
7	KP, RPBSPI	32	36, 37, 38, 40, 48, 51, 53, 54, 56, 58, 61, 67, 68, 73, 74, 75, 77, 84, 88, 92, 93, 94, 95, 96, 99, 101, 107, 109, 112, 116, 117, 118
8	TPAK, UMK	1	30
9	TPAK, RPBSPI	4	91, 100, 104, 119
10	KP, IPM, TPAK	10	52, 55, 59, 62, 63, 65, 66, 69, 70, 80

Kelompok	Variabel Independen Signifikan	Jumlah Kab/Kota	Kode Wilayah
11	KP, IPM, RPBSPI	1	72
12	KP, TPAK, UMK	18	2, 5, 8, 10, 11, 16, 17, 20, 22, 24, 27, 32, 35, 41, 42, 43, 44, 45
13	KP, IPM, TPAK, UMK	5	47, 49, 57, 76, 79
14	KP, TPAK, UMK, RPBSPI	5	12, 13, 15, 18, 19
15	KP, IPM, TPAK, UMK, RPBSPI	11	1, 3, 6, 7, 9, 21, 23, 26, 39, 71, 81

Berdasarkan Tabel 1, kelompok 7 merupakan kelompok dengan jumlah kabupaten/kota yang paling banyak, yaitu terdiri dari 32 kabupaten/kota. Kelompok 7 merupakan kelompok kabupaten/kota dengan kepadatan penduduk dan rata-rata pendapatan bersih sebulan pekerja informal sebagai variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka.



Gambar 2. Peta Kelompok Kabupaten/Kota Berdasarkan Variabel Independen yang Signifikan untuk Model GWR dengan Bandwidth CV

Pada Gambar 2 terlihat bahwa kelompok 1, kelompok 2, kelompok 3, dan kelompok 9 cenderung berada di wilayah Pulau Jawa bagian timur. Kelompok 7 cenderung berada di antara Jawa Timur dan Jawa Tengah. Kelompok 5, kelompok 6, kelompok 10, kelompok 11, dan kelompok 13 cenderung berada di wilayah tengah Pulau Jawa. Sementara kelompok 4, kelompok 8, kelompok 12, kelompok 14, dan kelompok 15 cenderung berada di wilayah barat Pulau Jawa.

**D. Pemodelan Geographically Weighted Regression (GWR) dengan Bandwidth Akaike Information Criterion Corrected (AICc)**

Nilai *bandwidth* optimum yang diperoleh dengan metode AICc adalah 1,020452. Sehingga fungsi pembobot spasial *Fixed Gaussian Kernel* yang digunakan untuk membentuk matriks pembobot spasial adalah

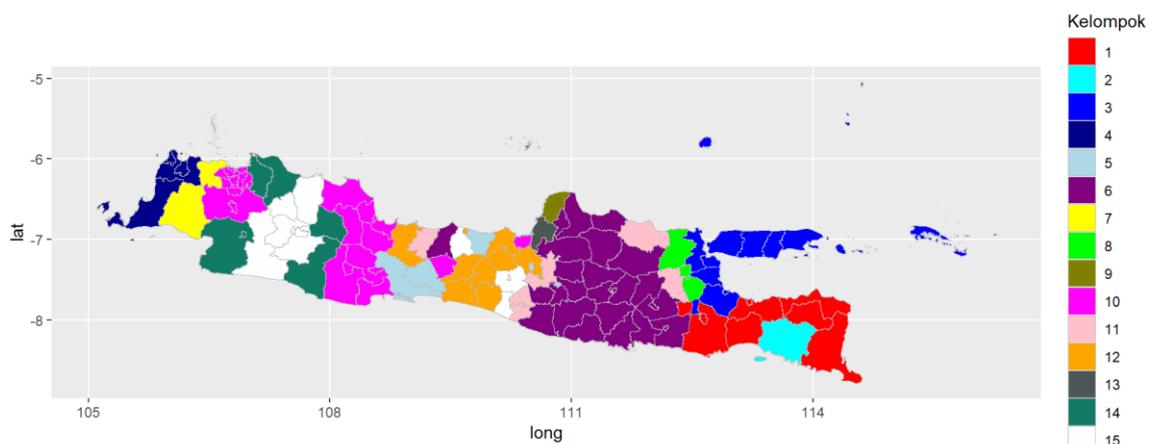
$$w_j(u_i, v_i) = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{1,020452} \right)^2 \right].$$

Tabel 2. Kelompok Variabel Independen yang Signifikan Mempengaruhi Tingkat Pengangguran Terbuka Berdasarkan Model GWR dengan Bandwidth AICc

Kelompok	Variabel Independen Signifikan	Jumlah Kab/Kota	Kode Wilayah
1	-	7	83, 87, 103, 105, 110, 114, 115
2	IPM	1	85
3	TPAK	11	82, 86, 89, 90, 97, 98, 102, 106, 108, 111, 113
4	UMK	4	28, 29, 33, 34
5	KP, TPAK	5	50, 52, 55, 59, 64
6	KP, RPBSPI	26	40, 48, 54, 56, 58, 61, 67, 68, 73, 74, 75, 77, 84, 88, 92, 93, 94, 95, 96, 101, 107, 109, 112, 116, 117, 118
7	TPAK, UMK	2	30, 31
8	TPAK, RPBSPI	3	91, 104, 119
9	KP, IPM, RPBSPI	1	60
10	KP, TPAK, UMK	24	2, 4, 5, 8, 10, 11, 14, 16, 17, 20, 22, 24, 25, 27, 32, 35, 41, 42, 43, 44, 45, 66, 69, 78

Kelompok	Variabel Independen Signifikan	Jumlah Kab/Kota	Kode Wilayah
11	KP, TPAK, RPBSPI	7	36, 37, 38, 51, 53, 99, 100
12	KP, IPM, TPAK, UMK	10	47, 49, 57, 62, 63, 65, 70, 76, 79, 80
13	KP, IPM, TPAK, RPBSPI	1	72
14	KP, TPAK, UMK, RPBSPI	6	12, 13, 15, 18, 19, 23
15	KP, IPM, TPAK, UMK, RPBSPI	10	1, 3, 6, 7, 9, 21, 26, 39, 71, 81

Berdasarkan pendugaan parameter model GWR dengan *bandwidth* AICc, maka model GWR yang dihasilkan berbeda-beda untuk setiap kabupaten/kota, yaitu terdapat 118 model GWR *bandwidth* AICc yang dihasilkan dari 118 kabupaten/kota yang ada di Pulau Jawa. Pemodelan GWR dengan *bandwidth* AICc menghasilkan lima belas kelompok kabupaten/kota berdasarkan variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka yang dijelaskan pada Tabel 2. Berdasarkan Tabel 2, kelompok 6 merupakan kelompok dengan jumlah kabupaten/kota yang paling banyak, yaitu terdiri dari 26 kabupaten/kota. Kelompok 6 merupakan kelompok kabupaten/kota dengan kepadatan penduduk dan rata-rata pendapatan bersih sebulan pekerja informal sebagai variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka.



Gambar 3 Peta Kelompok Kabupaten/Kota Berdasarkan Variabel Independen yang Signifikan untuk Model GWR dengan *Bandwidth* AICc

Pada Gambar 3 terlihat bahwa kelompok 1, kelompok 2, kelompok 3, kelompok 8, dan kelompok 11 cenderung berada di wilayah Pulau Jawa bagian timur. Kelompok 6 cenderung berada di antara Jawa Timur dan Jawa Tengah. Kelompok 5, kelompok 9, kelompok 12, dan kelompok 13 cenderung berada di wilayah tengah Pulau Jawa. Sementara kelompok 4, kelompok 7, kelompok 10, kelompok 14, dan kelompok 15 cenderung berada di wilayah barat Pulau Jawa.

**E. Pemodelan Geographically Weighted Regression (GWR) dengan Bandwidth Bayesian Information Criterion (BIC)**

Nilai *bandwidth* optimum yang diperoleh dengan metode BIC adalah 2,195162. Sehingga fungsi pembobot spasial *Fixed Gaussian Kernel* yang digunakan untuk membentuk matriks pembobot spasial adalah

$$w_j(u_i, v_i) = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{2,195162} \right)^2 \right].$$

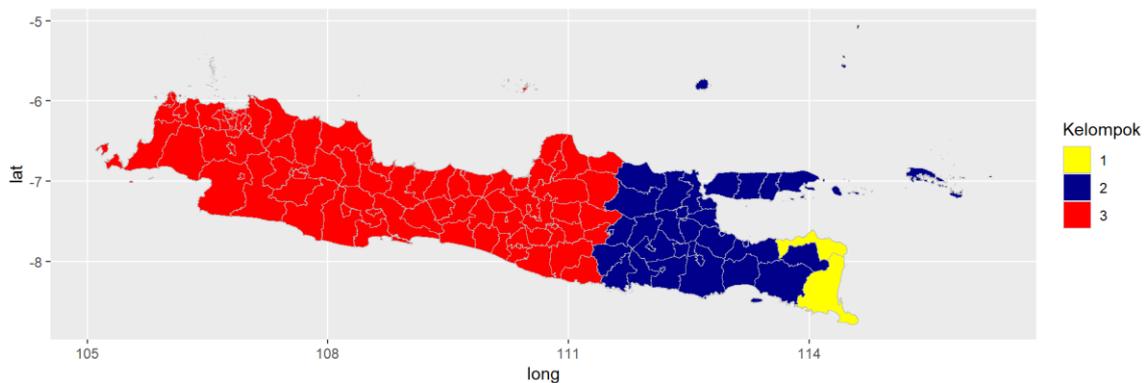
Berdasarkan pendugaan parameter model GWR dengan *bandwidth* BIC, maka model GWR yang dihasilkan berbeda-beda untuk setiap kabupaten/kota, yaitu terdapat 118 model GWR *bandwidth* BIC yang dihasilkan dari 118 kabupaten/kota yang ada di Pulau Jawa. Pemodelan GWR dengan *bandwidth* BIC menghasilkan tiga kelompok kabupaten/kota berdasarkan variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka yang dijelaskan pada Tabel 3.

Tabel 3 Kelompok Variabel Independen yang Signifikan Mempengaruhi Tingkat Pengangguran Terbuka Berdasarkan Model GWR dengan Bandwidth BIC

Kelompok	Variabel Independen Signifikan	Jumlah Kab/Kota	Kode Wilayah
1	TPAK	2	87, 103
2	KP, TPAK	33	82, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 104, 105, 106, 108, 109, 110, 111, 113, 114, 115, 116, 117, 118, 119

Kelompok	Variabel Signifikan	Independen	Jumlah Kab/Kota	Kode Wilayah
3	KP, TPAK, UMK		83	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 88, 107, 112

Berdasarkan Tabel 4, kelompok 3 merupakan kelompok dengan jumlah kabupaten/kota yang paling banyak, yaitu terdiri dari 83 kabupaten/kota. Variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka pada kelompok ini adalah kepadatan penduduk, tingkat partisipasi angkatan kerja, dan upah minimum kabupaten/kota. Sementara, kelompok 1 merupakan kelompok kabupaten/kota dengan tingkat partisipasi angkatan kerja sebagai variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka. Kelompok 2 merupakan kelompok kabupaten/kota dengan kepadatan penduduk dan tingkat partisipasi angkatan kerja sebagai variabel independen yang signifikan mempengaruhi tingkat pengangguran terbuka.



Gambar 4. Peta Kelompok Kabupaten/Kota Berdasarkan Variabel Independen yang Signifikan untuk Model GWR dengan *Bandwidth* BIC

Pada Gambar 4 terlihat bahwa kelompok 1 berada di Kabupaten Banyuwangi dan Kabupaten Situbondo, sedangkan sebagian besar kabupaten/kota lainnya di Provinsi Jawa Timur masuk ke kelompok 2. Adapun kelompok 3 berada di wilayah barat dan tengah Pulau Jawa tepatnya di Provinsi Jawa Barat, Provinsi Banten, Provinsi DKI Jakarta, Provinsi DI Yogyakarta, dan Provinsi Jawa Tengah.

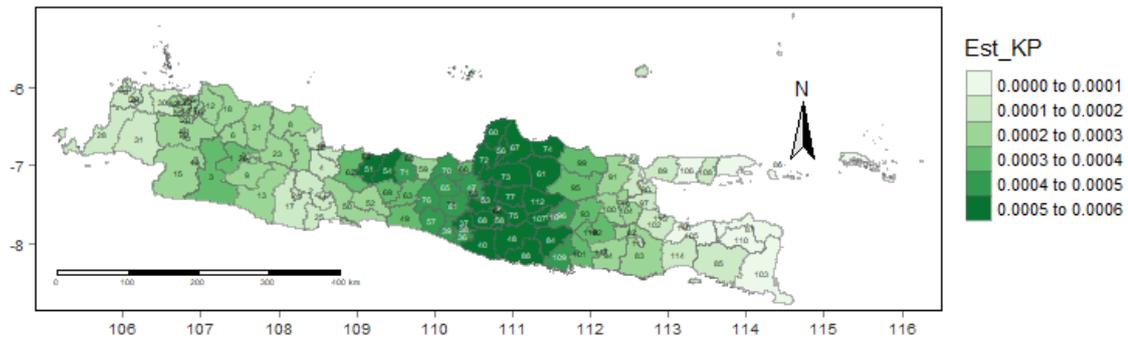
**F. Hasil Perbandingan Performa Model Geographically Weighted Regression dengan *Bandwidth* CV, AICc, dan BIC**

Tabel 4. Nilai RMSE,  $R^2$ , dan Adjusted- $R^2$  untuk Model GWR dengan *Bandwidth* CV, AICc, dan BIC

Model	RMSE	$R^2$	Adjusted- $R^2$
GWR <i>bandwidth</i> CV	1,0904	0,8539011	0,7937159
GWR <i>bandwidth</i> AICc	1,1336	0,8420887	0,7874906
GWR <i>bandwidth</i> BIC	1,3747	0,7677763	0,7331337

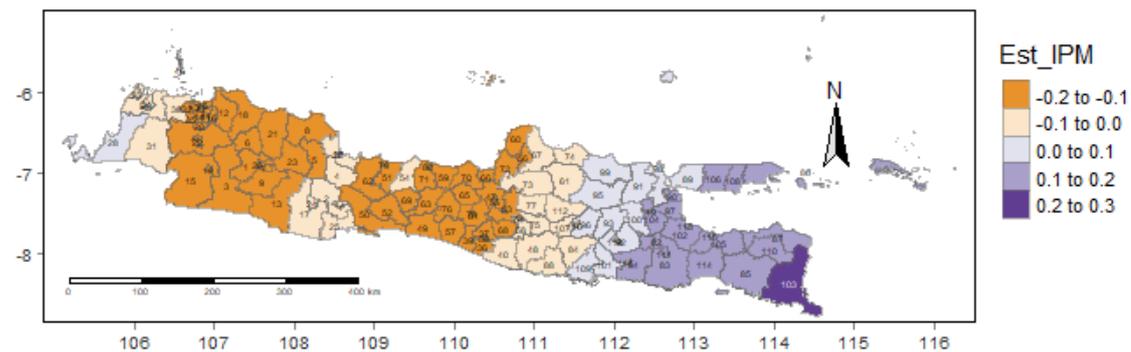
Berdasarkan Tabel 4, dapat dilihat bahwa model GWR dengan *bandwidth* CV mendapatkan nilai RMSE paling rendah dibandingkan model GWR dengan *bandwidth* AICc dan *bandwidth* BIC. Berdasarkan kriteria  $R^2$  dan Adjusted- $R^2$ , model GWR dengan *bandwidth* CV mendapatkan nilai  $R^2$  dan Adjusted- $R^2$  paling tinggi dibandingkan model GWR dengan *bandwidth* AICc dan *bandwidth* BIC. Hal ini berarti penggunaan *bandwidth* CV pada model GWR menghasilkan performa model yang lebih baik dibandingkan *bandwidth* AICc dan *bandwidth* BIC untuk menganalisis data pengangguran kabupaten/kota di Pulau Jawa tahun 2020.

Keragaman Spasial Penduga Parameter Model GWR dengan Performa Terbaik



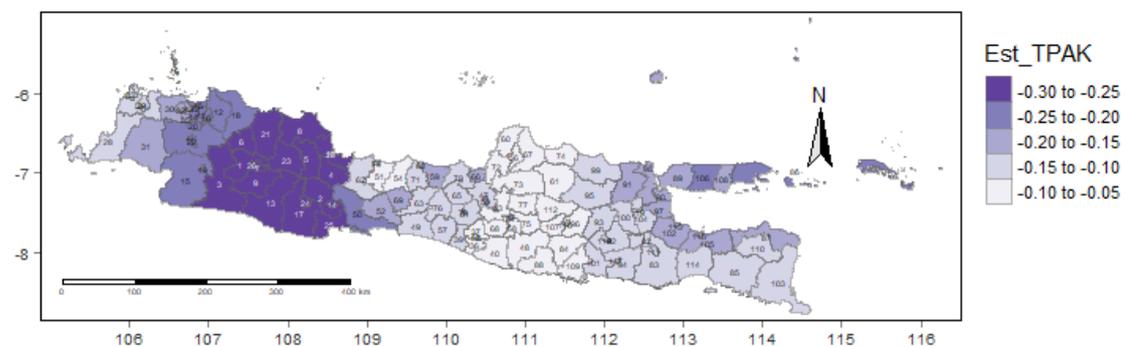
Gambar 5 Peta Keragaman Spasial  $\hat{\beta}_1$  (Koefisien Kepadatan Penduduk)

Pada Gambar 5 terlihat bahwa bubungan angka kepadatan penduduk dan angka tingkat pengangguran terbuka bersifat positif di seluruh lokasi kabupaten/kota di Pulau Jawa. Hubungan positif yang paling kuat berada di kabupaten/kota di Provinsi Jawa Tengah, sementara hubungan positif yang paling lemah berada di kabupaten/kota di Provinsi Jawa Timur. Hubungan positif ini berarti bahwa semakin tinggi angka kepadatan penduduk di suatu daerah maka semakin tinggi pula angka tingkat pengangguran terbuka di suatu daerah tersebut, dan sebaliknya.



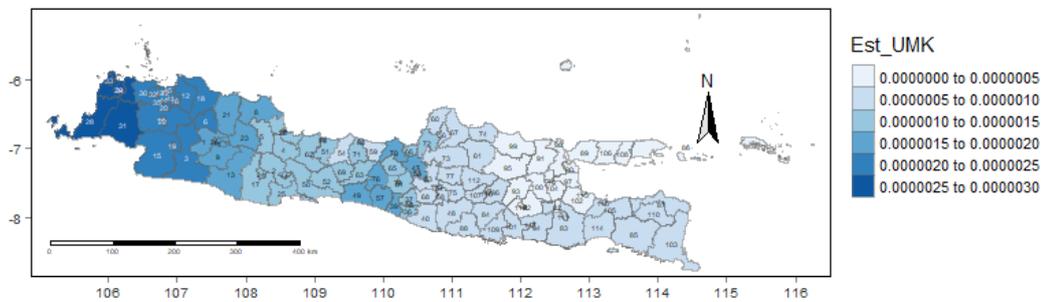
Gambar 6 Peta Keragaman Spasial  $\hat{\beta}_2$  (Koefisien Indeks Pembangunan Manusia)

Pada Gambar 6 terlihat bahwa hubungan indeks pembangunan manusia dan angka tingkat pengangguran terbuka bersifat negatif di hampir seluruh kabupaten/kota di wilayah barat dan tengah Pulau Jawa, sementara kabupaten/kota di wilayah timur Pulau Jawa memiliki hubungan yang positif. Hubungan negatif ini berarti bahwa semakin tinggi angka indeks pembangunan manusia maka angka tingkat pengangguran terbuka akan menurun, dan sebaliknya. Sementara, hubungan positif berarti bahwa semakin tinggi angka indeks pembangunan manusia maka angka tingkat pengangguran terbuka akan semakin tinggi pula, dan sebaliknya.



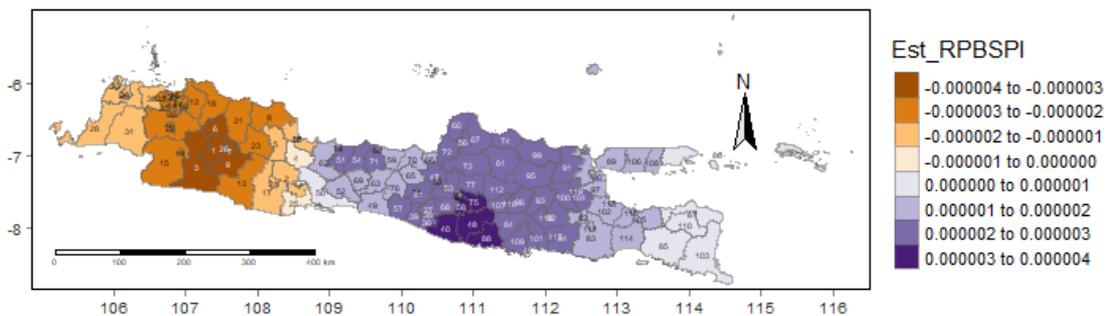
Gambar 7 Peta Keragaman Spasial  $\hat{\beta}_3$  (Koefisien Tingkat Partisipasi Angkatan Kerja)

Pada Gambar 7 terlihat bahwa hubungan angka tingkat partisipasi angkatan kerja dan angka tingkat pengangguran terbuka bersifat negatif di seluruh lokasi kabupaten/kota di Pulau Jawa. Hubungan negatif yang paling kuat berada di kabupaten/kota di Provinsi Jawa Barat, sementara hubungan negatif yang paling lemah berada di kabupaten/kota di Provinsi Jawa Tengah. Hubungan negatif ini berarti bahwa semakin tinggi angka tingkat partisipasi angkatan kerja di suatu daerah maka angka tingkat pengangguran terbuka di suatu daerah tersebut semakin menurun, dan sebaliknya.



Gambar 8 Peta Keragaman Spasial  $\hat{\beta}_4$  (Koefisien Upah Minimum Kabupaten/Kota)

Pada Gambar 8 terlihat bahwa hubungan upah minimum kabupaten/kota dan angka tingkat pengangguran terbuka bersifat positif di seluruh lokasi kabupaten/kota di Pulau Jawa. Hubungan positif yang paling kuat berada di kabupaten/kota di Provinsi Banten, sementara hubungan positif yang paling lemah berada di kabupaten/kota di Provinsi Jawa Timur. Hubungan positif ini berarti bahwa semakin tinggi upah minimum kabupaten/kota di suatu daerah maka angka tingkat pengangguran terbuka di suatu daerah tersebut juga semakin tinggi, dan sebaliknya.



Gambar 9 Peta Keragaman Spasial  $\hat{\beta}_6$  (Koefisien Rata-rata Pendapatan Bersih Sebulan Pekerja Informal)

Pada Gambar 9 terlihat bahwa hubungan rata-rata pendapatan bersih sebulan pekerja informal dan angka tingkat pengangguran terbuka bersifat negatif di seluruh kabupaten/kota di wilayah barat Pulau Jawa, sementara seluruh kabupaten/kota di wilayah timur dan tengah Pulau Jawa memiliki hubungan yang positif. Hubungan negatif ini berarti bahwa semakin tinggi rata-rata pendapatan bersih sebulan pekerja informal di daerah tersebut maka angka tingkat pengangguran terbuka di daerah tersebut akan semakin menurun, dan sebaliknya. Sementara, hubungan positif berarti bahwa semakin tinggi rata-rata pendapatan bersih sebulan pekerja informal di daerah tersebut maka angka tingkat pengangguran terbuka di daerah tersebut juga akan semakin tinggi, dan sebaliknya.

## V. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian yang telah dilakukan yaitu, berdasarkan pendugaan parameter model GWR dengan *bandwidth* CV, AICc, dan BIC pada data pengangguran kabupaten/kota di Pulau Jawa tahun 2020, maka model GWR yang dihasilkan berbeda-beda untuk setiap kabupaten/kota, yaitu terdapat 118 model GWR *bandwidth* CV, AICc, dan BIC yang dihasilkan dari 118 kabupaten/kota yang ada di Pulau Jawa. Pemodelan GWR dengan *bandwidth* CV memberikan hasil, terdapat 15 kelompok kabupaten/kota dengan signifikansi variabel independen yang berbeda-beda. Kelompok 7 adalah kelompok dengan jumlah kabupaten/kota terbanyak yang terdiri dari 32 kabupaten/kota di sebagian Jawa Timur, sebagian Jawa Tengah, dan sebagian besar DI Yogyakarta dengan variabel kepadatan penduduk dan rata-rata pendapatan bersih sebulan pekerja informal signifikan mempengaruhi tingkat pengangguran terbuka. Pemodelan GWR dengan *bandwidth* AICc memberikan hasil, terdapat 15 kelompok kabupaten/kota dengan signifikansi variabel independen yang berbeda-beda. Kelompok 6 adalah kelompok dengan jumlah kabupaten/kota terbanyak yang terdiri dari 26 kabupaten/kota di sebagian Jawa Timur, sebagian Jawa Tengah, dan sebagian kecil DI Yogyakarta dengan variabel kepadatan penduduk dan rata-rata pendapatan bersih sebulan pekerja informal signifikan mempengaruhi tingkat pengangguran terbuka. Pemodelan GWR dengan *bandwidth* BIC memberikan hasil, terdapat 3 kelompok kabupaten/kota dengan signifikansi variabel independen yang berbeda-beda. Kelompok 1 adalah kelompok yang hanya terdiri dari Kabupaten Banyuwangi dan Kabupaten Situbondo dengan variabel tingkat partisipasi angkatan kerja signifikan mempengaruhi tingkat pengangguran terbuka. Kelompok 2 adalah kelompok yang terdiri dari 33 kabupaten/kota di Jawa Timur dengan variabel kepadatan penduduk dan tingkat partisipasi angkatan kerja signifikan mempengaruhi tingkat pengangguran terbuka. Adapun kelompok 3 adalah kelompok dengan jumlah kabupaten/kota terbanyak yang terdiri dari 83 kabupaten/kota di Jawa Barat, Banten, DKI Jakarta, DI Yogyakarta, Jawa Tengah, dan sebagian kecil Jawa Timur dengan variabel kepadatan penduduk, tingkat partisipasi angkatan kerja, dan upah minimum kabupaten/kota signifikan mempengaruhi tingkat pengangguran terbuka. Model GWR dengan *bandwidth* CV memiliki

performa yang lebih baik dibandingkan model GWR dengan *bandwidth* AICc dan *bandwidth* BIC dalam menjelaskan data pengangguran kabupaten/kota di Pulau Jawa tahun 2020 karena memiliki nilai RMSE paling kecil dan memiliki nilai  $R^2$  dan *Adjusted-R<sup>2</sup>* paling besar.

Beberapa saran yang diajukan oleh penulis adalah

1. Hasil penelitian ini dapat digunakan sebagai bahan pertimbangan bagi pemerintah daerah di Pulau Jawa untuk membuat kebijakan dalam upaya mengurangi angka pengangguran dengan memperhatikan variabel-variabel yang signifikan di suatu daerah tersebut.
2. Hati-hati menggunakan metode BIC dalam menentukan *bandwidth* optimum pada model GWR, karena berdasarkan hasil penelitian ini, model GWR dengan *bandwidth* BIC menghasilkan performa model yang paling buruk dibandingkan dengan *bandwidth* CV dan AICc, serta model GWR yang dihasilkan menggunakan *bandwidth* BIC terlampau halus.

## DAFTAR PUSTAKA

- [1] BPS. (2021). *Tingkat Pengangguran Terbuka Menurut Provinsi (Persen), 2020-2021*. Available at: <https://www.bps.go.id/indicator/6/543/1/tingkat-pengangguran-terbuka-menurut-provinsi.html> [Accessed 16 Maret 2022].
- [2] Jumita, V.F., "Tingkat Pengangguran, Pengeluaran Pemerintah pada Sektor Infrastruktur, dan Investasi Terhadap Kesejahteraan Masyarakat di 13 Kabupaten/Kota Provinsi Jawa Tengah Tahun 2015-2019", Bachelor's thesis, Fakultas Ekonomi dan Bisnis, UIN Jakarta, Jakarta, 2021.
- [3] Ishak, K., "Faktor-Faktor Yang Mempengaruhi Pengangguran Dan Inflikasi Terhadap Indeks Pembangunan Di Indonesia", *Iqtishaduna: Jurnal Ilmiah Ekonomi Kita*, 7(1), pp.22-38. 2018.
- [4] Fotheringham, A.S., Brunson, C., & Charlton, M., *Geographically Weighted regression: The Analysis of Spatially Varying Relationships*. England: John Wiley and Sons, Ltd. 2002
- [5] Fotheringham, A.S., & Charlton, M., *Geographically Weighted Regression: White Paper*. Kildare: *National University of Ireland Maynooth*, 1-14. 2009.
- [6] Guo, L. M., "Comparison of Bandwidth Selection in Application of Geographically Weighted Regression: A Case Study", *Canadian Journal of Forest Research*, 38(9): 2526-2534. 2008.
- [7] Rosa, A. A., "Penggunaan Pembobot Fixed Gaussian Kernel dan Fixed Bisquare Kernel pada Model Geographically Weighted Regression", Skripsi, Fakultas MIPA, Universitas Hasanuddin, Makassar, 2015.
- [8] Mertha, W. P., "Analisis Hubungan Kondisional Sektor Ekonomi dan Penelitian terhadap Angka Kemiskinan di Jawa Timur menggunakan Metode GWR", Skripsi, Institut Teknologi Sepuluh November, Surabaya, 2008.
- [9] Ariyanto, D., "Perbandingan Bandwidth Cross Validation dan Bandwidth Akaike Information Criterion dalam Pembentukan Fungsi Pembobot Adaptive Gaussian Kernel pada Geographically Weighted Regression (Studi Kasus Produktivitas Padi Sawah di Kabupaten Tulungagung)", Tesis, Fakultas MIPA, Universitas Brawijaya Malang, Malang, 2017. <http://repository.ub.ac.id/id/eprint/2510>.
- [10] Amin, R.A., "PERFORMA BANDWIDTH CROSS VALIDATION DAN AKAIKE INFORMATION CRITERION CORRECTED PADA MODEL GEOGRAPHICALLY WEIGHTED REGRESSION (Studi Kasus: Jumlah Penduduk Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2018)", Skripsi, Fakultas MIPA, Universitas Hasanuddin, Makassar, 2021. <http://repository.unhas.ac.id/id/eprint/3800>.
- [11] Anselin, L., *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers. 1988.
- [12] Brunson, C., Fotheringham, S., & Charlton, M., *Geographically weighted regression-modelling spatial non-stationarity*. 47(3), 431-443. 2012.
- [13] Purhadi, & Yasin, H., "Mixed Geographically Weighted Regression Model (Case Study: the Percentage of Poor Households in Mojokerto 2008)", *European Journal of Scientific Research*, 69(2): 188-196. 2012.
- [14] Gujarati, D., *Ekonometrika Dasar*, Cetakan ketiga. Jakarta: Erlangga. 1993.
- [15] Harel, O., "The Estimation of  $R^2$  and Adjusted  $R^2$  in Incomplete Data Sets Using Multiple Imputation", *Journal of Applied Statistics*, 36(10): 1109- 1118. 2009.



© 2023 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (<http://creativecommons.org/licenses/by-sa/4.0/>).