

# Perbandingan Performa Bandwidth CV, AICc, dan BIC dalam Pembentukan Fungsi Pembobot Fixed Gaussian Kernel pada Model Geographically Weighted Regression: Aplikasi pada Data Pengangguran di Pulau Jawa = Comparison of CV, AICc, and BIC Bandwidth Performance in the Formation of Fixed Gaussian Kernel Weighted Function in Geographically Weighted Regression Models: Applications on Unemployment Data in Java

Carisa Putri Salsabila Purnamasari, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20528475&lokasi=lokal>

---

## 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 (TPT). Dibandingkan dengan metode analisis regresi linier, metode Geographically Weighted Regression (GWR) lebih diunggulkan karena dapat menangani masalah ketidakstasioneran spasial yang biasanya terjadi pada data fenomena sosial. Ketidakstasioneran spasial adalah situasi dimana hubungan antar variabel berbeda-beda secara signifikan di setiap lokasi pengamatan. Ketidakstasioneran spasial ini sering disebut juga dengan heterogen spasial. 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 sehingga model yang diperoleh berlaku hanya untuk lokasi tersebut. Penentuan pembobot bergantung pada nilai bandwidth. Bandwidth merupakan lingkaran dengan radius dari titik pusat lokasi pengamatan yang digunakan sebagai dasar penentuan pembobot setiap lokasi pengamatan. 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 CV, AICc, dan BIC dalam pembentukan fungsi pembobot Fixed Gaussian Kernel yang diterapkan pada data pengangguran di kabupaten/kota di Pulau Jawa. Variabel dependen yang digunakan dalam penelitian ini adalah tingkat pengangguran terbuka kabupaten/kota di Pulau Jawa, dan variabel independen yang digunakan 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. Hasil penelitian menunjukkan bahwa setiap kabupaten/kota memiliki model GWR yang berbeda-beda. 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 R2 dan Adjusted-R2 paling besar, yaitu 0,8539011 dan 0,7937159.

.....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 (TPT). Rather than using linear regression analysis, Geographically Weighted Regression (GWR) was preferable since it gave a better representative model by effectively resolve spatial non-stationary problem which is generally exist in spatial data of social phenomenon. Spatial non-stationary is a situation when the relationship between variables are significantly different in each location of observation point. This spatial non-stationary is often refer to spatial heterogeneity. 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 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 CV, AICc, and BIC bandwidth methods in the formation of Fixed Gaussian Kernel weighted function which is applied to unemployment data in districts/cities in Java. The dependent variable used in this study is the district/city open unemployment rate in Java, and the independent variables are population density, human development index, labor force participation rate, district/city minimum wage, the average monthly wage of formal workers, and the average monthly net income of informal workers. The results show that each district/city has a different GWR model. 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 R2 and Adjusted-R2 values, namely 0.8539011 and 0.7937159, respectively.