

Geographically Weighted Regression with Kernel Weighted Function on Poverty Cases in West Java Province*

Regresi Terboboti Geografis dengan Fungsi Pembobot Kernel pada Data Kemiskinan di Provinsi Jawa Barat

Winda Nurpadilah¹, I Made Sumertajaya^{2‡}, and Muhammad Nur Aidi³

^{1,2,3}Department of Statistics, IPB University, Indonesia, windanurfadilah17@gmail.com

[‡]corresponding author: imsjaya.stk@gmail.com

Copyright © 2021 Winda Nurpadilah, I Made Sumertajaya, and Muhammad Nur Aidi. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Spatial regression analysis is a form of regression model that considers spatial effects. Geographically weighted regression (GWR) is the spatial regression methods that can be used to deal with the problem of spatial diversity. This method generates local model parameter estimates for each observation location. The application of spatial statistics can be done in all areas such as the problem of poverty. Poverty can be influenced by factors of proximity between regions, so that in determining the poverty factor, the proximity factor of the region cannot be ignored. West Java Province is a province with the largest population, so this study aims to model the poverty data in West Java Province by incorporating spatial effects. The weighting function used for the GWR model is the function of the fixed and adaptive kernels. The analysis results show that the fixed exponential kernel function has the smallest cross validation (CV) value, so the weighting matrix used in the model is determined by the exponential kernel function. The largest R^2 value and the smallest AIC value are owned by the GWR model with an exponential kernel function. Based on the results obtained by the the ANOVA table to test GWR's global goodness, the GWR model is more effective than global regression. Therefore, the GWR model is the best model when it used in West Java's poverty cases. The effect of each explanatory variable on the percentage of poverty varies in each district/city in West Java Province.

Keywords: adaptive, fixed, geographically weighted regression, poverty, West Java.

* Received: Jan 2021; Reviewed: Mar 2021; Published: Mar 2021

1. Pendahuluan

Masalah statistika tentunya tidak terlepas dari data. Data yang memiliki informasi terkait dengan lokasi atau letak geografis suatu tempat termasuk dalam kategori spasial. Analisis regresi spasial merupakan bentuk model regresi yang mempertimbangkan pengaruh spasial. Menurut Anselin dan Getis (1992), pengaruh spasial antar lokasi dapat disebabkan oleh ketergantungan dan keragaman spasial. Beberapa metode regresi spasial yang diketahui diantaranya adalah model umum spasial (general spatial model, GSM), model galat spasial (spatial error model, SEM), model otoregresif spasial (spatial autoregressive model, SAR), dan model regresi terboboti geografis (geographically weighted regression, GWR). Salah satu metode regresi yang memperhatikan letak geografis pengamatan dan digunakan untuk menangani masalah keragaman spasial adalah metode regresi terboboti geografis (RTG) (Fotheringham et al., 2002). Metode RTG menghasilkan pendugaan parameter yang bersifat lokal pada setiap lokasi atau letak geografis suatu tempat pengamatan (Purhadi & Yasin, 2012).

Provinsi Jawa Barat adalah suatu provinsi yang diketahui mempunyai jumlah penduduk terbanyak, sekitar 48 ribu jiwa. Provinsi Jawa Barat pada bulan Maret 2018 berhasil meningkatkan garis kemiskinan sebesar 3.69% dan menurunkan jumlah penduduk miskin sekitar 159 ribu jiwa. Penerapan statistika spasial dapat dilakukan dalam semua bidang seperti permasalahan kemiskinan. Penerapan tersebut bertujuan untuk mengetahui faktor-faktor yang diduga memiliki pengaruh terhadap tinggi rendahnya suatu kemiskinan pada suatu daerah sehingga dapat dirumuskan sebuah kebijakan publik yang mampu mengurangi kemiskinan. Kemiskinan dapat dipengaruhi oleh faktor kedekatan antar wilayah. Oleh karena itu, penentuan faktor-faktor yang mempengaruhi kemiskinan tidak dapat mengabaikan faktor kedekatan antar wilayah. Kemiskinan pada suatu daerah tidak hanya dipengaruhi oleh faktor-faktor kemiskinan yang dimiliki oleh daerah tersebut, akan tetapi dipengaruhi oleh kemiskinan yang dimiliki daerah lain.

Penelitian ini memiliki tujuan untuk memodelkan data persentase kemiskinan di Provinsi Jawa Barat dengan memasukkan efek spasial. Metode yang akan diterapkan adalah metode regresi terboboti geografis. Dengan metode ini, peneliti dapat menilai variasi spasial (geografis) yang mungkin ada dalam hubungan antara peubah respon dan peubah penjelas pada seluruh lokasi atau letak geografis suatu tempat pengamatan (Clement et al., 2009). Pendugaan koefisien regresi dalam model dilakukan dengan menggunakan metode kuadrat terkecil terboboti atau weighted least square (WLS). Hal ini menunjukkan bahwa setiap lokasi di mana data dikumpulkan akan diberikan pembobot yang berbeda. Pemberian bobot untuk setiap data dilakukan berdasarkan kedekatan antar titik lokasi pengamatan (Caraka & Yasin, 2017). Dalam penelitian ini pendugaan parameter dilakukan menggunakan fungsi pembobot fix dan adaptif Kernel Gaussian.

2. Metodologi

2.1 Bahan dan Data

Data yang diterapkan dalam penelitian ini merupakan data sekunder yang bersumber

dari Badan Pusat Statistika (2018) yaitu data persentase kemiskinan sebagai peubah respon dan enam peubah penjelas yang diduga mempengaruhi persentase kemiskinan. Peubah penjelas yang digunakan dipilih berdasarkan penelitian-penelitian yang telah dilakukan sebelumnya. Jangkauan yang digunakan dalam penelitian adalah kabupaten/kota di Provinsi Jawa Barat tahun 2018. Peubah penjelas dan peubah respon yang digunakan dalam penelitian ini disajikan pada Tabel 1.

Tabel 1: Peubah yang digunakan dalam penelitian

Peubah	Keterangan
Y	Persentase kemiskinan (%)
X_1	Produk domestik regional bruto (Miliar Rupiah)
X_2	Angka melek huruf (AMH)
X_3	Persentase pengeluaran perkapita untuk makanan (%)
X_4	Persentase rumah tangga yang pernah membeli raskin (%)
X_5	Angka harapan lama sekolah (Tahun)
X_6	Indeks kesehatan

2.2 Metode Penelitian

Langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Mendeskripsikan data persentase penduduk miskin pada setiap kabupaten/kota di Provinsi Jawa Barat tahun 2018.
2. Menghitung nilai korelasi pearson untuk mengetahui pola hubungan peubah penjelas yang diduga mempengaruhi peubah respon.
3. Melakukan regresi global diantara sebagai berikut:
 - a. Pendugaan parameter.
 - b. Pengujian parsial parameter regresi.
 - c. Pengujian serentak parameter regresi.
 - d. Pengujian asumsi-asumsi klasik.
4. Melakukan pengujian multikolinearitas antar peubah-peubah penjelas dengan menggunakan nilai VIF, jika $VIF > 10$ maka terjadi multikolinearitas antar peubah.
5. Mengidentifikasi keragaman spasial pada data menggunakan uji *Breusch-Pagan* (BP). Hipotesis yang digunakan adalah sebagai berikut:

$$H_0: \beta_2 = \beta_3 = \dots = \beta_k = 0; k = 1, 2, \dots, p$$

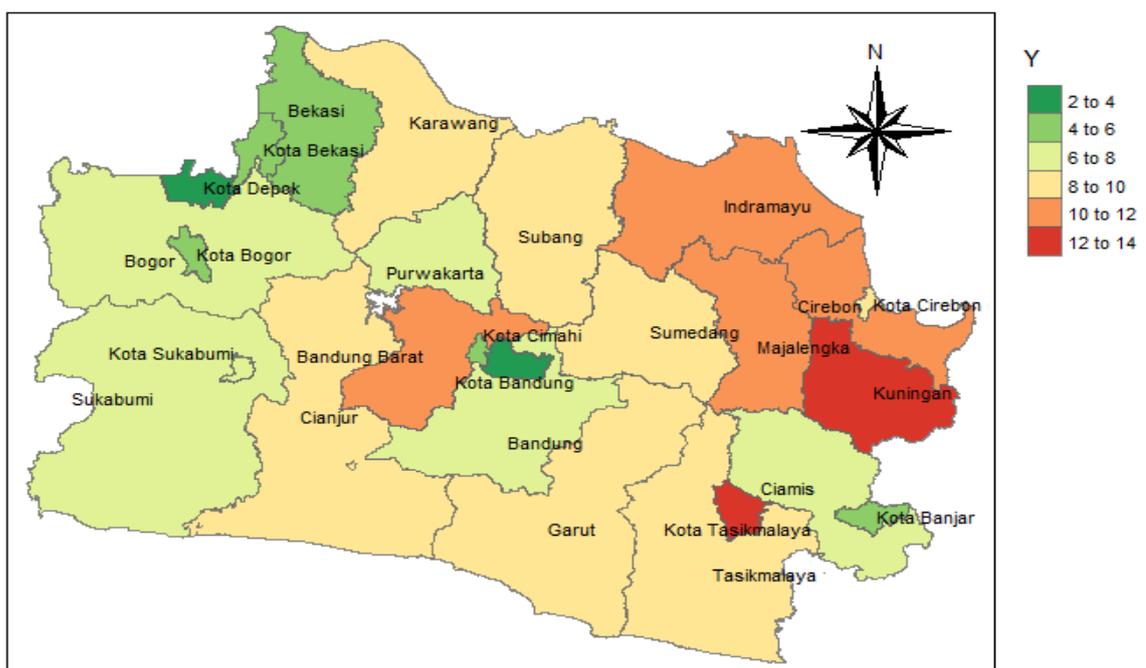
$$H_1: \sigma_i^2 \neq \beta_1 = \text{konstan}; i = 1, 2, \dots, n$$
 Statistik uji *Breusch-Pagan*: $BP = \frac{1}{2} (\sum_{i=1}^n x_i f_i)^T (\sum_{i=1}^n x_i x_i)^T (\sum_{i=1}^n x_i f_i) \sim \chi_{(k-1)}^2$
 dengan: $f_i = \left(\frac{\hat{z}_i}{\hat{\sigma}^2} - 1 \right)$, $\hat{z}_i = (y_i - \hat{\beta}^T x_i)$, $\hat{\sigma}^2 = \sum_{i=1}^n \hat{z}_i^2$. Kriteria uji yang digunakan adalah tolak H_0 jika $BP > \chi_{(k-1)}^2$ atau jika nilai-p < taraf nyata dengan k adalah banyaknya parameter (Arbia 2006).
6. Melakukan pemodelan RTG yang meliputi tahapan sebagai berikut:
 - a. Menghitung jarak Euclidean d_{ij} antara lokasi ke- i dan lokasi ke- j
 - b. Menentukan *bandwidth* spasial (h_s) optimum yang dipilih berdasarkan nilai *cross validation* (CV) yang minimum.
 - c. Menghitung matriks pembobot model dengan *bandwidth* optimum
 - d. Menghitung penduga parameter model RTG untuk setiap lokasi.
 - e. Melakukan pengujian parameter model RTG secara parsial.
7. Menentukan model yang paling sesuai untuk data persentase kemiskinan Provinsi Jawa Barat menggunakan kriteria nilai AIC dan nilai Pseudo R^2 .

8. Membuat peta sebaran faktor-faktor yang berpengaruh pada masing-masing kabupaten/kota di Provinsi Jawa Barat.

3. Hasil dan Pembahasan

3.1 Eksplorasi Data

Menurut Badan Pusat Statistika (2018), kemiskinan merupakan suatu hal yang dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Salah satu ukuran yang dapat digunakan untuk mengetahui kondisi kemiskinan disuatu wilayah adalah persentase penduduk miskin. Peta sebaran persentase kemiskinan kabupaten/kota di Provinsi Jawa Barat tahun 2018 dapat dilihat pada Gambar 1.



Gambar 1: Peta sebaran persentase kemiskinan di Provinsi Jawa Barat

Provinsi Jawa Barat memiliki 26 kabupaten/kota yang terdiri dari 17 kabupaten dan 9 kota. Nilai persentase kemiskinan kabupaten/kota di Provinsi Jawa Barat dipetakan pada Gambar 1. Dilihat dari segi lokasi, antar wilayah kabupaten/kota memiliki persentase kemiskinan yang berbeda-beda, kabupaten/kota dengan nilai persentase kemiskinan yang hampir sama cenderung berkelompok atau memiliki lokasi yang berdekatan. Hal ini menunjukkan bahwa kemiskinan dapat dipengaruhi faktor kedekatan antar wilayah. Kemiskinan pada suatu daerah tidak hanya dipengaruhi oleh faktor-faktor kemiskinan yang dimiliki oleh daerah tersebut, akan tetapi dipengaruhi oleh kemiskinan yang dimiliki daerah lain.

Berdasarkan Gambar 1 dapat diketahui bahwa Kota Tasikmalaya dan Kabupaten Kuningan merupakan daerah yang memiliki tingkat persentase kemiskinan tinggi, sedangkan Kota Depok dan Kota Bandung merupakan daerah yang memiliki tingkat persentase kemiskinan rendah. Kota Tasikmalaya merupakan daerah dengan tingkat persentase kemiskinan tertinggi yaitu sebesar 12.71%, sedangkan Kota Depok

merupakan daerah dengan tingkat persentase kemiskinan terendah di Provinsi Jawa Barat yaitu sebesar 2.14%.

Hubungan linier antara masing-masing peubah penjelas dengan peubah respon dapat diukur menggunakan nilai korelasi *Pearson*. Tabel 2 menunjukkan bahwa terdapat peubah penjelas yang memiliki hubungan linier positif dan linier negatif terhadap peubah respon. Hubungan linier antara peubah X_3 dan X_4 terhadap peubah Y masing-masing memiliki hubungan linier positif. Oleh karena itu, semakin tinggi persentase pengeluaran perkapita untuk makanan (X_3) dan persentase rumah tangga yang pernah membeli raskin (X_4) pada suatu kabupaten/kota di Provinsi Jawa Barat, maka semakin tinggi juga persentase kemiskinan di kabupaten/kota tersebut.

Tabel 2: Nilai korelasi peubah respon dengan peubah penjelas

Peubah	Korelasi <i>Pearson</i>
Y dengan X_1	-0.3921
Y dengan X_2	-0.2689
Y dengan X_3	0.6752
Y dengan X_4	0.5911
Y dengan X_5	-0.5183
Y dengan X_6	-0.6030

Hubungan linier antara peubah X_1 , X_2 , X_5 , dan X_6 terhadap peubah Y masing-masing memiliki hubungan linier negatif. Hal ini dapat diartikan bahwa semakin tinggi produk domestik regional bruto (X_1), angka melek huruf (X_2), angka harapan lama sekolah (X_5), dan indeks kesehatan (X_6) pada suatu kabupaten/kota di Provinsi Jawa Barat, maka akan semakin rendah persentase kemiskinan di kabupaten/kota tersebut.

3.2 Pengujian Multikolinearitas

Salah satu pengujian statistik yang digunakan untuk melihat indikasi adanya korelasi linier yang beresiko atau tidak antar peubah penjelas yaitu uji asumsi multikolinieritas. Pengujian asumsi multikolinieritas antar peubah dilakukan berdasarkan nilai VIF (*Variance Inflation Factor*). Korelasi linier dengan resiko cukup tinggi antar peubah penjelas dapat diindikasikan dengan nilai VIF yang lebih besar. Jika nilai VIF yang dihasilkan lebih besar daripada sepuluh ($VIF > 10$), maka mengindikasikan terjadinya multikolinieritas antar peubah penjelas. Tabel 3 menunjukkan bahwa seluruh nilai VIF peubah penjelas setiap tahunnya bernilai kurang dari 10, sehingga dianggap tidak terjadi multikolinieritas antar peubah penjelas. Hal ini menunjukkan bahwa korelasi linier antar peubah penjelas memiliki resiko yang rendah.

Tabel 3: Nilai VIF tiap peubah penjelas tahun 2018

Peubah	X_1	X_2	X_3	X_4	X_5	X_6
VIF	1.2603	1.2556	2.5252	1.6217	1.8605	2.3404

3.3 Pemodelan Regresi Linier

Berdasarkan pendugaan model regresi linier menggunakan metode kuadrat terkecil, model regresi untuk data persentase kemiskinan di Provinsi Jawa Barat tahun 2012-

2018 diperoleh sebagai berikut:

$$y = 22.6 - 0.0197 X_1 - 0.1317 X_2 + 0.2400 X_3 + 0.0375 X_4 - 0.2584 X_5 - 0.1610 X_6$$

Tabel 4 menunjukkan dari enam peubah penjelas yang digunakan dalam penelitian, hanya tiga peubah penjelas yang berpengaruh secara signifikan terhadap persentase kemiskinan di Provinsi Jawa Barat. Hal ini dapat dilihat dari nilai-p yang kurang dari taraf nyata 5%. Peubah penjelas yang berpengaruh signifikan terhadap persentase kemiskinan di Provinsi Jawa Barat adalah produk domestik regional bruto (X_1), persentase pengeluaran perkapita untuk makanan (X_3) dan persentase rumah tangga yang pernah membeli raskin (X_4).

Tabel 4: Pendugaan parameter regresi

Peubah	Koefisien	Galat baku	t-statistik	Nilai-p
Konstanta	22.6089	13.1223	1.7230	0.0867
X_1	-0.0197	0.0094	-2.1030	0.0369*
X_2	-0.1317	0.0831	-1.5850	0.1148
X_3	0.2400	0.0509	4.7120	4.98×10^{-6} *
X_4	0.0375	0.0089	4.2060	4.14×10^{-5} *
X_5	-0.2584	0.2695	-0.9590	0.3389
X_6	-0.1610	0.1061	-1.5170	0.1310

Keterangan: *) nyata pada $\alpha = 5\%$

3.4 Pemilihan Fungsi Pembobot Optimum

Pemodelan RTG menghasilkan model regresi yang berbeda pada setiap letak geografis atau lokasi pengamatan. Pemodelan dilakukan menggunakan fungsi kernel terbaik yang dipilih berdasarkan nilai *cross validation* (CV) terkecil. Berdasarkan Tabel 5 diperoleh bahwa nilai CV terkecil dihasilkan oleh fungsi kernel eksponensial, sehingga matriks pembobot untuk pemodelan RTG dalam penelitian ini ditentukan dengan fungsi kernel eksponensial.

Tabel 5: Pemilihan pembobot fungsi kernel terbaik

Fungsi Kernel	CV RTG	Lebar Jendela RTG
<i>Gaussian</i>	297.0695	0.1387
Eksponensial*	182.8707	0.0766
<i>Bisquare</i>	490.0332	0.5226
<i>Tricube</i>	526.8850	0.5349
Adaptif <i>Gaussian</i>	624.1060	24
Adaptif Eksponensial	557.2423	24
Adaptif <i>Bisquare</i>	299.7957	24
Adaptif <i>Tricube</i>	300.9386	24

Keterangan: *) Pembobot terbaik

3.5 Pemodelan Regresi Terboboti Geografis

Regresi Terboboti Geografis (RTG) menghasilkan penduga koefisien model regresi yang bersifat lokal. Hal ini dapat diartikan setiap kabupaten/kota memiliki model regresi yang berbeda. Pendugaan parameter menggunakan RTG dapat menghasilkan nilai pendugaan yang bernilai positif ataupun negatif untuk setiap lokasi yang berbeda, sehingga peubah penjelas yang sama dapat memberikan pengaruh positif maupun

negatif terhadap persentase kemiskinan pada kabupaten/kota yang berbeda. Tabel 6 menunjukkan rentang nilai penduga parameter model RTG.

Tabel 6: Deskripsi pendugaan parameter model RTG

	Minimum	Q_1	Q_2	Q_3	Maksimum
b_0	-187.2157	38.3476	73.0374	91.4197	179.8658
b_1	-0.2937	-0.0604	-0.0218	-0.0016	0.1231
b_2	-0.7786	-0.0754	0.0211	0.1149	1.1516
b_3	-0.5727	-0.2255	-0.0698	0.0271	0.1314
b_4	-0.0805	-0.0161	-0.0017	0.0100	0.0408
b_5	-6.2897	-1.7985	-1.1588	-0.5643	0.9609
b_6	-1.8007	-0.7422	-0.4119	0.2123	1.9143

3.6 Pemilihan Model Terbaik

Berdasarkan Tabel 7, pemodelan dengan fungsi pembobot kernel eksponensial menghasilkan nilai R^2 tertinggi yaitu sebesar 0.9858. Hal ini dapat diinterpretasikan bahwa 98.58% keragaman peubah respon yang digunakan mampu dijelaskan oleh peubah penjelas, sedangkan 1.42% sisanya dijelaskan oleh faktor lain yang tidak masuk kedalam model. Jika dibandingkan dengan model lainnya, pemodelan RTG dengan fungsi pembobot kernel eksponensial memiliki nilai AIC terkecil sehingga dapat disimpulkan model yang dihasilkan merupakan model regresi terboboti geografis terbaik.

Tabel 7: Nilai R^2 dan AIC model

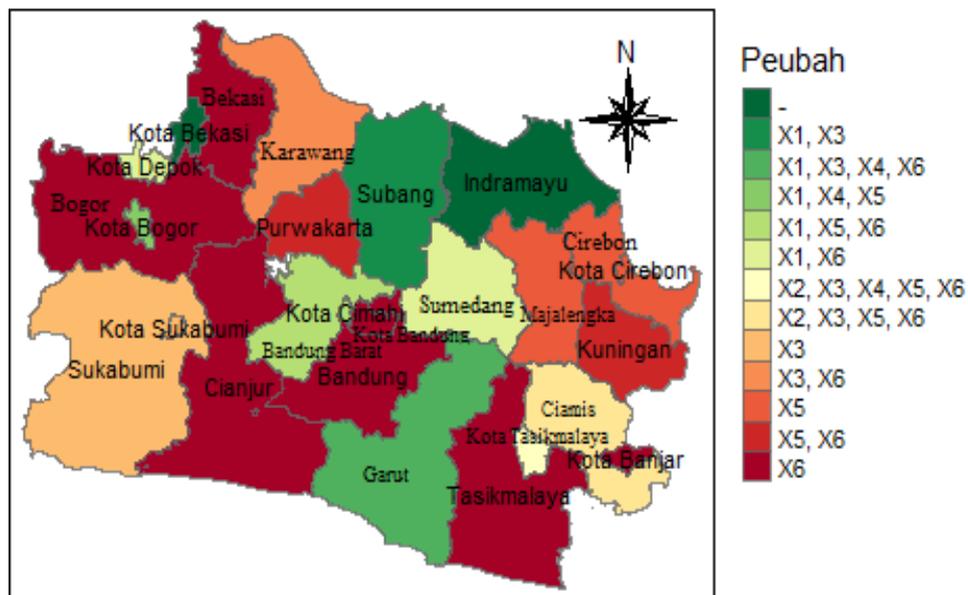
Model	AIC	R^2
Regresi Linier	823.4825	0.5615
RTG		
Gaussian	454.6848	0.9606
Eksponensial*	288.1561	0.9858
Bisquare	607.7869	0.8964
Tricube	627.9415	0.8829
Adaptif Gaussian	814.4385	0.5761
Adaptif Eksponensial	812.9876	0.5799
Adaptif Bisquare	814.3064	0.5765
Adaptif Tricube	814.4655	0.5761

Keterangan: *) Pembobot terbaik

Berdasarkan pengujian ANOVA yang telah dilakukan diperoleh nilai-p untuk pembobot kernel eksponensial kurang dari taraf nyata 5%, sehingga dapat dikatakan bahwa pemodelan menggunakan RTG dengan pembobot kernel eksponensial lebih baik digunakan dibandingkan dengan model regresi biasa dalam memodelkan peubah penjelas dengan peubah respon yang digunakan dalam penelitian ini. Hal tersebut menunjukkan bahwa besarnya pengaruh peubah penjelas pada masing-masing kabupaten/kota yang ada di Provinsi Jawa Barat secara signifikan berbeda-beda antar wilayah.

3.7 Pengujian Parameter Model RTG

Pengujian parameter model secara parsial dapat digunakan untuk mengetahui pengaruh peubah penjelas terhadap peubah respon pada setiap lokasi. Peubah-peubah penjelas dalam model RTG memiliki efek yang berbeda pada setiap daerah kabupaten/kota sehingga dapat dikatakan bahwa persentase kemiskinan di sebagian besar daerah dipengaruhi oleh faktor yang berbeda-beda. Jika faktor-faktor yang berpengaruh terhadap persentase kemiskinan tersebut dikelompokkan, maka akan terbentuk 11 kelompok peubah yang berpengaruh terhadap masing-masing wilayah. Gambaran sebaran pengaruh masing-masing peubah untuk setiap kabupaten/kota yang ada di Provinsi Jawa Barat tahun 2018 dapat dilihat pada Gambar 2.



Gambar 2: Sebaran peubah yang berpengaruh pada setiap kabupaten/kota

4. Simpulan

Provinsi Jawa Barat memiliki kabupaten/kota sebanyak 26 yang terdiri dari 17 kabupaten dan 9 kota. Dilihat dari segi lokasi, antar wilayah kabupaten/kota memiliki persentase kemiskinan yang berbeda-beda. Persentase penduduk miskin memiliki keragaman spasial sehingga dapat dilakukan pemodelan dengan regresi terboboti geografis (RTG). Fungsi kernel fixed eksponensial memiliki nilai cross validation (CV) terkecil, sehingga matriks pembobot dengan fungsi kernel eksponensial digunakan dalam pemodelan. Pemodelan dengan fungsi pembobot kernel eksponensial menghasilkan nilai R^2 tertinggi yaitu sebesar 0.9858. Hal ini dapat diinterpretasikan bahwa 98.58% keragaman peubah respon yang digunakan mampu dijelaskan oleh peubah penjelas, sedangkan 1.42% sisanya dijelaskan oleh faktor lain yang tidak masuk kedalam model. Jika dibandingkan dengan model lainnya, model RTG dengan fungsi pembobot kernel eksponensial menghasilkan nilai AIC terkecil sehingga dapat disimpulkan model yang dihasilkan merupakan model regresi terboboti geografis terbaik. Pendugaan parameter menggunakan RTG menghasilkan nilai pendugaan yang bernilai positif ataupun negatif untuk setiap lokasi yang berbeda, sehingga peubah penjelas yang sama dapat memberikan pengaruh positif maupun negatif

terhadap persentase kemiskinan pada kabupaten/kota yang berbeda. Jika faktor-faktor yang berpengaruh terhadap persentase kemiskinan tersebut dikelompokkan, maka akan terbentuk 11 kelompok peubah yang berpengaruh terhadap masing-masing wilayah

Daftar Pustaka

- Anselin, L., & Getis, A. (1992). Spatial statistical analysis and geographic information systems. *The Annals of Regional Science*, 26(1): 19–33.
- Badan Pusat Statistika. (2018). *Statistik Indonesia 2018*. Jakarta: Badan Pusat Statistika.
- Caraka, R., & Yasin, H. (2017). *Geographically Weighted Regression (GWR): Sebuah Pendekatan Regresi Geografis*. Yogyakarta: Mobius.
- Clement, F., Orange, D., Williams, M., & Mulley, C. (2009). Drivers of Afforestation in Northern Vietnam: Assessing Local Variations Using Geographically Weighted Regression. *International Journal of Applied Geography*, 29(4): 561–576.
- Fotheringham, A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.
- Purhadi, & Yasin, H. (2012). Mixed geographically weighted regression model (case study: the percentage of poor households in Mojokerto 2008). *European Journal of Scientific Research*, 69(2): 188–196.