

PEMODELAN DATA KEMATIAN BAYI DENGAN *GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION*

Riza F. Ramadhan¹, Robert Kurniawan²

^{1,2}Jurusan Komputasi Statistik, Sekolah Tinggi Ilmu Statistik (STIS), Jakarta

e-mail: ¹ 12.7353@stis.ac.id, ² robertk@stis.ac.id

DOI: 10.14710/medstat.9.2.95-106

Abstract

Overdispersion phenomenon and the influence of location or spatial aspect on data are handled using Binomial Geographically Weighted Regression (GWNBR). GWNBR is the best solution to form a regression analysis that is specific to each observation's location. The analysis resulted in parameter value which different from one observation to another between location. The Weighting Matrix Selection is done before doing The GWNBR modeling. Different weighting will resulted in different model. Thus this study aims to investigate the best fit model using infant mortality data that is produced by some kind of weighting such as fixed kernel Gaussian, fixed kernel Bisquare, adaptive kernel Gaussian and adaptive kernel Bisquare in GWNBR modeling. This region study covers all the districts/municipalities in Java because the number of observations are more numerous and have more diverse characteristics. The study shows that out of four kernel functions, infant mortality data in Java2012, the best fit model is produced by fixed kernel Gaussian function. Besides that GWNBR with fixed kernel Gaussian also shows better result than the poisson regression and negative binomial regression for data modeling on infant mortality based on the value of AIC and Deviance.

Keywords: *GWNBR, infant mortality, fixed gaussian, fixed bisquare, adaptive gaussian, adaptive bisquare.*

1. PENDAHULUAN

Indonesia melakukan upaya dalam menurunkan angka kematian bayi dan balita, dalam upaya mewujudkan sasaran pembangunan milenium (MDGs) untuk poin ke-4. Unicef Indonesia (2012) menyatakan bahwa pada tahun 1990-an Indonesia menunjukkan perkembangan dalam penurunan angka kematian balita bersama dengan komponen-komponennya, angka kematian bayi dan angka kematian bayi baru lahir. Akan tetapi, dalam beberapa tahun terakhir, penurunan angka kematian bayi (AKB) baru lahir tampak terhenti. Jika tren ini berlanjut, Indonesia berkemungkinan tidak mencapai target MDG ke empat. AKB ini merupakan indikator untuk menilai target penurunan angka kematian bayi yang dilihat dari Jumlah Kematian Bayi dibagi dengan Jumlah Kelahiran Hidup. Hubungan jumlah kematian bayi dengan faktor-faktor penyebabnya dapat diketahui salah satunya dengan metode analisis regresi.

Analisis regresi merupakan salah satu metode untuk menentukan adanya hubungan sebab akibat antara satu variabel dan variabel yang lain. Analisis regresi sendiri sangat luas pemakaiannya karena ada model pada analisis regresi yang dapat digunakan secara baik hanya pada kondisi tertentu. Salah satu contoh kondisinya ketika data variabel respon yang dijumpai adalah data cacah seperti jumlah kematian bayi dengan sebaran *Poisson* maka regresi *Poisson* menjadi cocok untuk digunakan. Namun, regresi *Poisson* mempunyai asumsi yang harus dipenuhi yaitu rata-rata variabel respon harus sama dengan variansinya yang dikenal dengan istilah equidispersi (Dobson dan Barnett, 2008). Namun, dalam kenyataannya tidak semua data cacah ini memiliki nilai rata-rata yang sama dengan variansinya. Giuffre *et al.* (2011) menyatakan bahwa yang sering terjadi pada data poisson adalah kondisi rata-rata yang lebih kecil dari variansinya atau lebih dikenal dengan istilah overdispersi.

Overdispersi dapat terjadi karena ada data yang berkelompok dalam populasi (McCullagh dan Nelder, 1989). Jika data dalam kelompok tersebut berkorelasi positif maka analisis dengan metode yang mengasumsikan kebebasan antar elemen akan menghasilkan penduga yang *underestimate* atau *varians* yang lebih kecil dari nilai sebenarnya (Astuti, 2006). Dari permasalahan overdispersi tersebut maka dibutuhkan model regresi yang lain, karena regresi *Poisson* menjadi kurang cocok digunakan untuk menganalisis data tersebut. Salah satu regresi yang tidak harus memenuhi asumsi equidispersi pada regresi *Poisson* adalah regresi *Negative Binomial* (NB). Regresi ini merupakan salah satu metode campuran *poisson-gamma* yang distribusi gamma-nya digunakan untuk mengatasi data overdispersi yang terjadi pada regresi *Poisson* (Hardin dan Hilbe 2007).

Pada dasarnya regresi *Poisson* maupun regresi *Negative Binomial* (NB) mempunyai kegunaan untuk menganalisis hubungan antara variabel respon (Y) data cacah dan satu atau lebih variabel penjelas (X). Akan tetapi, metode ini kurang representatif jika diterapkan pada data spasial atau data yang mengandung kondisi geografis (Widodo dkk, 2013). Di dalam penelitiannya Afri (2013) dijelaskan bahwa salah satu alat analisis untuk mengatasi kasus data spasial tersebut dikembangkan oleh McMillen dan McDonald tahun 1997 dan model ini dinamai *Geographically Weighted Regression* (GWR) oleh Fotheringham *et al.* (2002). Kemudian, untuk kasus penanganan fenomena overdispersi dan pengaruh lokasi atau aspek spasial pada data, penelitian selanjutnya menggunakan *Geographically Weighted Binomial Regression* (GWNBR).

GWNBR adalah salah satu solusi yang tepat untuk membentuk analisis regresi yang bersifat lokal untuk setiap lokasi pengamatan. Hasil analisisnya adalah nilai-nilai parameternya berlaku hanya pada tiap lokasi pengamatan dan berbeda dengan lokasi lainnya (Rahmawati dan Djuraidah, 2010). GWR merupakan bagian dari analisis spasial yang mempunyai pembobot berdasarkan posisi atau jarak satu lokasi dengan lokasi pengamatan lain. Unsur pembobot tersebut menentukan semakin dekat suatu lokasi, bobot pengaruhnya akan semakin besar.

Pemilihan matriks pembobot dilakukan sebelum masuk untuk pemodelan GWNBR. Berbeda dalam memilih pembobot maka berbeda pula hasil pada model yang dihasilkan. Sehingga dalam penelitian ini bertujuan menginvestigasi model terbaik yang dihasilkan oleh beberapa jenis pembobot diantaranya Kernel Tetap *Gaussian*, Kernel Tetap *Bisquare*, Kernel *Adaptive Gaussian*, dan Kernel *Adaptive Bisquare* dalam pembentukan model GWNBR untuk data Kematian Bayi. Kemudian pembobot yang menghasilkan model

terbaik dipilih untuk memodelkan data Kematian Bayi. Model terbaik adalah model yang menghasilkan nilai AIC terkecil (Widarjono, 2007) dan devian terkecil (Afri, 2013)

Afri (2013) menggunakan Model GWNBR untuk Data Kematian Bayi pada tahun 2008 di Jawa Timur. Jumlah kematian bayi adalah kejadian langka yang terjadi pada satuan waktu yang satu daerah dengan daerah lainnya memberikan pengaruh yang berbeda, sehingga variabel-variabel penelitian ini mengacu pada penelitian Afri (2013). Variabel yang didapat kemudian disesuaikan dengan kondisi data yang ada pada tahun 2012 terkait ketersediaan dan kelengkapan data. Kemudian wilayah yang dipilih untuk penelitian ini adalah kabupaten/kota di Pulau Jawa karena jumlah pengamatan lebih banyak dan karakteristik pengamatan lebih beragam.

2. KAJIAN PUSTAKA

2.1. Regresi *Poisson*

Regresi *Poisson* yang melibatkan n pasangan data, variabel respon (Y_i) dan variabel penjelas (X_i) dengan $i = 1, 2, 3, \dots, n$ memiliki hubungan antar variabel. Hubungan tersebut dinyatakan oleh Myers (1990) sebagai berikut:

$$Y_i = \mu_i + \varepsilon_i \quad (1)$$

dengan ε_i adalah galat atau error pada pengamatan ke- i dan μ_i adalah nilai harapan dari Y_i . Untuk mengetahui variabel X_i maka μ_i dinyatakan sebagai:

$$\mu_i = E(Y_i) = \exp\left(\sum_{j=0}^p \beta_j x_{ij}\right) \quad (2)$$

2.2. Regresi Binomial Negatif

Regresi binomial negatif merupakan salah satu model regresi terapan dari GLM. Fungsi padat peluang dari sebaran Binomial Negatif (Astuti, 2006) dapat ditulis sebagai berikut:

$$f(y; \mu, \varphi) = \frac{\Gamma(y + \varphi^{-1})}{\Gamma(\varphi^{-1}) y!} \left(\frac{1}{1 + \varphi\mu}\right)^{\varphi^{-1}} \left(\frac{\varphi\mu}{1 + \varphi\mu}\right)^y \quad (3)$$

dengan rata-rata dan varians Binomial Negatif dinyatakan dalam bentuk:

$$E(y) = \mu \text{ dan } Var(y) = \mu + \varphi\mu^2$$

2.3. Keragaman Spasial

Keragaman spasial adalah fenomena dalam model spasial. Karena keragaman spasial terjadi karena adanya pengaruh dari perbedaan karakteristik wilayah dan letak geografi antar wilayah pengamatan (Charlton dan Fotheringham, 2009). Variansi spasial ini dapat diketahui dengan menggunakan pengujian *Breusch-Pagan* (BP) (Anselin 1988) dengan observasinya berupa wilayah. Hipotesisnya adalah:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2 \text{ (antarlokasi mempunyai variansi yang sama)}$$

$$H_1 : \sigma_i^2 \neq \sigma^2 \text{ Minimal ada satu yang tidak sama (tidak terdapat variansi antarlokasi)}$$

Statistik Uji:

$$BP = \left(\frac{1}{2}\right) f^T Z(Z^T Z)^{-1} Z^T f \sim \chi^2_{(k)} \quad (4)$$

dengan

$$f = (f_1, f_2, \dots, f_n)^T, \quad f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right), \quad \text{dan} \quad e_i = y_i - \hat{y}_i$$

Z : Matriks dengan ukuran $n(k+1)$, berisi vektor yang sudah di normal bakukan (z) untuk setiap pengamatan

e_i^2 : Kuadrat sisaan untuk pengamatan ke- i

σ^2 : Ragam dari y

Kaidah pengambilan keputusan untuk pengujian keragaman spasial menggunakan BP, yaitu jika nilai p -value nya kurang dari tingkat signifikansi (α) 5 %, maka kesimpulannya menolak H_0 .

2.4. Fungsi Kernel

Terdapat dua jenis fungsi Kernel dalam GWR, yaitu fungsi Kernel tetap atau *Fixed Kernel* dan fungsi Kernel Adaptif atau *Adaptive Kernel* (Fotheringham *et al.*, 2002).

Fungsi *Fixed Spatial Kernel* memiliki *bandwith* (b) yang sama pada setiap titik lokasi pengamatan. Dua jenis fungsi kernel tetap yang digunakan dalam GWR adalah

1. Fungsi Kernel Tetap *Gaussian*

$$w_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right] \quad (5)$$

2. Fungsi Kernel Tetap *Bi-square*

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b}\right)^2\right]^2, & \text{jika } d_{ij} < b \\ 0, & \text{lainnya} \end{cases} \quad (6)$$

Fungsi *Adaptive Spatial Kernel* memiliki *bandwidth* yang berbeda untuk setiap titik lokasi pengamatan. Dua jenis fungsi kernel adaptif dalam GWR yaitu:

1. Fungsi Kernel Adaptif *Gaussian*

$$w_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b_{i(q)}}\right)^2\right] \quad (7)$$

2. Fungsi Kernel Adaptif *Bi-square*

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_{i(q)}} \right)^2 \right]^2, & \text{jika } d_{ij} < b \\ 0, & \text{lainnya} \end{cases} \quad (8)$$

dengan $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak *euclidean* antara titik lokasi pengamatan ke- i dan titik lokasi pengamatan ke- j (Fotheringham *et al.*, 2002).

Pemilihan *bandwidth* optimum salah satunya adalah *Akaike Information Criterion* (AIC). Menurut Akaike (1974) fungsi AIC adalah:

$$AIC = 2k - 2\log(L) \quad (9)$$

dengan k adalah banyaknya parameter penduga pada model dan L merupakan fungsi *Likelihood* dari model. Proses pemilihan lebar *bandwidth* optimum menggunakan teknik Golden Section Search. Teknik ini dilakukan secara iterasi dengan mengevaluasi nilai AIC terkecil pada interval jarak minimum dan maksimum lokasi pengamatan sehingga diperoleh nilai AIC minimum.

2.5. *Geographically Weighted Negative Binomial Regression*

Model *Geographically Weighted Negative Binomial Regression* (GWNBR) adalah salah satu metode yang cukup efektif menduga data yang memiliki spasial heterogenitas untuk data cacah yang memiliki overdispersi (Widodo dkk, 2013). Model GWNBR akan menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter berbeda. Model GWNBR dapat dirumuskan sebagai berikut:

$$y_i \square NB \left[\exp(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i)), \varphi(u_i, v_i) \right], i = 1, 2, \dots, n \quad (10)$$

Sehingga fungsi padat peluang dapat ditulis dalam bentuk persamaan berikut:

$$f(y_i; \mu_i, \varphi(u_i, v_i)) = \frac{\Gamma(y_i + \varphi^{-1})}{\Gamma(\varphi^{-1})y_i!} \left(\frac{1}{1 + \varphi\mu_i} \right)^{\varphi^{-1}} \left(\frac{\varphi\mu_i}{1 + \varphi\mu_i} \right)^{y_i} \quad (11)$$

dengan

y_i : Nilai observasi respon ke- i

μ_i : $\exp(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i))$ (rata-rata untuk setiap lokasi (u_i, v_i))

$\varphi(u_i, v_i)$: parameter dispersi untuk setiap lokasi (u_i, v_i)

3. METODOLOGI PENELITIAN

3.1. Variabel Penelitian dan Sumber Data

Variabel-variabel yang digunakan antara lain : Jumlah kematian bayi (Y), rata-rata sarana kesehatan (X1), rasio tenaga kesehatan (X2), persentase keluarga yang anggotanya menjadi buruh tani (X3), jumlah penduduk miskin (X4), jumlah sekolah negeri dan swasta

(X5), jumlah balita penderita gizi buruk (X6), jumlah bayi yang diberi ASI eksklusif (X7), rasio keluarga yang berada di pemukiman kumuh per 10000 penduduk (X8), dan Jumlah Persalinan oleh tenaga selain medis (X9).

Data yang digunakan adalah data sekunder yang diperoleh melalui Dinas Kesehatan dan Melalui Publikasi BPS pada tahun 2012 dapat dilihat di Tabel 1.

Tabel 1. Rincian Variabel Jumlah Kematian Bayi dan Faktor yang Mempengaruhinya

No	Nama Variabel	Simbol	Sumber Data
1.	Jumlah Kematian Bayi pada tiap kabupaten/kota di Pulau Jawa 2012.	Y	Profil Kesehatan 2012 Provinsi: Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Jogjakarta, Jawa Timur
2.	Rata-rata Sarana Kesehatan (Puskesmas, Poskesdes, Posyandu) per desa/kelurahan tiap kabupaten/kota di Pulau Jawa 2012	X1	Profil Kesehatan 2012 Provinsi: Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Jogjakarta, Jawa Timur
3.	Rasio Tenaga Kesehatan puskesmas dan rumah sakit (Dokter, Bidan, dan Perawat) per 10000 penduduk tiap kabupaten/kota di Pulau Jawa 2012	X2	Data Dasar 2012-Dinkes Prov.Jateng Profil Kesehatan 2012 Provinsi: Banten, DKI Jakarta, Jawa Barat, DI Jogjakarta, Jawa Timur
4.	Persentase keluarga yang anggotanya menjadi buruh tani pada tiap kabupaten/kota.	X3	SUSENAS 2012
5.	Jumlah Penduduk Miskin (dan hampir miskin) yang dicakup ASKESKIN dalam setahun terakhir pada tiap kabupaten/kota.	X4	Data dan Informasi Kemiskinan kabupaten/kota 2012
6.	Jumlah sekolah negeri dan swasta (SD, SMP, dan SMA) pada tiap kabupaten/kota.	X5	Daerah Dalam Angka 2012: Banten ,DKI Jakarta, Jawa Barat,Jawa Tengah,DI Jogjakarta,Jawa Timur.
7.	Jumlah balita penderita gizi buruk pada tiap kabupaten/kota.	X6	Profil Kesehatan 2012 Provinsi: Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Jogjakarta, Jawa Timur
8.	Jumlah bayi yang diberi ASI eksklusif pada tiap kabupaten/kota.	X7	Profil Kesehatan 2012 Provinsi: Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Jogjakarta, Jawa Timur
9.	Rasio Keluarga yang berada di Pemukiman Kumuh per 10000 penduduk pada tiap kabupaten/kota	X8	Profil Desa
10.	Jumlah Persalinan oleh tenaga selain medis pada tiap kabupaten/kota.	X9	Profil Kesehatan 2012 Provinsi: Banten, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Jogjakarta, Jawa Timur

3.2. Tahapan Penelitian

Tahapan penelitian yang digunakan untuk mencapai tujuan penelitian ini adalah sebagai berikut:

1. Memeriksa distribusi variabel respon. Variabel respon yang digunakan adalah data dengan mengikuti distribusi binomial negatif, nilai parameter dari distribusi dinyatakan dalam bentuk $E(Y) = \mu$ dan $V(Y) = \mu + \phi\mu^2$.
2. Memeriksa adanya overdispersi. Pemeriksaan overdispersi ini bisa didapat dari nilai dispersi pearson *Chi-Square* atau bisa didapat dari *deviance* yang dibagi dengan derajat bebasnya diperoleh nilai lebih besar dari 1 (Pratama dan Wulandari, 2015)
3. Memeriksa multikolinearitas. Uji Multikolinearitas digunakan untuk mengetahui ada atau tidaknya penyimpangan pada asumsi klasik yaitu adanya hubungan linear antar variabel penjelas dalam model regresi.
4. Menguji variansi spasial dengan menggunakan uji BP (*Breusch-Pagan*) . Hipotesis yang digunakan adalah variansi antar pengamatan sama. Keputusan tolak Hipotesis berarti variansi antar lokasi berbeda.
5. Menganalisis data menggunakan GWNBR. Adapun langkah dalam proses GWNBR sebagai berikut.
 - a. Menghitung matriks jarak *euclidean* antar pengamatan
 - b. Menentukan besar *bandwidth* tiap pengamatan dengan cara *Golden Section Search*.
 - d. Menguji masing-masing kernel atau matriks pembobot.
 - e. Menghitung parameter GWNBR dengan metode iterasi *Newton Raphson*.
6. Membandingkan masing-masing nilai AIC dan *Deviance* setiap kernel pada model GWNBR kemudian memilih kernel dengan AIC dan *Deviance* terkecil untuk memodelkan data kematian bayi.
7. Melakukan Interpretasi model yang didapat.

4. HASIL DAN PEMBAHASAN

4.1. Variabel Respon

Variabel respon yang dipakai adalah data cacah yaitu jumlah kematian bayi dengan nilai terendah adalah 4 dan nilai tertinggi adalah 508. Dengan rata-rata sebesar 166.69 dan *varians* 11586.64. Adanya keragaman antar wilayah menjadikan variasi yang lebih besar daripada rata-rata atau terjadi overdispersi. Dengan demikian variabel respon berdistribusi Binomial Negatif.

4.2. Memeriksa Multikolinearitas

Hasil dari uji VIF pada Tabel 2 menunjukkan bahwa variabel penjelas X9 mempunyai nilai lebih dari 10 dan variabel X5 lebih dari 5. Menurut Gujarati (1995) jika VIF lebih dari 10, terjadi multikolinearitas dan jika VIF bernilai 5 sampai dengan 10, memiliki potensi masalah. Sehingga variabel penjelas X9 dan X5 dikeluarkan dari model yang dapat dilihat Tabel 3.

Tabel 2. Nilai VIF Semua Variabel

Var	X1	X2	X3	X4	X5	X6	X7	X8	X9
VIF	1.85	2.05	3.77	3.94	9.28	2.17	2.21	1.25	10.9

Tabel 3. Nilai VIF Tanpa Variabel X5 dan X9

Var	X1	X2	X3	X4	X6	X7	X8
VIF	1.73	1.76	2.39	2.28	1.87	1.79	1.19

4.3. Memeriksa Keragaman Spasial

Hasil dari uji diperoleh nilai BP sebesar 45.741 yang nilai peluangnya adalah 9.816e-08. Nilai peluangnya kurang dari tingkat signifikansi 5%. Cara lain yaitu dengan membandingkan nilai BP dengan $\chi^2_{(0.05,7)} = 14.06714045$. Antarlokasi mempunyai variansi yang berbeda bila $p\text{-value} < \alpha$ atau BP $> \chi^2_{(\alpha, k)}$ yang artinya menolak H_0 atau variansi antar lokasi berbeda.

4.4. Geographically Weighted Negative Binomial Regression

a. Matriks Jarak

Jarak yang digunakan adalah jarak *euclidean*. Jarak ini didapat melalui variabel koordinat masing-masing pengamatan.

b. Bandwidth, AIC dan Deviance

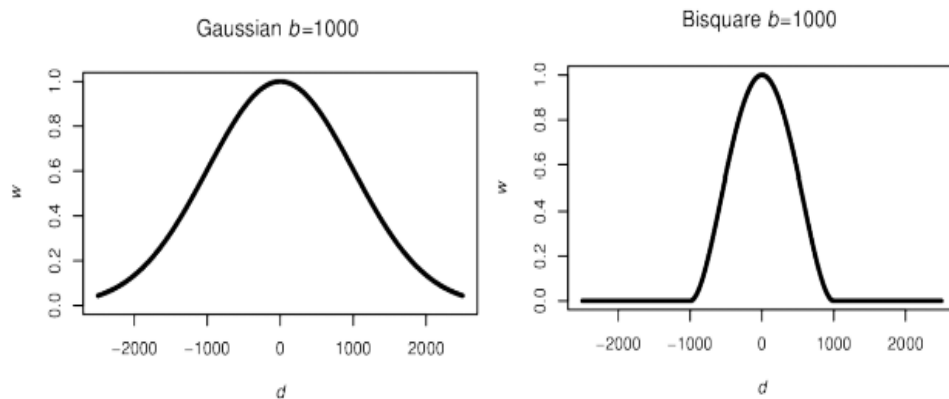
Bandwidth yang diujicobakan untuk kasus ini adalah *bandwidth* tetap dengan mengevaluasi AIC terkecil dengan teknik *golden section search*. Pada Tabel 4 nilai AIC terkecil adalah kernel Fixed Gaussian. *Bandwidth* terpilih adalah 1.754969 (dalam satuan koordinat) atau jika di konversi kira-kira sebesar 194.694 km.

Tabel 4. *Bandwidth*, AIC, dan *Deviance* untuk masing-masing Fungsi Kernel pada Model GWNBR

No	Fungsi Kernel	<i>Bandwidth</i>	AIC	<i>Deviance</i>
1	Tetap <i>Gaussian</i>	1.754969	696.5925	32.88764
2	Tetap <i>Bisquare</i>	4.240107	728.3535	33.11355
3	Adaptif <i>Gaussian</i>	3.830045 (obs 1)	877.7416	34.26246
4	Adaptif <i>Bisquare</i>	6.036699 (obs 1)	752.3044	33.27063

Hasil ini ternyata sesuai dengan penelitian Guo, Ma, dan Zhang (2013) yang menyatakan bahwa model GWR akan lebih fit dengan ukuran *bandwidth* yang lebih kecil kemudian hasil dari fungsi kernel tetap menghasilkan distribusi spasial yang lebih halus dibandingkan dengan fungsi kernel adaptif. Selain itu, penelitian ini juga mempunyai kemiripan hasil dengan penelitian dari Bidanset dan Lombard (2014) yang menyatakan bahwa fungsi tetap *gaussian* mempunyai AIC paling kecil dibandingkan dengan fungsi lainnya.

Jika dibandingkan gaussian dengan bisquare, ternyata fungsi bisquare cenderung tidak memperhatikan lagi wilayah yang berada diluar *bandwidth*-nya. Padahal belum tentu wilayah yang jauh tidak memberikan pengaruh terhadap wilayah penelitian itu sendiri. Berbeda dengan gaussian yang masih memperhatikan wilayah lain yang berada diluar *bandwidth*-nya karena *bandwidth* hanya menyatakan pengaruh besar jika wilayah lain ada didalam area *bandwidth* dan pengaruh kecil jika diluar. Hal ini bisa dilihat pada gambar berikut (Gollini *et al.*, 2015).



Gambar 1. Plot Fungsi Kernel Gaussian dan Bisquare, dengan *Bandwidth* $b = 1000$, $w =$ weight, dan d adalah jarak.

c. Matriks Pembobot Spasial

Bandwidth yang terpilih kemudian digunakan untuk mencari matriks pembobot spasial dengan menggunakan kernel tetap *Gaussian*. Hasilnya berupa matriks $n \times n$, dengan n adalah banyaknya observasi.

d. GWNBR dengan Iterasi Numerik Newton Raphson

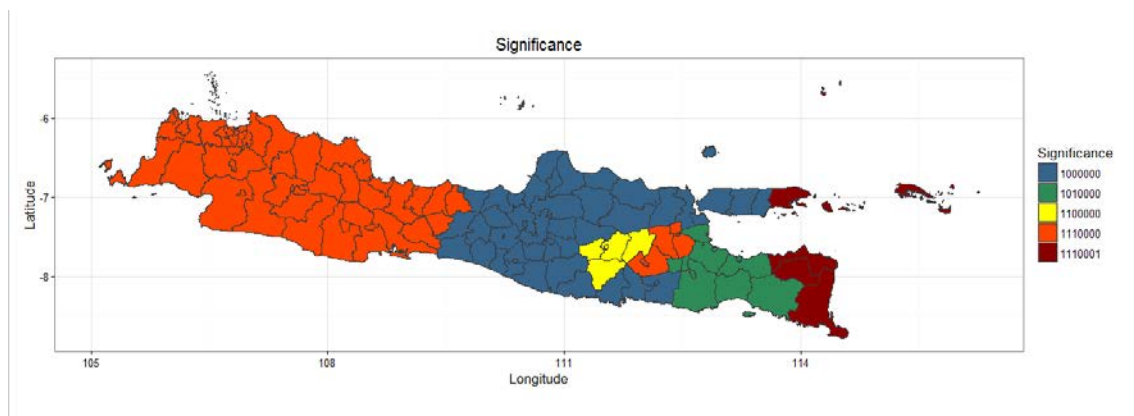
Melalui formula iterasi Newton Raphson dengan parameter awal merupakan parameter yang didapat dari regresi binomial negatif. Parameter ini kemudian diiterasi pada tiap pengamatan sampai konvergen sehingga menjadi parameter lokal pengamatan itu sendiri. Parameter yang signifikan dengan taraf nyata 5% di setiap kabupaten/kota yang telah dikelompokkan dapat dilihat pada Tabel 5.

Pengelompokan pada Tabel 5 bisa juga disajikan dalam bentuk peta. Pemetaannya terdapat pada Gambar 2. Pada gambar tersebut warna biru merupakan kelompok 1. Warna hijau merupakan kelompok 2. Warna kuning merupakan kelompok 3. Warna orange merupakan kelompok 4. Warna merah merupakan kelompok 6.

Selanjutnya hasil penilaian terhadap Model GWNBR dilakukan berdasar nilai devian dan AIC terkecil. Hasil perhitungan menunjukkan devian GWNBR senilai 32.88764 dan AIC sebesar 696.5925. Jika dibandingkan dengan model regresi *Poisson* dan model regresi Negatif Binomial, GWNBR kernel tetap Gaussian pada studi kasus kematian bayi lebih bagus. Nilai Devian dan AIC masing-masing model ada pada Tabel 6.

Tabel 5. Pengelompokan Kabupaten/Kota berdasarkan Variabel yang Signifikan dengan model GWNBR Kernel Tetap *Gaussian*

Kabupaten/Kota			Kelompok	Variabel yang Signifikan
Kab. Banjarnegara Kab. Kebumen Kab. Purworejo Kab. Wonosobo Kab. Magelang Kab. Boyolali Kab. Klaten Kab. Sukoharjo Kab. Wonogiri Kab. Karanganyar Kab. Sragen Kab. Grobogan Kab. Rembang Kab. Pati Kab. Kudus Kab. Jepara	Kab. Demak Kab. Semarang Kab. Temanggung Kab. Kendal Kab. Batang Kota Magelang Kota surakarta Kota salatiga Kota semarang Kota pekalongan Kab. Kulon progo Kab. Bantul Kab. Gunung kidul Kab. Sleman Kota Yogyakarta	Kab. Pacitan Kab. Trenggalek Kab. Tulungagung Kab. Blitar Kab. Blora Kab. Ngawi Kab. Bojonegoro Kab. Tuban Kab. Lamongan Kab. Gresik Kab. Bangkalan Kab. Sampang Kab. Pamekasan Kota Blitar Kota Surabaya	1	X1
Kab. Malang Kab. Lumajang Kab. Jember Kab. Probolinggo	Kab. Pasuruan Kab. Sidoarjo Kota Malang Kota Probolinggo	Kota Pasuruan Kota Mojokerto Kota Batu	2	X1,X3
Kab. Ponorogo Kab. Nganjuk	Kab. Madiun Kab. Magetan	Kota Kediri Kota Madiun	3	X1,X2
Kab. Kep. Seribu Kota Jakarta selatan Kota Jakarta timur Kota Jakarta pusat Kota Jakarta barat Kota Jakarta utara Kab. Bogor Kab. Sukabumi Kab. Cianjur Kab. Bandung Kab. Garut Kab. Tasikmalaya Kab. Ciamis Kab. Kuningan Kab. Cirebon Kab. Majalengka Kab. Sumedang	Kab. Indramayu Kab. Subang Kab. Purwakarta Kab. Karawang Kab. Bekasi Kab. Bandung Barat Kota Bogor Kota Sukabumi Kota Bandung Kota Cirebon Kota Bekasi Kota Depok Kota Cimahi Kota Tasikmalaya Kota Banjar Kab. Cilacap Kab. Banyumas	Kab. Purbalingga Kab. Pekalongan Kab. Pematang Kab. Tegal Kab. Brebes Kota Tegal Kab. Kediri Kab. Mojokerto Kab. Jombang Kab. Pandeglang Kab. Lebak Kab. Tangerang Kab. Serang Kota Tangerang Kota Cilegon Kota Serang Kota Tangerang Slt	4	X1,X2,X3
Kab. Banyuwangi Kab. Bondowoso	Kab. Situbondo	Kab. Sumenep	5	X1,X2,X3,X7



Gambar 2. Peta Hasil Pengelompokan Kabupaten/Kota berdasarkan Variabel yang signifikan

Tabel 6. Perbandingan AIC dan Devians Regresi *Poisson*, Regresi Binomial Negatif dan GWNBR

Model Regresi	AIC	Deviance
Regresi <i>Poisson</i>	5803.768	4998.129
Regresi Binomial Negatif	1369.409	124.4262
GWNBR Kernel Tetap <i>Gaussian</i>	696.5925	32.88764

5. KESIMPULAN

Setelah membandingkan keempat fungsi pembobot, model terbaik berdasarkan nilai AIC dan *Deviance* dihasilkan ketika pembobot yang digunakan untuk data kematian bayi di Pulau Jawa 2012 adalah Kernel tetap *Gaussian*. Kemudian GWNBR kernel tetap *Gaussian* juga memberikan hasil yang lebih baik dibanding regresi *Poisson* dan regresi Negatif Binomial untuk pemodelan data kematian bayi berdasarkan nilai AIC dan *Deviance*.

DAFTAR PUSTAKA

- Afri, L. E., 2013, Model Regresi Binomial Negatif Terboboti Geografis Untuk Data Kematian Bayi. *Jurnal Edu Research*, 2(1), pp 15-26.
- Akaike, H., 1974, A New Look at The Statistical Model Identification. *Automatic Control, IEEE Transactions on*, 19(6), pp 716-723.
- Anselin L., 1988, *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Astuti, E. T., 2006, Uji Trend/Regresi untuk Data Over Dispersi, *Jurnal Statistika Ikatan Perstatistikan Indonesia*, 2, pp 179-186.
- Bidanset, P. E., dan Lombard, J. R., 2014, The Effect of Kernel and *Bandwidth* Specification in Geographically Weighted Regression Models on The Accuracy and Uniformity of Mass Real Estate Appraisal, *Journal of Property Tax Assessment & Administration*, 11(3), 5.

- Charlton, M., dan Fotheringham, A. S., 2009, Geographically Weighted Regression. White paper, National Centre for Geocomputation National University of Ireland Maynooth.
- Dobson, A. J., dan Barnett, A., 2008, *An Introduction to Generalized Linear Models*. CRC press.
- Fotheringham, A. S., Brunson, C., dan Charlton, M., 2002, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons.
- Giuffrè, O., Granà, A., Roberta, M., dan Corriere, F., 2011, Handling Underdispersion in Calibrating Safety Performance Function at Urban, Four-Leg, Signalized Intersections. *Journal of Transportation Safety & Security*, 3(3), pp 174-188.
- Chiu, S.L., 1994, Fuzzy Model Identification Based on Cluster Estimation, *Journal of Intelligent and Fuzzy Systems*, Vol. 2, pp 267-278.
- Gollini, I., Lu, B., Charlton, M., Brunson, C., and Harris, P., 2015, GW Model: an R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *arXiv preprint arXiv:1306.0413*.
- Gujarati, N. D., 1995, *Basic Econometrics* 3rd Edition. McGraw Hill, New York.
- Guo, L., Ma, Z., and Zhang, L., 2008, Comparison of *Bandwidth* Selection in Application of Geographically Weighted Regression: a Case Study. *Canadian Journal of Forest Research*, 38(9), pp 2526-2534.
- Hardin, J. W., Hilbe, J. M., dan Hilbe, J., 2007, *Generalized Linear Models and Extensions*. Stata Press.
- McCullagh, P., dan J. A. Nelder, 1989, *Generalized Linear Models*, 2nd Ed., Chapman and Hall, New York.
- Myers RH., 1990, *Classical and Modern Regression with Applications Second Edition*, PWS-KENT, New York.
- Pratama, W., dan Wulandari, S. P., 2015, Pemetaan Dan Pemodelan Jumlah Kasus Penyakit Tuberculosis (TBC) di Provinsi Jawa Barat Dengan Pendekatan Geographically Weighted Negative Binomial Regression (GWNBR). *Jurnal Sains dan Seni ITS*, 4(1), D37-D42.
- Rahmawati, R., dan Djuraidah, A., 2010, Regresi Terboboti Geografis dengan Pembobot Kernel Kuadrat Ganda untuk Data Kemiskinan di Kabupaten Jember. In *Forum Statistika dan Komputasi*, Vol. 15, No. 2.
- UNICEF Indonesia, 2012, Kesehatan Ibu & Anak. Jakarta: UNICEF Indonesia.
- Widarjono, A., 2007, *Ekonometrika Teori dan Aplikasi untuk Ekonomi dan Bisnis*, Edisi Kedua, Penerbit Ekonisia Fakultas Ekonomi UII, Yogyakarta.
- Widodo. T.C, Sunaryo. S, dan Purhadi, 2013, *Pemodelan Spasial Balita Gizi buruk dengan Geographically Weighted Negative Binomial Regression dan Flexibly Shaped Spatial Scan Statistic*. Surabaya: ITS [Thesis]. Tidak dipublikasikan.