

KAJIAN SIMULASI OVERDISPERSI PADA REGRESI POISSON DAN BINOMIAL NEGATIF TERBOBOTI GEOGRAFIS UNTUK DATA BALITA GIZI BURUK*

Puput Cahya Ambarwati¹, Indahwati^{2‡}, and Muhammad Nur Aidi³

¹Department of Statistics, IPB University, Indonesia, puputcahya.ambarwati15@gmail.com

²Department of Statistics, IPB University, Indonesia, indah.stk@gmail.com

³Department of Statistics, IPB University, Indonesia, nuraidi@yahoo.com

‡corresponding author

Indonesian Journal of Statistics and Its Applications (eISSN:2599-0802)

Vol 4 No 3 (2020), 484 - 497

Copyright © 2020 Puput Cahya Ambarwati, Indahwati, 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

One type of geographically weighted regression (GWR) that can be used to explain the relationship between the response variables in the form of count data and explanatory variables is the geographically weighted Poisson regression (GWPR). In the GWPR, there is an assumption that should be fulfilled called equidispersion, a condition where the variance equals the mean. If that condition is ignored, overdispersion will occur. Overdispersion is a condition when the variance is greater than the mean. The use of GWPR analysis in an overdispersion situation will produce a smaller standard error than it should be (underestimate). This may produce a significant test result leading to the rejection of the null hypothesis. One of the classic approaches commonly used to handle overdispersion in GWR is geographically weighted negative binomial regression (GWNBR). GWNBR is derived from a mixture of Poisson and Gamma distributions which is similar to the negative binomial distribution. Simulation data and real data were used in this study. The results showed that the application of GWPR on overdispersion data could increase the number of rejections of H_0 or the number of p-values. The application of GWNBR on the East Java malnutrition toddler data in 2017 showed that the GWNBR model is better than GWPR based on the comparison of AIC, Pseudo R^2 , and RMSE.

Keywords: geographically weighted poisson regression, geographically weighted negative binomial regression, overdispersion.

* Received Jul 2020; Accepted Nov 2020; Published online on Nov 2020

1. Pendahuluan

Data cacah merupakan data yang diperoleh dari hasil mencacah atau menghitung. Salah satu model regresi yang dapat digunakan untuk menjelaskan hubungan antara peubah respon yang berupa data cacah dengan peubah penjelas adalah model regresi Poisson (Agresti, 2003). Dalam model regresi Poisson terdapat asumsi yang harus dipenuhi yaitu kondisi ragam dan rata-rata sama atau *equidispersi*. Jika kondisi ini diabaikan maka akan terjadi *overdispersi*. *Overdispersi* adalah kondisi ragam lebih besar daripada rata-rata (McCullagh & Nelder, 1989). Penggunaan analisis regresi Poisson jika terjadi *overdispersi* akan menghasilkan galat baku yang terlalu kecil daripada seharusnya (*underestimate*). Hal ini akan menghasilkan uji signifikansi peubah bebas cenderung menolak hipotesis. Salah satu pendekatan klasik yang biasa digunakan untuk mengatasi *overdispersi* berkaitan dengan sebaran Poisson yaitu pendekatan yang diperoleh dari turunan sebaran Poisson dan gamma yang mirip dengan fungsi sebaran binomial negatif.

Data spasial merupakan data hasil pengamatan yang mempunyai informasi tentang koordinat lokasi di mana data dikumpulkan selain informasi tentang peubah yang sedang diamati. Salah satu model yang dapat mengatasi keragaman dan ketergantungan spasial yaitu *geographically weighted regression* (GWR) atau regresi terboboti geografis (RTG). RTG merupakan model regresi linier dengan pendekatan titik. Model ini akan menghasilkan parameter model yang bersifat lokal untuk setiap lokasi dimana data itu dikumpulkan (Fotheringham *et al.*, 2003). RTG merupakan pemodelan untuk peubah respon menyebar normal dengan data kontinu. Namun, pada kenyataannya peubah respon seringkali merupakan data cacah. Menurut Nakaya *et al.* (2005) pemodelan RTG dengan peubah respon mengikuti sebaran Poisson dapat menggunakan pendekatan regresi Poisson terboboti geografis (RPTG). Namun, pemodelan RPTG sering tidak memenuhi asumsi *overdispersi* seperti pada regresi Poisson. Sama halnya pada regresi Poisson, pendekatan untuk mengatasi *overdispersi* pada data cacah dengan keragaman spasial yaitu pendekatan regresi binomial negatif terboboti geografis (RBNTG) (da Silva & Rodrigues, 2014).

Penanganan *overdispersi* menggunakan RBNTG untuk data cacah dengan keragaman spasial diduga memerlukan waktu komputasi yang cukup lama jika dibandingkan RPTG. Hal ini dikarenakan parameter yang diduga lebih banyak yaitu adanya tambahan parameter dispersi untuk setiap lokasi pada RBNTG. Oleh karena itu perlu dikaji sampai sejauh mana RPTG masih tepat digunakan untuk data cacah dengan keragaman spasial. Simulasi yang akan dilakukan yaitu membuat beberapa kondisi *overdispersi* yaitu mendekati 0 atau tidak ada *overdispersi*, *overdispersi* mendekati 1 atau kecil dan *overdispersi* menjauhi 1 atau besar untuk melihat batas toleransi *overdispersi* yang masih tepat dimodelkan menggunakan RPTG.

Salah satu indikator umum yang dapat menggambarkan tingkat kesehatan masyarakat yaitu jumlah gizi buruk pada balita. Jawa Timur merupakan salah satu dari 4 provinsi yang selama 5 tahun berturut-turut termasuk ke dalam kategori provinsi dengan kasus gizi buruk tertinggi. Penelitian tentang gizi buruk yang mempertimbangkan aspek lokasi telah dibahas oleh Sofia (2018). Penelitian tersebut menghasilkan kesimpulan bahwa peubah yang berpengaruh nyata terhadap gizi buruk

antara lain jumlah bayi dengan berat lahir rendah (BBLR), jumlah rumah tangga dengan kategori rumah sehat, jumlah rumah tangga dengan penggunaan air bersih, jumlah posyandu aktif, jumlah balita yang mendapatkan vitamin A, dan jumlah bayi dengan pemberian air susu ibu (ASI) eksklusif. Miranti & Purhadi (2016) mengatakan faktor-faktor yang berpengaruh nyata terhadap kejadian balita gizi buruk adalah persentase BBLR, persentase balita mendapat ASI dan persentase rumah tangga miskin dengan model RBNTG. Penelitian ini ingin mengetahui toleransi overdispersi yang masih tepat dimodelkan menggunakan RPTG serta ingin menerapkan model RPTG dan RBNTG pada data balita gizi buruk di Jawa Timur pada tahun 2017 menggunakan BBLR, pemberian ASI eksklusif, jumlah posyandu, pemberian vitamin A, akses air bersih dan rumah memenuhi syarat rumah sehat sebagai peubah penjelas.

2. Metodologi

Data yang digunakan dalam penelitian ini merupakan data simulasi dan data nyata (*real*). Data simulasi merupakan data bangkitan menggunakan *software* R. Data nyata (*real*) pada penelitian ini merupakan data jumlah balita gizi buruk di Jawa Timur pada tahun 2017. Tahapan analisis dalam penelitian ini sesuai data yang digunakan yaitu tahapan simulasi dan tahapan penerapan pada data nyata (*real*).

2.1 Data Simulasi

Pembangkitan data simulasi pada penelitian ini merupakan pengembangan dari simulasi yang telah dilakukan oleh Liu *et al.* (2017). Lokasi pengamatan terdiri titik $m \times m$ dengan jarak $l = 6 / (m - 1)$. Dalam hal ini $m = 7$, dan ukuran sampel $n = 49$ pengamatan. Langkah-langkah pembangkitan data sebagai berikut:

1. Menentukan koordinat lokasi (u_i, v_i) di mana pengamatan yang dihitung dapat dinyatakan sebagai berikut:

$$(u_i, v_i) = \left(\text{mod}(i - 1, m), \text{mod} \left(\text{int} \left(\frac{i-1}{m} \right), m \right) \right)$$

$$(i = 1, 2, \dots, m^2)$$

di mana $\text{mod}(i - 1, m)$ adalah sisa dari $i - 1$ dibagi dengan m

$\text{int} \left(\frac{i-1}{m} \right)$ adalah nilai integer dari $\text{int} \left(\frac{i-1}{m} \right)$.

2. Membuat data set yang mempunyai keragaman spasial dengan inisialisasi awal seperti berikut:

$$\beta_{0_i} = 3 + (u_i + v_i), \beta_{1_i} = \frac{1}{2} [3(2 - u_i)], \beta_{2_i} = 2 [9 - (6 - v_i)]$$

3. Membangkitkan peubah bebas X_1 dan X_2 yang mengikuti sebaran seragam yang berkisar antara $(0, 100)$.

4. Membangkitkan Y mengikuti $y_i \sim \text{Binomial negatif}$

$$\left[\exp(\beta_0(u_i, v_i) + \sum_{j=1}^p \beta_j(u_i, v_i) x_{ij}), \theta(u_i, v_i) x_{ij} \right] .$$
5. Mengulang langkah 4 sebanyak 100 kali.

2.2 Data Balita Gizi Buruk

Data jumlah gizi buruk pada balita berasal dari buku Profil Kesehatan Jawa Timur Tahun 2017 yang diterbitkan oleh Dinas Kesehatan Jawa Timur dan buku Saku Pemantauan Status Gizi Tahun 2017 yang diterbitkan oleh Kementerian Kesehatan Republik Indonesia. Dalam penelitian ini, terdapat 38 kabupaten/kota sebagai lokasi pengamatan. Peubah penelitian yang digunakan adalah sebagai berikut :

Tabel 1: Peubah Penelitian

Kode	Peubah
Y	Jumlah gizi buruk pada balita
X_1	Persentase BBLR
X_2	Persentase balita dengan pemberian ASI eksklusif
X_3	Persentase balita yang mendapatkan vitamin A
X_4	Persentase posyandu aktif
X_5	Persentase rumah tangga dengan akses air bersih
X_6	Persentase rumah memenuhi syarat rumah sehat

2.3 Analisis Data Simulasi

Langkah-langkah analisis yang dilakukan untuk data simulasi adalah sebagai berikut:

1. Membuat kondisi equidispersi yaitu θ mendekati 0 dan overdispersi dengan 2 kondisi yaitu mendekati 1 dan menjauhi 1 sebagai berikut:

$$\theta_1(u_i, v_i) = 0.000001^{(v_i+1)}$$

$$\theta_2(u_i, v_i) = 1.11^{(v_i+1)}$$

$$\theta_3(u_i, v_i) = 3.61 + 1.19^{(v_i)}$$

Ketiga inisialisasi θ menggunakan fungsi v_i dibuat sesuai tujuan peneliti untuk menghasilkan θ yang bernilai mendekati 0, mendekati 1 (rentang 1-2) dan menjauhi 1 (rentang 4-7).

2. Memodelkan menggunakan RPTG dan RBNTG sebagai kontrol.
3. Mengkaji toleransi overdispersi untuk masing-masing kondisi berdasarkan jumlah p -value yang signifikan.

2.4 Analisis Data Balita Gizi Buruk

Langkah-langkah analisis yang dilakukan untuk data balita gizi buruk adalah sebagai berikut:

1. Mengeksplorasi data balita gizi buruk beserta faktor faktor yang memengaruhi.
2. Menentukan lebar jendela optimum
 - a. Menentukan fungsi pembobot dengan menggunakan pembobot *adaptive bisquare*.
 - b. Menghitung matriks pembobot $W(u_i, v_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in})$
 - c. Menghitung $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$
 - d. Meminimumkan nilai validasi silang berdasarkan h dengan rumus:

$$\text{Cross Validation}(CV) \text{ atau } (h) = \sum_{i=1}^n [y_{\neq i} - \hat{y}_{\neq i}(h)]^2$$

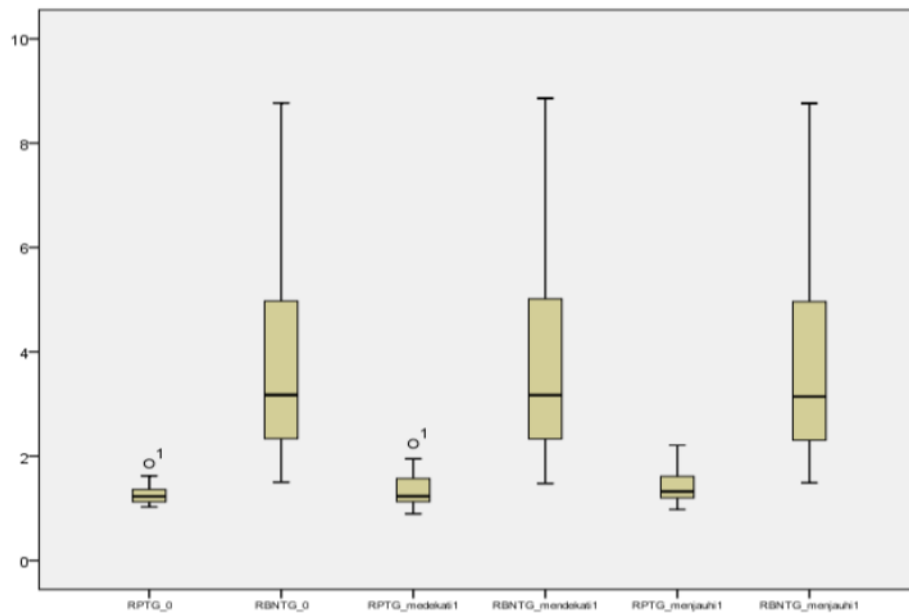
3. Menduga parameter menggunakan model RPTG dan RBNTG.

3. Hasil dan Pembahasan

3.1 Hasil Kajian Data Simulasi

Pada simulasi ini, data yang dibangkitkan dibuat mirip seperti struktur data RTG yang mempunyai overdispersi dan keragaman spasial. Simulasi yang dilakukan mempunyai 49 lokasi yang mempunyai koordinat longitudinal dan latitudinal. 3 kondisi yang dibuat yaitu mendekati 0 (rentang 1e-06 sampai 1e-42), mendekati 1 (rentang 1.110 sampai 2.076) dan menjauhi 1 (rentang 4.610 sampai 7.450). Jumlah peubah penjelas yang dibangkitkan yaitu 2 peubah penjelas yang mengikuti sebaran seragam dengan rentang 0 sampai dengan 100. Lama waktu komputasi untuk RPTG kurang lebih 5 menit sedangkan untuk RBNTG 1 jam untuk setiap kondisi overdispersi. Lama waktu komputasi RBNTG dikarenakan RBNTG menduga parameter overdispersi untuk setiap lokasi. Jumlah ulangan pada simulasi ini sebanyak 100 kali ulangan.

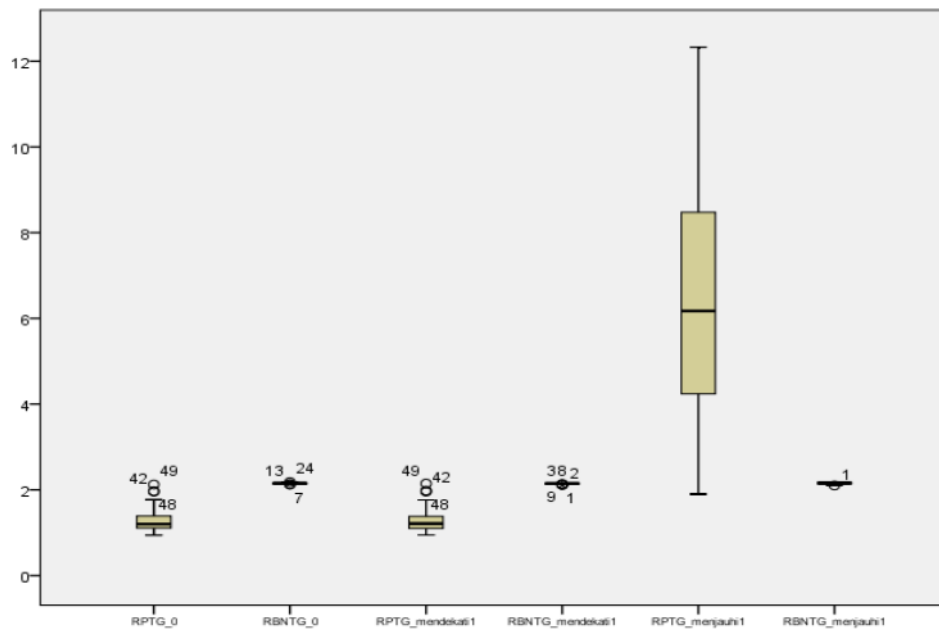
RMSE merupakan salah satu kriteria untuk melihat kebaikan model. Semakin kecil nilai RMSE maka model akan dianggap lebih baik. Gambar 1 menunjukkan perbandingan rata-rata RMSE beta 0 untuk 3 kondisi overdispersi pada model RPTG dan RBNTG. Untuk model RPTG, rata-rata RMSE ketika overdispersi mendekati 0 menunjukkan nilai terkecil jika dibandingkan dengan kondisi lainnya. Rata-rata RMSE pada 2 kondisi yaitu overdispersi mendekati 0 dan mendekati 1 mempunyai rata-rata yang hampir sama. Berdasarkan perhitungan, rata-rata RMSE pada kondisi overdispersi mendekati 0 mempunyai rata-rata RMSE yang lebih kecil yaitu 1.348 dari pada *overdispersi* mendekati 1 yaitu 1.398. Selisih dari 2 kondisi di atas sebesar 0.050. Sedangkan rata-rata RMSE pada kondisi overdispersi menjauhi 1 mempunyai nilai 1.254 dan merupakan nilai RMSE yang terbesar daripada 2 kondisi lainnya. Untuk model RBNTG, rata-rata RMSE 3 kondisi overdispersi menunjukkan yang hampir sama yaitu 3.749, 3.718 dan 3.735. Rata-rata RMSE beta 0, 3 kondisi untuk model RPTG lebih kecil jika dibandingkan dengan model RBNTG



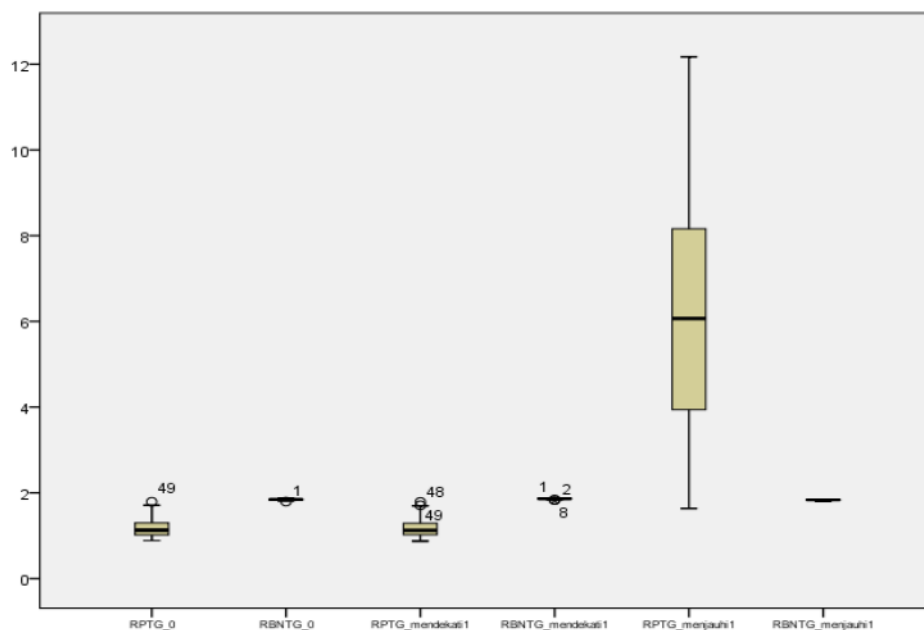
Gambar 1: Perbandingan boxplot rata-rata RMSE beta 0

Gambar 2 menunjukkan perbandingan rata-rata RMSE beta 1 yang hampir sama pada 2 kondisi yaitu overdispersi mendekati 0 dan overdispersi mendekati 1 untuk model RPTG. Berdasarkan perhitungan, rata-rata RMSE pada kondisi overdispersi mendekati 0 mempunyai rata-rata RMSE yang lebih kecil yaitu 1.278 dari pada overdispersi mendekati 1 yaitu 1.285. Sedangkan rata-rata RMSE pada kondisi overdispersi menjauhi 1 mempunyai nilai 6.341 dan merupakan yang terbesar daripada 2 kondisi lainnya. Untuk model RBNTG, rata-rata RMSE 3 kondisi overdispersi menunjukkan yang hampir sama yaitu 2.144, 2.151 dan 2.153. Rata-rata RMSE pada kondisi overdispersi mendekati 0 mempunyai rata-rata RMSE yang lebih kecil dari pada overdispersi mendekati 1 yaitu 2.151 dan overdispersi menjauhi 1 yaitu 2.153. Dari semua kondisi, overdispersi menjauhi 1 untuk model RPTG mempunyai nilai yang paling besar untuk rata-rata RMSE beta 1.

Gambar 3 menunjukkan rata-rata RMSE beta 2 yang hampir sama pada 2 kondisi yaitu overdispersi mendekati 0 dan overdispersi mendekati 1. Berdasarkan perhitungan, rata-rata RMSE pada kondisi overdispersi mendekati 0 mempunyai rata-rata RMSE yang lebih kecil yaitu 1.185 dari pada overdispersi mendekati 1 yaitu 1.189. Selisih dari 2 kondisi ini sangat sedikit yaitu sebesar 0.004. Sedangkan rata-rata RMSE pada kondisi overdispersi menjauhi 1 mempunyai nilai 6.130 dan merupakan yang terbesar daripada 2 kondisi lainnya. Untuk model RBNTG, rata-rata RMSE 3 kondisi overdispersi menunjukkan yang hampir sama yaitu 1.859, 1.831 dan 1.844. Rata-rata RMSE model RPTG pada kondisi overdispersi menjauhi 1 mempunyai nilai terbesar daripada kondisi lainnya.



Gambar 2: Perbandingan boxplot rata-rata RMSE beta 1



Gambar 3: Perbandingan boxplot rata-rata RMSE beta 2

Overdispersi adalah kondisi ketika nilai ragam lebih besar daripada rata-rata. Efek dari overdispersi adalah semakin besarnya peluang menerima H_1 atau menolak H_0 jika tetap dimodelkan menggunakan RPTG.. Semakin banyak p -value yang signifikan menandakan model semakin kurang baik. Tabel merupakan jumlah penolakan H_0 atau signifikansi pada masing-masing kondisi yaitu overdispersi mendekati 0, overdispersi mendekati 1 dan overdispersi menjauhi 1 untuk model RPTG.

Tabel 2: Perbandingan jumlah dan rata-rata p -value yang signifikan pada pemodelan RPTG dan RBNTG.

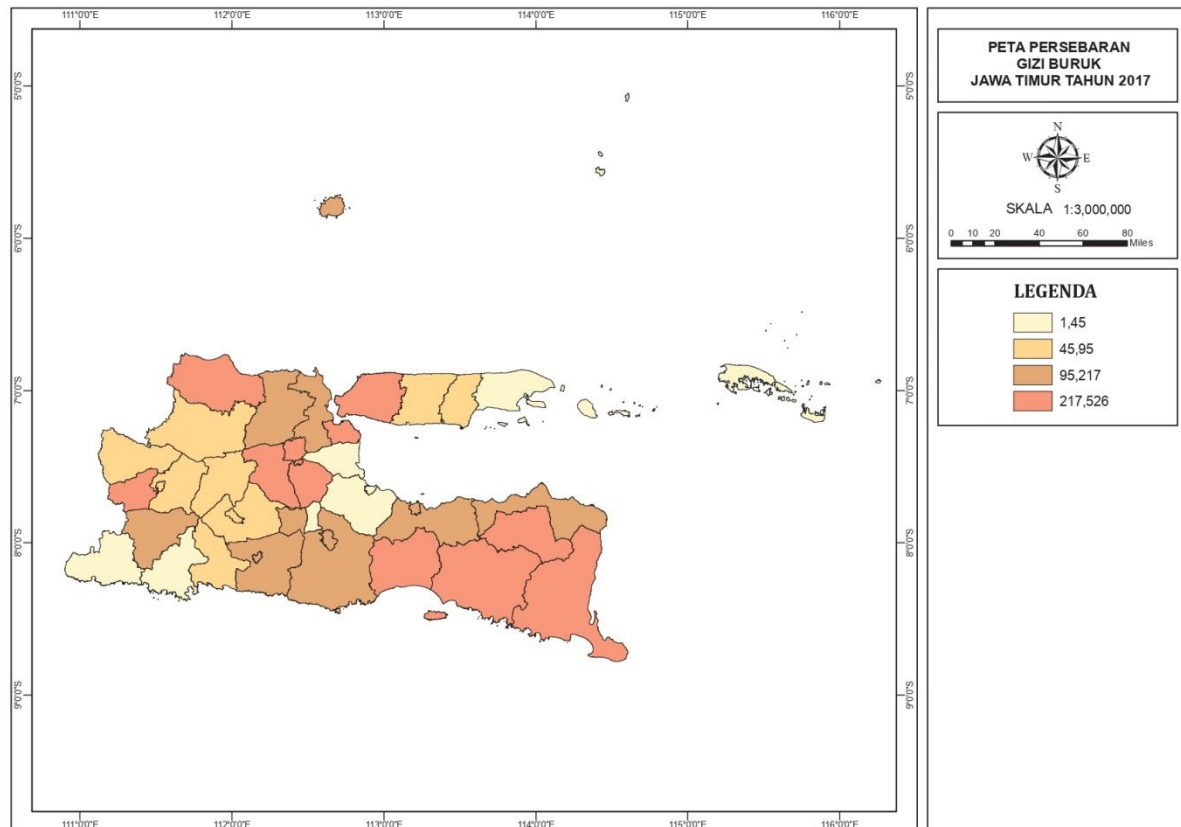
Model	Mendekati 0			Mendekati 1			Menjauhi 1		
	Beta0	Beta1	Beta2	Beta0	Beta1	Beta2	Beta0	Beta1	Beta2
Jumlah RPTG	2624	1776	1214	4888	4323	4251	4887	4327	4255
Jumlah RBNTG	2714	2233	1652	2764	2272	1776	2860	2363	1855
Rata-rata (%) RPTG	53.55	36.24	24.77	99.75	88.22	86.75	99.73	88.31	86.84
Rata-rata (%) RBNTG	55.39	45.57	33.71	56.41	46.37	36.25	58.37	48.22	37.86

Berdasarkan Tabel 2 jumlah dan rata-rata p -value yang signifikan model RPTG pada saat overdispersi mendekati 0 mempunyai nilai yang lebih kecil jika dibandingkan dengan kondisi overdispersi mendekati 1 dan menjauhi 1. Hal ini menunjukkan bahwa RPTG tepat jika digunakan untuk data tanpa overdispersi. Pada saat overdispersi mendekati 1 dan menjauhi 1, jumlah dan rata-rata p -value pada saat RPTG lebih besar daripada RBNTG. Hal ini menunjukkan bahwa RPTG kurang tepat untuk memodelkan data dengan overdispersi. Jumlah dan rata-rata p -value yang signifikan model RBNTG lebih kecil dari pada model RPTG pada 2 kondisi yaitu overdispersi mendekati 1 maupun menjauhi 1. Selisih jumlah maupun rata-rata p -value pada saat overdispersi mendekati 1 dan menjauhi 1 model RPTG dengan model RBNTG mempunyai nilai yang jauh antara rentang 1960 sampai 2479 atau 40% sampai 50.59%. Besarnya jumlah dan rata-rata overdispersi pada model RPTG pada kondisi overdispersi mendekati dan menjauhi 1 menunjukkan bahwa terdapat kecenderungan efek overdispersi yaitu sering terjadinya penolakan H_0 . Pada model RBNTG kondisi overdispersi mendekati 0, jumlah dan rata-rata p -value lebih besar jika dibandingkan model RPTG. Hal ini menunjukkan RBNTG memang tepat untuk data overdispersi. Jumlah dan rata-rata p -value pada model RBNTG untuk kondisi overdispersi menjauhi 1 mempunyai nilai yang lebih besar jika dibandingkan dengan kondisi overdispersi mendekati 1. Hal ini menunjukkan bahwa batas toleransi overdispersi yang masih tepat digunakan yaitu overdispersi mendekati 1.

3.2 Hasil Kajian Data Balita Gizi Buruk

Status gizi masyarakat merupakan indikator kesehatan yang dapat menggambarkan tingkat kesehatan masyarakat di suatu wilayah. Salah satu indikator umumnya yaitu jumlah gizi buruk pada balita. Usia balita merupakan usia keemasan karena 1000 hari

pertama merupakan masa balita tumbuh dan berkembang. Kekurangan gizi pada balita dapat mengakibatkan terganggunya pertumbuhan dan perkembangan fisik, mental, dan spiritual sehingga akan mengakibatkan rendahnya kualitas sumber daya manusia. Jawa Timur merupakan salah satu 4 provinsi yang selama 5 tahun berturut-turut termasuk ke dalam kategori provinsi dengan kasus gizi buruk tertinggi.



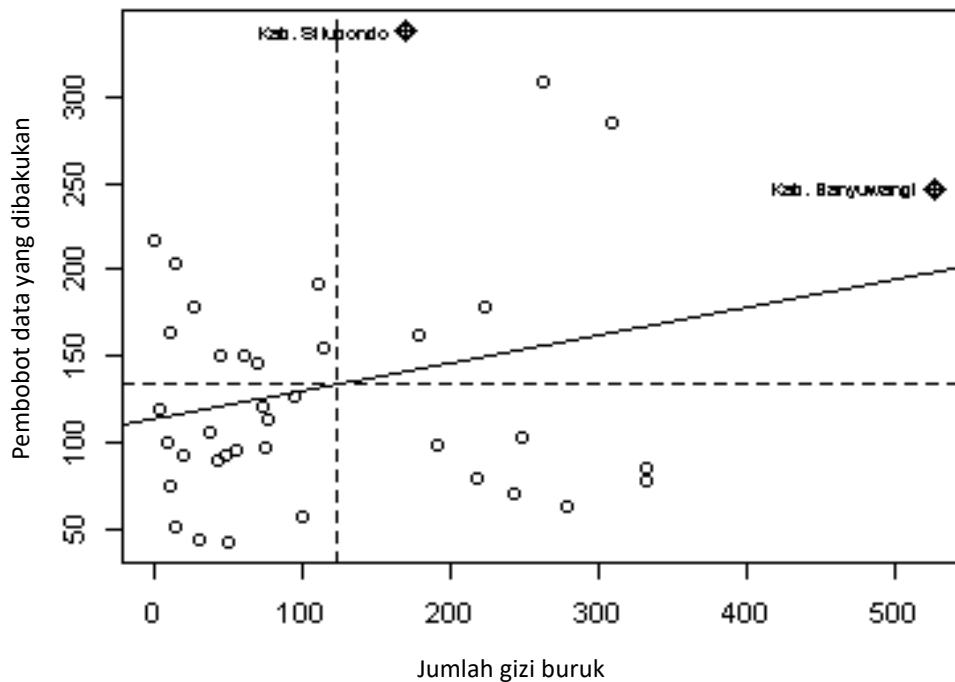
Gambar 4: Persebaran gizi buruk Jawa Timur tahun 2017

Gambar 4 menunjukkan bahwa terdapat beberapa kabupaten/kota yang berdekatan cenderung menggerombol ditandai dengan persamaan warna pada peta. Pada 4 kabupaten yang terletak di ujung timur Provinsi Jawa Timur mempunyai warna merah yang artinya jumlah gizi buruk pada tahun 2017 di 4 kabupaten tersebut relatif sama dan tinggi dibandingkan dengan kabupaten/kota yang lain. 5 kabupaten/kota di ujung barat Provinsi Jawa Timur juga cenderung menggerombol dengan ditandai warna kuning. Kabupaten/kota tersebut mempunyai rentang jumlah gizi buruk pada tahun 2017 antara 45-95 balita. Penggerombolan pada data balita gizi buruk Jawa Timur tahun 2017 mengindikasikan bahwa data menyebar secara binomial negatif.

3.2.1 Pengujian Sebaran Data dan Asumsi Spasial

Pengujian sebaran dilakukan untuk mengetahui sebaran dari data. Uji yang dilakukan

yaitu uji Kolmogorov-Smirnov. Hasil dari uji menunjukkan p -value yang kurang dari $\alpha = 0.05$ yang berarti data balita gizi buruk Jawa Timur tahun 2017 tidak mempunyai sebaran Poisson. Pengujian selanjutnya yaitu pengujian overdispersi. Pengujian overdispersi pada data spasial dilakukan untuk setiap lokasi. Hasil tes overdispersi menunjukkan nilai yang lebih dari 0 yang berarti bahwa terdapat overdispersi pada setiap lokasi di Provinsi Jawa Timur tahun 2017 sesuai lampiran 19. Adanya overdispersi menunjukkan sebaran data balita gizi buruk Jawa Timur tahun 2017 mempunyai sebaran binomial negatif.

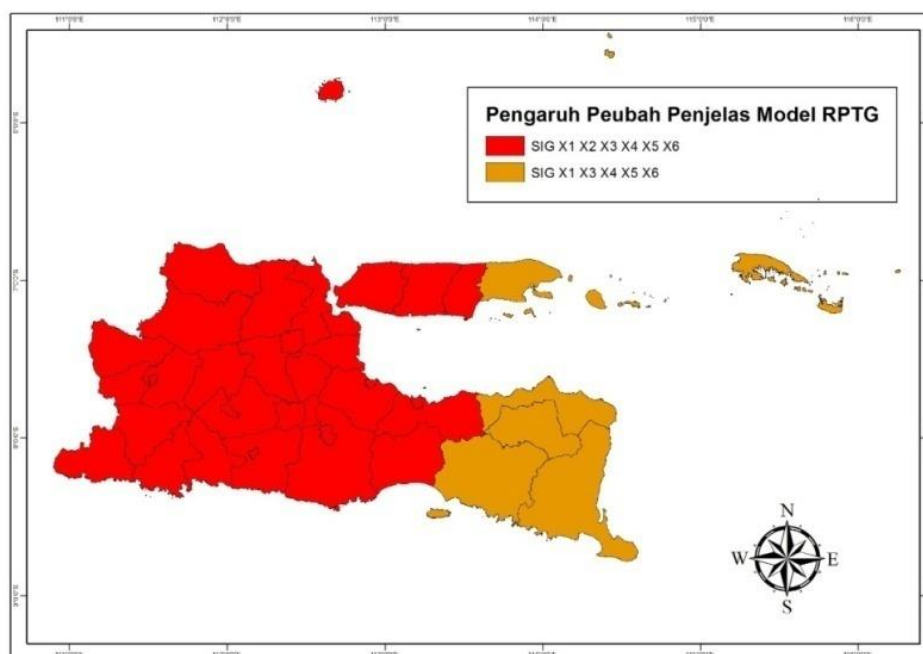


Gambar 5: Moran scatterplot.

Moran scatterplot adalah analisis eksplorasi secara visual untuk mendeteksi autokorelasi spasial. Hasil yang ditampilkan merupakan data yang telah di standarisasi. Gambar 5 menunjukkan bahwa lokasi pada kuadran I *High-high* yaitu lokasi yang mempunyai balita gizi buruk tinggi dikelilingi oleh lokasi yang mempunyai balita gizi buruk tinggi juga adalah Kabupaten Banyuwangi, Kabupaten Situbondo. Pada kuadran III terdapat 14 lokasi artinya 14 lokasi tersebut mempunyai balita gizi buruk yang rendah dikelilingi oleh lokasi yang mempunyai balita gizi buruk rendah juga. Pengujian heterogenitas spasial (keragaman spasial) dilakukan dengan menghitung nilai Breusch-Pagan. Nilai Breusch-Pagan yang diperoleh yaitu sebesar 14.45. Pada taraf 5% nilai $\chi^2_{(6;0.05)}$ sebesar 12.59 sehingga menghasilkan kesimpulan bahwa cukup bukti untuk menyatakan bahwa data gizi buruk Jawa Timur tahun 2017 terdapat keragaman spasial.

3.2.3 Regresi Poisson Terboboti Geografis

Model RPTG adalah pengembangan model RTG dengan peubah respon merupakan data cacahan. Hasil pengelompokan kota/kabupaten berdasarkan peubah yang signifikan pada gambar 6 menunjukkan bahwa kesamaan peubah yang berpengaruh signifikan sebanyak 2 kelompok. Kelompok pertama yang terdiri dari semua kabupaten/kota di Jawa Timur yang berpengaruh terhadap 6 peubah. Kelompok kedua adalah semua kabupaten/kota di Jawa Timur kecuali Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo dan Kabupaten Sumenep yang signifikan pada peubah persentase balita dengan pemberian ASI eksklusif (X_2). Kejadian balita gizi buruk di 5 kabupaten yaitu Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo dan Kabupaten Sumenep dapat diminimalisir dengan meningkatkan frekuensi sosialisasi terhadap ibu-ibu agar memberikan ASI eksklusif untuk anaknya.

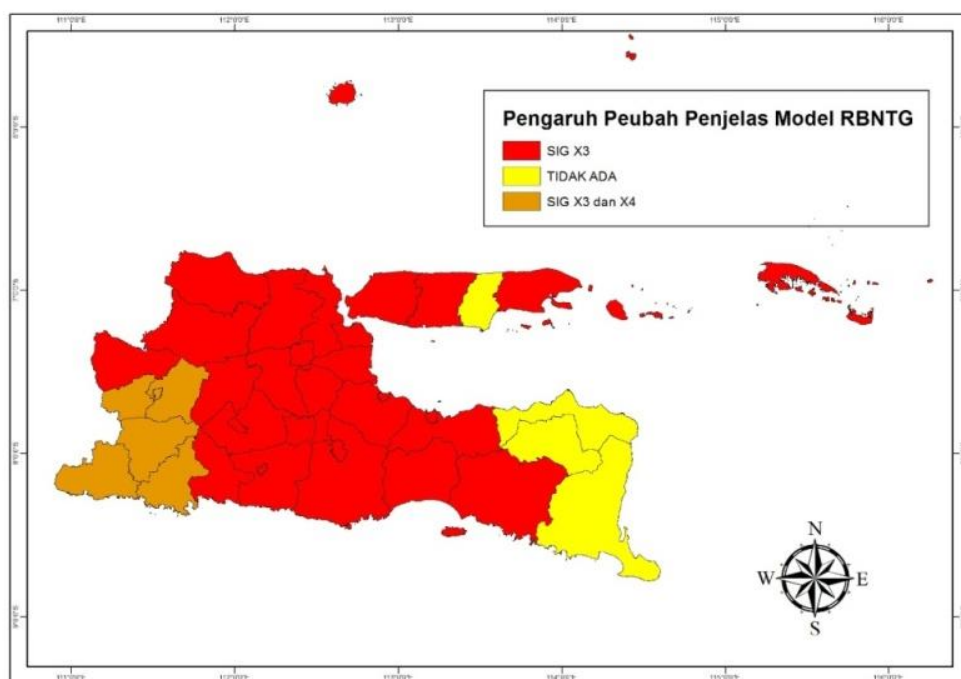


Gambar 6: Peta peubah penjelas yang signifikan model RPTG

Hasil pendugaan parameter pada RPTG masih kurang tepat digunakan pada kasus gizi buruk Provinsi Jawa Timur tahun 2017 karena adanya overdispersi pada data. Sesuai kajian simulasi, efek overdispersi pada model RPTG hampir semua peubah disetiap lokasi berpengaruh signifikan terhadap balita gizi buruk.

3.2.4 Regresi Binomial Negatif Terboboti Geografis

RBNTG tepat digunakan pada data yang mempunyai overdispersi dan keragaman spasial. Hasil pengelompokan kota/kabupaten berdasarkan peubah yang signifikan pada gambar 7 menunjukkan bahwa kesamaan peubah yang berpengaruh signifikan sebanyak 3 kelompok. Kelompok pertama adalah kelompok yang signifikan terhadap persentase balita yang mendapatkan vitamin A (X_3) yang terdiri dari semua kabupaten/kota di Jawa Timur kecuali Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo dan Kabupaten Pamekasan. Kelompok kedua adalah Kabupaten Madiun, Kota Madiun, Kabupaten Pacitan, Kabupaten Ponorogo, Kabupaten Trenggalek dan Kabupaten Magetan yang signifikan pada peubah Persentase posyandu aktif (X_4). Kejadian balita gizi buruk di 4 kabupaten yaitu Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo dan Kabupaten Pamekasan dapat diminimalisir dengan memastikan setiap balita mendapatkan vitamin A. Kelompok ketiga adalah Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo dan Kabupaten Pamekasan yang tidak signifikan terhadap 6 peubah.



Gambar 7: Peta peubah penjelas yang signifikan model RBNTG

Sesuai dengan kajian simulasi, untuk overdispersi menjauhi 1 pendugaan parameter model RBNTG menghasilkan penolakan hipotesis yang lebih sedikit daripada model RPTG. Hal ini mengindikasikan bahwa efek dari overdispersi pada kasus gizi buruk Provinsi Jawa Timur tahun 2017 sudah tertangani.

3.2.5 Pemilihan Model Terbaik

Pemilihan model terbaik bertujuan untuk mengetahui model mana yang lebih baik antara model RPTG dan RBNTG dalam memodelkan data balita gizi buruk di Jawa Timur tahun 2017. Berdasarkan Tabel 3 dapat dilihat bahwa model RBNTG memiliki nilai *AIC* dan *RMSE* terkecil serta nilai *Pseudo R²* terbesar sehingga model RBNTG lebih baik dalam memodelkan data balita gizi buruk di Jawa Timur tahun 2017.

Tabel 3: Nilai *AIC*, *Pseudo R²* dan *RMSE*

Model	<i>AIC</i>	<i>Pseudo R²</i>	<i>RMSE</i>
RPTG	464.134	49.130	117.428
RBNTG	110.888	65.481	63.748

4. SIMPULAN

Kesimpulan berdasarkan data simulasi yaitu batas toleransi yang masih tepat dimodelkan menggunakan RPTG yaitu overdispersi mendekati 1 yaitu rentang overdispersi 1.110 sampai 2.067. Semakin overdispersi menjauhi 1 atau rentang overdispersi 4.610 sampai 7.450, jumlah *p-value* yang signifikan atau jumlah penolakan H_0 semakin banyak. Model RBNTG merupakan model yang terbaik untuk data overdispersi pada kasus balita gizi buruk di Provinsi Jawa Timur tahun 2017 jika dibandingkan RPTG berdasarkan perbandingan *AIC*, *Pseudo R²* dan *RMSE*.

Daftar Pustaka

- Agresti, A. (2003). *Categorical data analysis* (Vol. 482). New York (US): John Wiley & Sons.
- da Silva, A. R., & Rodrigues, T. C. V. (2014). Geographically weighted negative binomial regression—incorporating overdispersion. *Statistics and Computing*, 24(5): 769–783.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester (UK): John Wiley & Sons.
- Liu, J., Zhao, Y., Yang, Y., Xu, S., Zhang, F., Zhang, X., ... Qiu, A. (2017). A mixed geographically and temporally weighted regression: Exploring spatial-temporal variations from global and local perspectives. *Entropy*, 19(2)(53): 1–20. <https://doi.org/10.3390/e19020053>
- McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models. Second edition*. London (UK): CRC Monographs on Statistics & Applied Probability.

- Miranti, Z., & Purhadi. (2016). Pemetaan Jumlah Balita Gizi Buruk di Kota Surabaya dengan GWNBR dan Flexibly Shaped Spatial Scan Statistic. *Jurnal Sains Dan Seni ITS*, 5(2): D247–D252. <https://doi.org/10.12962/j23373520.v5i2.16567>
- Nakaya, T., Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, 24(17): 2695–2717.
- Sofia, A. (2018). *Regresi Poisson Dan Spasial Otoresif Poisson Dalam Menduga Faktor-Faktor Kasus Gizi Buruk Balita di Pulau Jawa [Thesis]*. Bogor (ID): Institut Pertanian Bogor.