

Pendekatan *Flexibly Shaped Spatial Scan Statistic* untuk Deteksi Wilayah Kantong Penyakit DBD Melalui Pemodelan Regresi Binomial Negatif

Fefy Dita Sari dan Purhadi

Jurusan Statistika, FMIPA, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

E-mail: purhadi@statistika.its.ac.id, fefydita429@gmail.com

Abstrak— Menurut Laporan Profil Kesehatan Jawa Timur tahun 2013, Jawa Timur adalah provinsi dengan jumlah kejadian luar biasa (KLB) DBD tertinggi di Indonesia dengan angka kematian masih di bawah target yakni 1,04 persen. Dalam penelitian ini dilakukan pendeteksian wilayah kantong penyakit DBD dengan *Flexibly Shaped Spatial Scan Statistic*. Untuk melakukan pendeteksian tersebut dibutuhkan nilai ramalan jumlah kasus DBD di setiap kabupaten. Nilai ramalan dapat diperoleh dengan melakukan pemodelan regresi binomial negatif. Hasil penelitian menunjukkan bahwa variabel yang berpengaruh signifikan terhadap jumlah kasus DBD di Jawa Timur adalah persentase rumah tangga berperilaku hidup bersih dan sehat dan persentase sarana pendidikan dibina lingkungan kesehatannya. Hasil *Flexibly Shaped Spatial Scan Statistic* menunjukkan bahwa terdapat dua belas kantong DBD. Daerah paling rawan yaitu Kota Surabaya yang memiliki nilai resiko relatif sebesar 3,16. Daerah dengan resiko terbesar kedua yakni Kabupaten Bondowoso dan Kabupaten Jember dengan resiko relatif sebesar 2,10. Kabupaten dengan resiko relatif tertinggi ketiga yakni Kabupaten Sampang dengan resiko relatif 1,90.

Kata Kunci— DBD, *Flexibly Shaped Spatial Scan Statistic*, Regresi Binomial Negatif

I. PENDAHULUAN

DBD merupakan salah satu penyakit yang sampai saat ini masih menjadi masalah kesehatan masyarakat karena perjalanan penyakitnya cepat dan dapat menyebabkan kematian dalam waktu singkat. Indonesia yang berada di wilayah tropis pada daerah ekuator memungkinkan perkembangbiakan *Aedes aegypti* yang merupakan vektor dari virus *dengue*, sehingga Indonesia memiliki resiko besar untuk terjangkit dan tertular DBD dengan cepat [1]. Pada beberapa tahun terakhir sering terjadi peningkatan kejadian luar biasa (KLB) DBD terutama di pulau Jawa. Jawa Timur adalah provinsi dengan KLB paling banyak di Indonesia dengan angka kematian atau *case fatality rate* (CFR) masih berada di bawah target Nasional yakni 1,04% [2]. Oleh karena itu perlu dilakukan pencegahan dini penularan penyakit DBD di Jawa Timur. Agar pencegahan lebih tepat sasaran dan berlangsung secara efektif dan efisien maka dibutuhkan informasi wilayah kantong penyakit DBD.

Informasi tentang wilayah kantong penyakit DBD dapat diperoleh menggunakan metode *scan statistics*. Karena kantong wilayah DBD yang terbentuk tidak selalu berdekatan dan

memiliki bentuk kantong yang beraturan maka pendekatan yang digunakan adalah *Flexibly Shaped Spatial Scan Statistic*. Pada metode ini dibutuhkan nilai ramalan jumlah kasus DBD berdasarkan faktor-faktor yang mempengaruhi. Untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus DBD dapat digunakan analisis regresi [3].

Jumlah kasus DBD merupakan salah satu contoh data cacahan, sehingga analisis yang tepat digunakan untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus DBD di Jawa Timur adalah regresi Poisson. Dalam analisis regresi Poisson, sering kali muncul fenomena overdispersi [4]. Salah satu metode yang digunakan dalam mengatasi overdispersi dalam regresi Poisson adalah regresi Binomial Negatif.

Penelitian ini dilakukan untuk mengidentifikasi wilayah kantong penyakit DBD di Jawa Timur dengan *Flexibly Shaped Spatial Scan Statistic* dan untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus DBD di Jawa Timur, sehingga upaya pencegahan dan penanggulangan DBD dapat berjalan secara efektif, efisien dan tepat sasaran.

II. TINJAUAN PUSTAKA

A. Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu himpunan data, sehingga memberikan informasi yang berguna [5]. Contoh statistika deskriptif adalah rata-rata, koefisien varians, nilai minimum, nilai maksimum dan peta tematik. Salah satu bentuk penyajian statistika deskriptif adalah menggunakan peta tematik dengan metode *natural break* [6].

B. Multikolinieritas

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinearitas. Pendeteksian kasus multikolinieritas dilakukan koefisien korelasi dengan kriteria nilai VIF. Terdapat multikolinieritas jika koefisien koersali lebih dari 0,95. Nilai VIF (*Variance Inflation Factor*) lebih besar dari 10 menunjukkan adanya multikolinieritas. Nilai VIF dinyatakan sebagai berikut.

$$VIF_k = \frac{1}{1 - R_k^2} \quad (1)$$

dengan R_k^2 adalah koefisien determinasi antara satu variabel prediktor dengan variabel prediktor lainnya [7].

C. Regresi Poisson

Regresi Poisson merupakan model regresi nonlinear yang tepat digunakan untuk memodelkan data cacahan [8]. Jika variabel random diskrit (y) merupakan distribusi Poisson dengan parameter μ maka fungsi probabilitas dinyatakan sebagai berikut.

$$f(y, \mu) = \frac{e^{-\mu} \mu^y}{y!}; y = 0, 1, 2, \dots \quad (2)$$

dengan μ merupakan rata-rata dari y yang berdistribusi Poisson. Persamaan model regresi Poisson dapat ditulis sebagai berikut.

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p}) \quad (3)$$

Estimasi parameter dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) yang memaksimumkan fungsi log *likelihood*. Selanjutnya, estimasi parameter dilakukan dengan iterasi *Newton Raphson* dari turunan pertama fungsi log *likelihood* yang diturunkan terhadap β^T dan disamadengankan nol. Syarat agar diperoleh estimasi parameter dengan MLE adalah dihasilkannya matrik Hessian yang definit negatif. Matrik Hessian adalah turunan kedua fungsi log *likelihood* terhadap β . Fungsi log *likelihood* dirumuskan sebagai berikut.

$$\ln L(\beta) = - \sum_{i=1}^n e^{\mathbf{X}_i^T \beta} + \sum_{i=1}^n y_i \mathbf{X}_i^T \beta - \sum_{i=1}^n \ln(y_i!) \quad (4)$$

Pengujian signifikansi parameter terdiri dari uji serentak dan parsial. Uji signifikansi secara serentak menggunakan *Maximum Likelihood Ratio Test* (MLRT) dengan hipotesis sebagai berikut [9].

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0, j = 1, 2, \dots, p$$

$$\text{Statistik Uji: } D(\hat{\beta}) = -2 \ln \Lambda = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad (5)$$

Tolak H_0 jika $D(\hat{\beta}) > \chi^2_{(\alpha, p)}$ yang berarti minimal ada satu

parameter yang berpengaruh secara signifikan. Kemudian dilakukan pengujian parameter secara parsial dengan hipotesis sebagai berikut.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0; j = 1, 2, \dots, p$$

$$\text{Statistik Uji: } z_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (6)$$

Tolak H_0 jika $|z_{hitung}| > z_{\alpha/2}$ dengan α merupakan tingkat signifikansi yang ditentukan.

Overdispersi merupakan nilai dispersi *pearson Chi-square* atau *deviance* yang dibagi dengan derajat bebasnya, diperoleh nilai lebih besar dari 1 [10].

D. Regresi Binomial Negatif

Model regresi Binomial Negatif mempunyai fungsi distribusi sebagai berikut [11].

$$f(y, \mu, \theta) = \frac{\Gamma(y+1/\theta)}{\Gamma(1/\theta)y!} \left(\frac{1}{1+\theta\mu} \right)^{1/\theta} \left(\frac{\theta\mu}{1+\theta\mu} \right)^y \quad (7)$$

$$y = 0, 1, 2, \dots; \mu_i = \mathbf{X}_i^T \beta$$

Estimasi model regresi Binomial Negatif dinyatakan sebagai berikut.

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p}) \quad (8)$$

Metode *Maximum Likelihood Estimation* (MLE) digunakan untuk estimasi parameter. Selanjutnya, dengan iterasi *Newton Raphson* dari turunan pertama fungsi log *likelihood* yang diturunkan terhadap β, θ , dan disamadengankan nol. Syarat agar diperoleh estimasi parameter dengan MLE adalah dihasilkannya matrik Hessian yang definit negatif. Matrik Hessian adalah turunan kedua fungsi log *likelihood* terhadap β . Fungsi log *likelihood* dari regresi Binomial Negatif adalah sebagai berikut.

$$\ln\{L(\beta, \theta)\} = \sum_{i=1}^n \left[\left(\sum_{r=1}^{y_i-1} \ln(r + \theta^{-1}) \right) - \ln(y_i!) + y_i \ln(\theta\mu_i) - (\theta^{-1} + y_i) \ln(1 + \theta\mu_i) \right] \quad (9)$$

Pengujian signifikansi secara serentak untuk estimasi parameter model regresi Binomial Negatif dengan hipotesis sebagai berikut [12].

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1: \text{minimal ada satu } \beta_j \neq 0; j = 1, 2, \dots, p$$

Statistik Uji:

$$D(\hat{\beta}) = -2 \ln \Lambda = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad (10)$$

dimana $L(\hat{\omega})$ merupakan fungsi *likelihood* berdasarkan persamaan (9) tanpa melibatkan variabel prediktor, sedangkan $L(\hat{\Omega})$ merupakan fungsi *likelihood* dengan melibatkan variabel prediktor. Tolak H_0 jika $D(\hat{\beta}) > \chi^2_{(\alpha, p)}$ yang berarti minimal ada

satu parameter yang berpengaruh secara signifikan. Kemudian dilakukan pengujian parameter secara parsial dengan hipotesis sebagai berikut.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0; j = 1, 2, \dots, p$$

Statistik Uji:

$$z_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (11)$$

Tolak H_0 jika $|z_{hitung}| > z_{\alpha/2}$ yang berarti bahwa parameter ke- j signifikan terhadap model regresi Binomial Negatif.

E. Flexibly Shaped Spatial Scan Statistic

Spatial Scan Statistic merupakan salah satu metode statistik yang digunakan untuk mendeteksi *cluster* pada sebuah lokasi yang berupa titik maupun data agregat. Metode *Flexibly Spatial Scan Statistic* mempunyai *power* lebih tinggi daripada metode *Circular Spatial Scan Statistic* saat *cluster* yang dideteksi adalah *non-circular* dan fleksibel terhadap bentuk kantong yang dihasilkan sehingga tidak terbatas pada bentuk lingkaran saja [13].

Algoritma yang digunakan untuk mendapatkan Z dengan *pre-specified maximum length* L adalah sebagai berikut.

1. Membentuk himpunan wilayah Z

$$Z = \{Z_{il(m)} | 1 \leq i \leq n, 1 \leq l \leq L, 1 \leq m \leq m_{il}\}$$

dimana n adalah banyaknya data pengamatan, L adalah banyak wilayah maksimum dalam satu kantong dan m adalah banyak data bangkitan dengan simulasi monte-Carlo.

2. Himpunan wilayah Z dibagi menjadi dua himpunan wilayah baru yakni Z_0 dan Z_1 , dimana Z_0 adalah himpunan wilayah dari Z yang acak dan Z_1 juga merupakan himpunan wilayah acak dari Z selain Z_0
3. Buat 2 himpunan baru yakni Z_0' terdiri dari wilayah Z_1 yang berbatasan dengan Z_0 dan Z_1' terdiri dari wilayah Z_1 yang tidak berbatasan dengan Z_0
4. Ulangi langkah 3 hingga Z_0 dan Z_1 menjadi himpunan kosong.
5. Jika Z_1 menjadi himpunan kosong terlebih dahulu maka Z adalah berhubungan, jika Z_0 kosong terlebih dahulu maka Z tidak berhubungan
6. Ulangi langkah 2-5 hingga terbentuk maksimum L wilayah.

Untuk setiap wilayah i dan panjang dari *scanning window*, hipotesis alternatifnya adalah minimal ada satu *window Z* yang mempunyai peluang resiko lebih tinggi (*elevated risk*) daripada di luar *window*. Dengan kata lain, hipotesis yang digunakan adalah sebagai berikut.

$$H_0: E(Y(Z)) = \mu(Z) \text{ untuk semua } Z$$

$$H_1: E(Y(Z)) > \mu(Z) \text{ untuk beberapa } Z$$

dimana $Y(\cdot)$ melambangkan jumlah kasus yang random dan $\mu(\cdot)$ merupakan nilai harapan dari kasus *window* tertentu.

Statistik Uji:

$$\lambda = \left\{ \left(\frac{y(Z)}{\mu(Z)} \right)^{n(Z)} \left(\frac{y(Z^c)}{\mu(Z^c)} \right)^{n(Z^c)} I \left(\frac{y(Z)}{\mu(Z)} > \left(\frac{y(Z^c)}{\mu(Z^c)} \right) \right) \right\} \quad (12)$$

dimana Z^c melambangkan semua wilayah di luar *window Z*, $y(\cdot)$ melambangkan jumlah kasus dalam *window* yang ditentukan, $n(Z)$ adalah banyaknya wilayah dalam *window Z*, $n(Z^c)$ adalah banyaknya wilayah di luar *window Z*, dan $I(\cdot)$ merupakan fungsi indikator. Ketika memeriksa kantong dengan *high rates*, maka $I(\cdot)$ memiliki nilai 1 jika Z memiliki peluang lebih besar dan 0 jika lainnya.

Prosedur untuk mendapatkan *p-value* dengan pendekatan *Monte Carlo* adalah sebagai berikut.

1. Hitung penjumlahan nilai *log likelihood ratio* tertinggi t_0 untuk data riil.
2. Membangun data acak yang ukurannya sama dengan data riil yang dibangun di bawah kondisi H_0 .
3. Melakukan proses pembentukan *scanning window Z* dari data acak yang dibangun berdasarkan kondisi H_0 .
4. Mencari nilai *log likelihood ratio* dari setiap *scanning window*, dan dicatat apakah jumlah kasus yang diamati lebih besar atau lebih kecil dari yang diestimasi, kemudian menjumlahkan nilai *log likelihood ratio* yang jumlah kasusnya lebih besar dari jumlah yang diestimasi, untuk setiap *scanning window*. Langkah selanjutnya, mendapatkan penjumlahan nilai *log likelihood ratio* yang tertinggi dari simulasi pertama pembangunan data acak tersebut.
5. Mengulang langkah 2, 3, dan 4 sebanyak m kali pengulangan/simulasi, sehingga memperoleh m penjumlahan nilai *log likelihood ratio* tertinggi dari data acak dan data riil.

$$\text{Hitung } p\text{-value, } p = \frac{\text{banyaknya } (T(x) \geq t_0)}{m+1}$$

t_0 menyatakan penjumlahan nilai *log likelihood ratio* tertinggi yang dimiliki suatu *scanning window Z* dari data riil. $T(x)$ adalah penjumlahan nilai *log likelihood ratio* dari data acak yang dibangun di bawah kondisi H_0 . m adalah banyaknya simulasi untuk membangun data di bawah kondisi H_0 . Jika *P-value* < α maka tolak H_0 yang berarti bahwa *window Z* adalah wilayah kantong yang signifikan.

F. Pemilihan Model Terbaik

Kriteria yang digunakan untuk memilih model regresi terbaik adalah *Akaike Information Criterion* (AIC). AIC merupakan kriteria yang digunakan untuk memilih model terbaik yang didefinisikan sebagai berikut.

$$AIC = -2 \ln L(\beta) + 2K \quad (13)$$

dimana $\ln L(\beta)$ merupakan nilai likelihood yang di $\hat{\beta}$ i persamaan (4) untuk regresi Poisson dan persamaan (9) untuk regresi Binomial Negatif.

G. DBD

DBD adalah penyakit yang disebabkan oleh virus dengue yang masuk ke peredaran darah manusia melalui gigitan nyamuk dari genus *Aedes aegypti*. Penyakit DBD dapat muncul sepanjang tahun dan dapat menyerang seluruh kelompok umur. Penyakit ini berkaitan dengan kondisi lingkungan dan perilaku masyarakat. Penyakit ini masih menjadi permasalahan kesehatan masyarakat karena fatalitasnya dalam menyebabkan kematian dan kerapnya Kejadian Luar Biasa (KLB) yang terjadi pada bulan tertentu. Adapun nyamuk *Aedes aegypti* memiliki kemampuan terbang mencapai radius 100-200 meter. Oleh karena itu, jika di suatu lingkungan terkena kasus DBD, maka masyarakat yang berada pada radius tersebut harus waspada. Nyamuk *Aedes aegypti* lebih menyukai tempat yang gelap, berbau, dan lembap. Tempat perindukan yang sering dipilih *Aedes aegypti* adalah kawasan yang padat dengan sanitasi yang kurang memadai, terutama digenangan air dalam rumah, seperti pot, vas bunga, bak mandi atau tempat penyimpanan air lainnya seperti tempayan, drum, atau ember plastik [14].

Faktor lingkungan memegang peranan penting dalam penularan penyakit, terutama lingkungan rumah yang tidak memenuhi syarat. Lingkungan rumah merupakan salah satu faktor yang memberikan pengaruh besar terhadap status kesehatan penghuninya [15]. Penyakit DBD sering terjadi di daerah tropis dan muncul pada musim penghujan. Kurangnya kesadaran manusia dalam menjaga kebersihan lingkungan juga merupakan hal yang berpengaruh terhadap penyakit DBD [16].

III. METODOLOGI PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui publikasi data profil kesehatan di Dinas Kesehatan Provinsi Jawa Timur tahun 2014. Unit observasi penelitian sebanyak 38 kabupaten/kota di Provinsi Jawa Timur. Software yang digunakan adalah ArcView, FleXscan v3.1.2 dan program R. Penjelasan masing-masing variabel adalah sebagai berikut.

1. Jumlah kasus Demam Berdarah *Dengue* di tiap kabupaten/kota (Y)
2. Kepadatan Penduduk (X_1)

3. Persentase rumah/bangunan bebas jentik nyamuk Aedes (X_2)
4. Persentase rumah tangga berperilaku hidup bersih dan sehat (PHBS) (X_3)
5. Persentase rumah sehat (X_4)
6. Persentase sarana pendidikan yang dibina lingkungan kesehatannya (X_5)

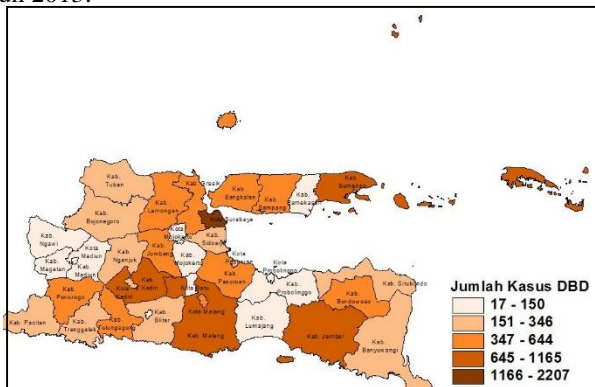
Langkah analisis yang dilakukan dalam penelitian ini yang didasarkan pada tujuan penelitian adalah sebagai berikut.

1. Mendeskripsikan karakteristik jumlah kasus DBD dan faktor-faktor yang mempengaruhinya di Provinsi Jawa Timur tahun 2013 menggunakan pemetaan wilayah untuk masing-masing variabel.
2. Pengujian kasus multikolinieritas berdasarkan kriteria korelasi dan VIF.
3. Menganalisis model regresi Poisson.
4. Menganalisis model regresi Binomial Negatif.
5. Melakukan pemilihan model terbaik dengan kriteria AIC.
6. Mendeteksi kantong DBD tingkat kabupaten/kota di provinsi Jawa Timur dengan *Flexibly Shaped Spatial Scan Statistic*.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Jumlah Kasus DBD Tahun 2013 di Provinsi Jawa Timur

Jumlah kasus DBD di Provinsi Jawa Timur sangat beragam di setiap kabupaten/kota. Keragaman data jumlah kasus DBD yang tinggi ini terjadi karena terdapat wilayah tertentu yang jumlah kasus DBD nya sangat rendah yakni Kota Mojokerto dengan jumlah kasus DBD hanya 17 kasus sepanjang tahun 2013 sedangkan kondisi yang berbeda terjadi di Kota Surabaya yakni dengan jumlah kasus DBD sebanyak 2.207 kasus pada tahun 2013.



Gambar 1. Persebaran Jumlah Kasus DBD di Jawa Timur (Y)

Pada Gambar 1 dengan indikator warna coklat tua merupakan kabupaten/kota dengan jumlah kasus DBD tinggi antara rentang 1.166 hingga 2.207 kasus yaitu Kota Surabaya.

B. Pemeriksaan Multikolinieritas

Berikut adalah salah satu cara mendeteksi multikolinieritas dengan kriteria korelasi dan VIF.

Tabel 1. Korelasi antar Variabel Prediktor

	X_1	X_2	X_3	X_4
X_2	0,184			
X_3	0,285	-0,080		

X_4	0,374	0,123	0,395	
X_5	0,201	0,409	-0,064	0,430

Tabel 2. Nilai VIF dari Variabel Prediktor

Variabel Prediktor	VIF
X_1	1,234
X_2	1,236
X_3	1,337
X_4	1,658
X_5	1,555

Tabel 1 dan 2 menunjukkan nilai koefisien korelasi kurang dari 0,95 dan VIF dari masing-masing variabel prediktor memiliki nilai VIF kurang dari 10, sehingga dapat disimpulkan bahwa tidak terdapat kasus multikolinieritas.

C. Regresi Poisson

Berikut ini merupakan estimasi parameter model regresi Poisson.

Tabel 3. Estimasi Parameter Model Regresi Poisson

	<i>Estimate</i>	<i>Std.Error</i>	<i>Z Value</i>	<i>P Value</i>
<i>(Intercept)</i>	5,815	0,1079	53,899	<2x10 ^{-16*}
X₁	6,064x10 ⁻⁵	4,044 x10 ⁻⁶	14,996	<2x10 ^{-16*}
X₂	-0,005211	1,288 x10 ⁻³	-4,045	5,24x10 ^{-5*}
X₃	0,02372	7,632 x10 ⁻⁴	31,079	<2x10 ^{-16*}
X₄	-0,001271	4,500 x10 ⁻⁴	-2,824	0.00474*
X₅	-0,007571	4,582 x10 ⁻⁴	-16,522	<2x10 ^{-16*}
Deviance: 8.757,8			DF: 31	
AIC: 9.044,8				

*) signifikan dengan taraf nyata 5%

Tabel 3 menunjukkan nilai devians sebesar 8.757,8 dengan taraf signifikansi 5% didapatkan $\chi^2_{(5;0,05)}$ sebesar 11,070 yang berarti minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Berdasarkan hasil pengujian secara individu diperoleh nilai *P-value* yang sangat kecil sehingga dapat disimpulkan bahwa semua variabel prediktor dalam model secara individu berpengaruh signifikan terhadap jumlah kasus DBD di Jawa Timur. Sehingga didapatkan model regresi Poisson sebagai berikut.

$$\ln(\hat{\mu}) = 5,815 + 6,064 \times 10^{-5} X_1 - 0,005211 X_2 + 0,02372 X_3 - 0,001271 X_4 - 0,007571 X_5$$

D. Regresi Binomial Negatif

Berikut ini merupakan estimasi parameter model regresi Binomial Negatif.

Tabel 4. Estimasi Parameter Model Regresi Binomial Negatif

	Estimate	Std.Error	Z Value	P Value
(Intercept)	5,985	1,806	3,314	0,00235*
X_1	3,612x10 ⁻⁵	7,024 x10 ⁻⁵	0,514	0,61076
X_2	-2,788x10 ⁻⁴	0,0211	-0,013	0,98954
X_3	0,01689	0,01098	1,539	0,13407*
X_4	-1,420x10 ⁻⁴	0,00729	-0,019	0,98459
X_5	-0,01099	0,008325	-1,320	0,19665*

Deviance: 31

DF: 31

AIC: 522,12

*) signifikan dengan taraf nyata 25%

Tabel 4 menunjukkan nilai devians sebesar 31 dengan taraf signifikansi 25% didapatkan $\chi^2_{(5;0,25)}$ sebesar 6,62568 yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Berdasarkan hasil pengujian secara individu hanya terdapat dua variabel prediktor yang signifikan, yaitu X_3 dan X_5 . Berikut ini merupakan model regresi Binomial Negatif.

$$\ln(\hat{\mu}) = 5,985 + 3,612 \times 10^{-5} X_1 - 2,788 \times 10^{-4} X_2 + 0,01689 X_3 - 1,420 \times 10^{-4} X_4 - 0,01099 X_5$$

Berdasarkan model yang terbentuk dapat diketahui bahwa setiap pertambahan 1 jiwa/Km² (X_1) maka akan menambah rata-rata jumlah kasus DBD sebesar $\exp(0,00003612) \approx 1$ kasus dengan asumsi variabel lain konstan. Selanjutnya setiap pertambahan persentase rumah/bangunan bebas jentik aedes (X_2) sebesar 1 persen maka jumlah kasus DBD akan berkurang sebesar $\exp(0,0002788) \approx 1$ kasus dengan asumsi variabel lain konstan. Penambahan 1 persen rumah tangga berperilaku hidup bersih dan sehat (X_3) akan meningkatkan jumlah kasus DBD sebesar $\exp(0,01689) \approx 1$ kasus DBD di setiap kabupaten/kota dengan asumsi variabel lain konstan. Hal ini tidak sesuai karena seharusnya semakin tinggi rumah tangga yang berperilaku hidup bersih dan sehat maka semakin sedikit jumlah kasus DBD nya. Namun ketidaksesuaian tersebut bukan berarti bahwa semakin tinggi persentase rumah tangga yang berperilaku hidup bersih dan sehat justru akan meningkatkan jumlah kasus DBD. Hal ini mungkin saja terjadi karena dampak dari perilaku hidup bersih dan sehat pada masyarakat tidak langsung memberikan dampak pada penurunan jumlah kasus DBD namun dampaknya dirasakan setelah beberapa interval waktu tertentu.

E. Pemilihan Model Terbaik

Pemilihan model terbaik dengan kriteria AIC adalah sebagai berikut.

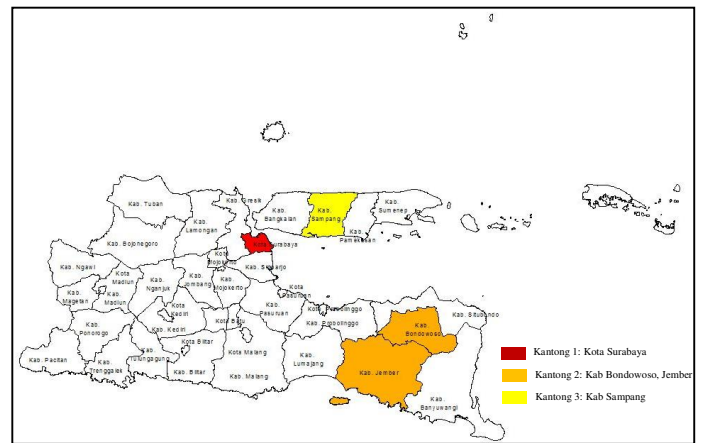
Tabel 5. Pemilihan Model Terbaik

Model	AIC
Regresi Poisson	9.044,8
Regresi Binomial Negatif	522,12

Tabel 5 menunjukkan bahwa model regresi Binomial Negatif memiliki nilai AIC lebih kecil dibanding dengan regresi poisson, sehingga dapat disimpulkan bahwa model regresi Binomial Negatif lebih baik dalam memodelkan jumlah kasus DBD di kabupaten/kota Jawa Timur.

F. Pendeteksian Kantong DBD dengan Flexibly Shaped Spatial Scan Statistic

Berikut adalah hasil pendeteksian kantong DBD di Jawa Timur dengan pengulangan sebanyak 9999.



Gambar 3. Peta Pendeteksian Kantong DBD di Jawa Timur

Tabel 8. Hasil Deteksi Kantong DBD di Jawa Timur

Kantong DBD	Juml. Kab	Kab/Kota	Resiko Relatif
1	1	Kota Surabaya	3,46
2	2	Kab. Jember, Bondowoso	2,10
3	1	Kab. Sampang	1,90
4	2	Kab. Malang, Kota Malang	1,86
5	1	Kab. Malang	1,80
6	1	Kab. Sumenep	1,72
7	2	Kab. Trenggalek, Kab. Tulungagung	1,45
8	1	Kab. Ponorogo	1,42
9	1	Kab. Bangkalan	1,28
10	1	Kab. Jombang	1,27
11	1	Kab. Situbondo	1,25
12	1	Kab. Pasuruan	1,24

Kota Surabaya merupakan kota yang paling rawan terkena DBD dimana Kota Surabaya beresiko terkena DBD 3,46 kali lebih besar dibanding kabupaten/kota lain. Kantong kedua terdiri dari Kabupaten Jember dan Kabupaten Bondowoso dengan jumlah kasus adalah resiko relatif 2,10. Kantong ketiga yakni Kabupaten Sampang yang memiliki resiko 1,90 kali lebih tinggi untuk terkena DBD dibanding kab/kota lain di luar kantong 3.

V. KESIMPULAN DAN SARAN

Hasil pemodelan regresi Binomial Negatif menunjukkan bahwa variabel yang signifikan adalah persentase rumah tangga berperilaku hidup bersih dan sehat dan persentase sarana pendidikan di bina lingkungan kesehatannya.

Hasil pendeteksian kantong DBD di Jawa Timur menunjukkan bahwa terbentuk dua belas kantong DBD dimana 1 wilayah yakni Kota Surabaya merupakan daerah paling rawan terkena DBD dengan resiko 3,46 kali di banding wilayah lainnya.

Saran untuk penelitian selanjutnya dapat dilakukan analisis dengan *Geographically Weighted Negative Binomial Regression* dengan menambah variabel-variabel baru.

DAFTAR PUSTAKA

- [1] World Health Organization. *Dengue Guidelines for Diagnosis, treatment, Prevention and Control*. Diakses dari <http://www.who.int/tdr/> pada 22 Desember 2015, 2009.
- [2] Dinas Kesehatan Provinsi Jawa Timur. *Profil Kesehatan Jawa Timur Tahun 2013*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2014.
- [3] Draper, N. R., dan Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). New York: John Wiley & Sons, Inc.
- [4] Cameron, A. C., & Trivedi, P. K. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press, 1998.
- [5] Walpole, R. E. (1995). *Pengantar Statistika*. Jakarta: Gramedia Pustaka Utama.
- [6] Expert Health Data Programming (2014). *What is Jenks Natural Breaks??* diakses dari <http://www.ehdp.com/vitalnet/breaks-1.htm> pada 30 desember 2015
- [7] Hocking, R.R. *Method and Applications of Linier Models*. New York: John Wiley and Sons, Inc, 1996.
- [8] Agresti, A. *Categorical Data Analysis Second Edition*. New York: John Wiley & Sons, 2002.
- [9] Mc Cullagh, P. and Nelder, J.A. *Generalized Linear Models Second Edition*. London: Chapman & Hall, 1989.
- [10] Hardin, J. W., & Hilbe, J. M. *Generalized Linear Models and Extensions Second Edition*. Texas: Stata Press, 2007.
- [11] Greene, W. *Functional Forms for the Negative Binomial Model for Count Data, Foundation, and Trends in Econometrics*, 99, 585-590. New York: New York University, 2008.
- [12] Hosmer, David Watson and Lemeshow, Sticher. *Applied Logistic Regression*. New York: John Wiley and Sons Inc, 1995.
- [13] Tango, T. dan Takahashi, k. *A Flexibly Shaped Spatal Scan Statistic For Detecting Clusters*. *International Journal of Health Geographics*. Volume 4:11, 2005.
- [14] Tobing, TMDNL. *Pemodelan Kasus DBD (DBD) di Jawa Timur dengan Model Poisson dan Binomial Negatif*. Bogor: Thesis Institut Pertanian Bogor, 2011.
- [15] Notoatmodjo, S. *Ilmu Kesehatan Masyarakat, Prinsip-Prinsip Dasar*. Jakarta: Rineka Cipta, 2003.
- [16] Sigarlaki, H.J.O. *Karakteristik, Pengetahuan, dan Sikap Ibu terhadap Penyakit Demam Berdarah Dengue*. Jakarta. Universitas Kristen Indonesia, 2012.