

PEMODELAN JUMLAH KASUS DEMAM BERDARAH *DENGUE* (DBD) DI JAWA TENGAH DENGAN *GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION* (GWNBR)

Indah Suryani¹, Hasbi Yasin², Puspita Kartikasari³

^{1,2,3} Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

Email: indahsuryani0610@gmail.com

ABSTRACT

Dengue Hemorrhagic Fever (DHF) is one of the diseases with unusual occurrence in Central Java and spread throughout the regency/city. The number sufferers of this disease is still high because the mortality rate is still above the national target. Regarding the less handling of DHF spread, it is necessary to make a plan by identify the factors that allegedly affect that case. Characteristics of data the DHF cases is count data, so this research is carried out using poisson regression. If in poisson regression there is overdispersion, it can be overcome using negative binomial regression. Meanwhile to see the spatial effect, we can use the Geographically Weighted Negative Binomial Regression (GWNBR) method. GWNBR modeling uses a fixed exponential kernel for weighting function. GWNBR is better at modeling the number of DHF cases because it has the smallest AIC value than poisson regression and negative binomial regression. The results of research with poisson regression obtained three variables that have a significant effect on dengue cases. For negative binomial regression, two variables have a significant effect on DHF cases. While the GWNBR method obtained two groups of districts/cities based on significant variables. The variables affecting the number of DHF cases in all districts/cities in Central Java are the percentage of healthy houses, the percentage of clean water quality, and the ratio of medical personnel.

Keywords: DHF, GWNBR, Poisson Regression, Binomial Negative Regression, Fixed Exponential Kernel

1. PENDAHULUAN

Demam Berdarah *Dengue* (DBD) merupakan salah satu jenis penyakit endemis di daerah yang beriklim tropis seperti Indonesia. Penyakit ini disebabkan oleh virus *dengue* yang penularannya melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus* betina. Penderita DBD sebagian besar adalah anak berumur kurang dari 15 tahun, namun dapat juga orang dewasa. Apabila jumlah penderita DBD atau wilayah yang terjangkit cenderung meningkat secara berkala, maka dapat dianggap sebagai Kejadian Luar Biasa (KLB).

Menurut DinKes Jateng (2018), Angka Kematian atau *Case Fatality Rate* (CFR) DBD di Provinsi Jawa Tengah pada tahun 2018 sebesar 1,05 persen, mengalami penurunan bila dibandingkan tahun 2017 yaitu 1,24 persen. Walaupun demikian, angka tersebut masih lebih tinggi dibandingkan dengan target nasional (<1%). Hal ini menunjukkan bahwa penanganan kasus penyebaran penyakit DBD belum maksimal. Untuk mengatasi permasalahan tersebut dapat dilakukan analisis pola penyebaran dan mengidentifikasi faktor-faktor yang diduga mempengaruhinya.

Jumlah penderita DBD di Jawa Tengah tahun 2018 merupakan data diskrit (*count*), sehingga metode regresi poisson dapat digunakan untuk menganalisis faktor-faktor yang berpengaruh terhadap jumlah kasus DBD. Dalam analisis regresi poisson, terdapat asumsi yang harus terpenuhi yaitu *variance* dari variabel responnya sama dengan *mean*. Akan tetapi, sering kali data *count* memiliki *variance* yang lebih besar dari *mean*, atau biasa disebut fenomena overdispersi (Cameron & Trivedi, 1998). Jika terjadi *overdispersion*, regresi poisson tidak sesuai untuk memodelkan data dan model yang akan terbentuk menghasilkan estimasi parameter yang bias.

Salah satu metode yang digunakan untuk mengatasi overdispersi dalam regresi poisson adalah regresi binomial negatif. Model binomial negatif lebih fleksibel dibandingkan model poisson karena asumsi mean dan variance tidak harus sama. Dengan memperhatikan aspek spasial (wilayah), maka digunakan metode Geographically Weighted Negative Binomial Regression (GWNBR).

Aspek spasial perlu diperhatikan karena setiap wilayah memiliki kondisi geografisnya yang berbeda. Kondisi tersebut menyebabkan adanya perbedaan jumlah kasus DBD antara satu dengan wilayah lainnya dikaitkan dengan kondisi lingkungan dalam rumah tangga penduduk. Kejadian DBD dapat berpindah dari suatu wilayah ke wilayah lain. Penyebarannya semakin meluas seiring dengan meningkatnya mobilitas dan kepadatan penduduk. Di samping itu, penyakit tersebut dapat ditularkan dengan sangat mudah dan tidak terbatas dengan wilayah administrasi. Maka dapat dikatakan DBD merupakan penyakit menular yang bersifat lintas batas. Hal ini juga sesuai dengan hukum pertama Tobler dalam Anselin (1988), yang menyatakan bahwa “Segala sesuatu memiliki hubungan dengan yang lainnya, akan tetapi sesuatu yang berdekatan akan memiliki hubungan yang lebih dari pada sesuatu yang berjauhan”.

Berdasarkan uraian tersebut, pada penelitian ini dilakukan pemodelan faktor-faktor yang mempengaruhi jumlah kasus DBD di Jawa Tengah menggunakan metode Geographically Weighted Negative Binomial Regression (GWNBR) dengan menggunakan pembobot *fixed exponential kernel*.

2. TINJAUAN PUSTAKA

2.1 Regresi Poisson

Regresi poisson merupakan model regresi nonlinier dengan variabel respon (Y) mengikuti distribusi poisson untuk mengatasi data *count*. fungsi probabilitas dari distribusi poisson dapat dinyatakan sebagai berikut (Myers, 1990).

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

dengan μ merupakan rata-rata variabel respon yang berdistribusi poisson dimana nilai rata-rata dan varian dari y mempunyai nilai lebih dari 0.

Persamaan model regresi poisson dapat ditulis sebagai berikut.

$$\hat{\mu}_i = \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_p X_{pi})$$

dengan $\hat{\mu}_i$ merupakan rata-rata jumlah kejadian yang terjadi dalam interval waktu tertentu.

Metode yang digunakan untuk menduga parameter regresi poisson adalah metode *Maximum Likelihood Estimation* (MLE). Dalam regresi poisson, parameter yang diestimasi adalah β_k . Untuk mendapatkan nilai taksiran langkah yang dilakukan adalah membentuk fungsi *likelihood* dari fungsi poisson dan dilanjutkan dengan iterasi Newton Raphson.

Pengujian parameter regresi poisson adalah untuk mengetahui pengaruh dari suatu parameter terhadap model dengan tingkat signifikansi tertentu. Pengujian kelayakan model regresi poisson dilakukan dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) yang memiliki hipotesis sebagai berikut.

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

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

Statistik uji:

$$G^2 = -2 \ln \Lambda = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right)$$

$L(\hat{\omega})$ dan $L(\hat{\Omega})$ merupakan dua fungsi *likelihood* yang berhubungan dengan model regresi yang diperoleh. $L(\hat{\omega})$ adalah nilai maksimum *likelihood* untuk model tanpa melibatkan

variabel prediktor dan $L(\hat{\Omega})$ adalah nilai maksimum *likelihood* untuk model yang melibatkan variabel prediktor.

Keputusan: Tolak H_0 jika nilai $G^2 > \chi^2_{(\alpha,p)}$, yang berarti minimal ada satu parameter yang berpengaruh secara signifikan terhadap model. G^2 merupakan statistik rasio likelihood yang mengikuti distribusi chi-kuadrat dengan derajat bebas p (McCullagh & Nelder, 1989).

Kemudian dilakukan pengujian parameter secara parsial untuk melihat signifikansi parameter terhadap model dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0; k = 1, 2, \dots, p$$

Statistik uji yang digunakan mengikuti distribusi z yaitu,

$$t_{hitung} = \frac{\hat{\beta}_k}{S\hat{E}(\hat{\beta}_k)}$$

$S\hat{E}(\hat{\beta}_k)$ adalah *standart error*, didapatkan dari elemen diagonal ke- $(k + 1)$ dari $\text{var}(\hat{\beta})$ dengan $\text{var}(\hat{\beta}) = -E(\mathbf{H}^{-1}(\hat{\beta}))$. H_0 akan ditolak jika nilai dari $|t_{hitung}|$ lebih besar dari nilai $t_{\alpha/2}$ dimana α adalah tingkat signifikansi parameter.

2.2 Overdispersi Regresi Poisson

Overdispersi merupakan kondisi dimana nilai varians lebih besar dari nilai mean $\{Var(Y) > E(Y)\}$. Overdispersi dapat menyebabkan suatu kesimpulan yang diperoleh tidak valid karena nilai *standart error* menjadi *under estimate*. Hal ini disebabkan karena parameter koefisien regresi yang dihasilkan tidak efisien meskipun koefisien regresinya tetap konsisten (Cameron & Trivedi, 1990).

Pemeriksaan overdispersi menggunakan statistik uji *Pearson's Chi-square* sebagai berikut.

$$\sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\sigma_i^2} \sim \chi^2_{(n-p)}; \text{ dengan } p = k + 1$$

Data mengalami overdispersi jika statistik uji *Pearson's Chi-square* dibagi dengan derajat bebas akan menghasilkan nilai lebih dari 1.

2.3 Regresi Binomial Negatif

Pada regresi binomial negatif, variabel respon diasumsikan berdistribusi binomial negatif yang dihasilkan dari distribusi *mixture poisson-gamma*. Untuk fungsi massa peluang binomial negatif adalah sebagai berikut (Greene, 2008).

$$f(y, \mu, \theta) = \frac{\Gamma(y + 1/\theta)}{\Gamma(1/\theta)y!} \left(\frac{1}{1 + \theta\mu}\right)^{1/\theta} \left(\frac{\theta\mu}{1 + \theta\mu}\right)^y; y = 0, 1, 2, \dots, n$$

Model regresi binomial negatif dinyatakan dalam bentuk kombinasi linier antara parameter (μ) dengan parameter regresi yang akan ditaksir yaitu:

$$\mu_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi})$$

Estimasi parameter model regresi binomial negatif menggunakan metode maksimum likelihood dengan prosedur Newton Raphson.

Pengujian signifikansi secara serentak untuk estimasi parameter model regresi Binomial Negatif menggunakan uji devians dengan hipotesis sebagai berikut (Hosmer & Lemeshow, 1995).

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

H_1 : minimal ada satu $\beta_k \neq 0; k = 1, 2, \dots, p$

Statistik uji:

$$G^2 = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2(\ln L(\hat{\Omega}) - \ln L(\hat{\omega}))$$

Tolak H_0 jika nilai $G^2 > \chi^2_{(\alpha, p)}$

Pengujian signifikansi secara parsial untuk mengetahui parameter mana saja yang memberikan pengaruh yang signifikan terhadap model dengan hipotesis sebagai berikut.

$H_0 : \beta_k = 0$

$H_1 : \beta_k \neq 0; k = 1, 2, \dots, p$

Statistik Uji:

$$t_{hitung} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}$$

H_0 ditolak jika statistik uji $|t_{hitung}|$ lebih besar dari $t_{(n-k; \alpha/2)}$, tolak H_0 artinya bahwa parameter ke-p signifikan terhadap model regresi Binomial Negatif.

2.4 Pengujian Aspek Data Spasial

Regresi spasial merupakan salah satu metode yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor dengan memperhatikan aspek lokasi atau spasial. Aspek spasial yang dimaksud adalah data yang digunakan memiliki error saling berkorelasi dan memiliki heterogenitas spasial (Anselin, 1988). Terdapat dua pengujian spasial yaitu uji heterogenitas spasial dan uji dependensi spasial.

Pengujian heteroskedastisitas spasial dilakukan untuk melihat apakah terdapat kekhasan pada setiap lokasi pengamatan. Pengujian heteroskedastisitas spasial dilakukan menggunakan statistik uji *Breusch-Pagan* (BP) dengan hipotesis sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (varians antar lokasi sama)

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$ (varians antar lokasi berbeda)

Statistik uji:

$$BP = (1/2) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f}$$

Keputusan: tolak H_0 jika nilai $BP > \chi^2_{(\alpha, k)}$ atau $p\text{-value} < \alpha$ yang berarti terjadi heteroskedastisitas dalam model.

Pengujian dependensi spasial dilakukan untuk melihat apakah pengamatan di suatu lokasi berpengaruh terhadap pengamatan di lokasi lain yang letaknya berdekatan. Pengujian dependensi spasial dilakukan menggunakan statistik uji *Moran's I* dengan hipotesis sebagai berikut.

$H_0 : I = 0$ (tidak terdapat dependensi spasial)

$H_1 : I \neq 0$ (terdapat dependensi spasial)

Statistik uji:

$$Z_{hitung} = \frac{\hat{I} - E(\hat{I})}{\sqrt{Var(\hat{I})}}$$

dengan

$$\hat{I} = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{(\sum_{i=1}^n \sum_{j=1}^n w_{ij}) \sum_{i=1}^n (y_i - \bar{y})^2}$$

dimana y_i merupakan nilai pengamatan pada lokasi ke-i dan y_j merupakan nilai pengamatan pada lokasi ke-j. \bar{y} merupakan rata-rata dari y_i dan n merupakan banyaknya pengamatan, sedangkan w_{ij} adalah elemen matriks pembobot kernel *fixed exponential*.

Rumus persamaan nilai *mean* dan varians *Moran's I* adalah sebagai berikut.

$$E(\hat{I}) = \frac{-1}{(n-1)}$$

$$Var(\hat{I}) = \frac{n^2s - ns_2 - 3s_0^2}{(n^2-1)s_0^2} - E(\hat{I})^2$$

Keputusan: tolak H_0 jika nilai $|Z_{hitung}| > Z_{\alpha/2}$, yang berarti terdapat dependensi spasial dalam model.

2.5 Model Geographically Weighted Negative Binomial Regression (GWNBR)

Model GWNBR akan menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda-beda. Model GWNBR dapat dirumuskan sebagai berikut (Ricardo & Carvalho, 2013).

$$y_i \sim NB \left[\exp \left(\sum_p \beta_p(u_i, v_i) x_{ip} \right), \theta(u_i, v_i) \right]; i = 1, 2, \dots, n$$

dimana,

y_i : nilai observasi respon ke- i

x_{ip} : nilai observasi variabel prediktor ke- p pada pengamatan lokasi (u_i, v_i)

$\beta_p(u_i, v_i)$: koefisien regresi variabel prediktor ke- p untuk setiap lokasi (u_i, v_i)

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

Fungsi sebaran binomial negatif untuk setiap lokasi dapat ditulis dalam bentuk persamaan sebagai berikut.

$$f(y_i | x_{ij} \beta_j(u_i, v_i), \theta(u_i, v_i)) = \frac{\Gamma(y_i + 1/\theta_i)}{\Gamma(1/\theta_i) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/\theta_i} \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right)^{y_i}$$

dengan $i = 1, 2, \dots$; dimana $\mu_i = \exp(\mathbf{X}_i^T \boldsymbol{\beta}(u_i, v_i))$ dan $\theta_i = \theta(u_i, v_i)$.

Estimasi parameter model GWNBR menggunakan metode maksimum *likelihood* dan didapatkan fungsi *likelihood* sebagai berikut.

$$L(\boldsymbol{\beta}(u_i, v_i), \theta_i | y_i, x_i) = \prod_{i=1}^n \left(\prod_{r=0}^{y_i-1} \left(r + \frac{1}{\theta_i} \right) \right) \left(\frac{1}{y_i!} \right) \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/\theta_i} \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/y_i}$$

Pengujian GWNBR terdiri dari pengujian kesamaan model GWNBR dengan regresi binomial negatif, pengujian signifikansi parameter model GWNBR secara serentak dan parsial.

Pengujian kesamaan model GWNBR dengan regresi Binomial Negatif dilakukan untuk melihat terdapat perbedaan yang signifikan atau tidak antara model GWNBR dengan regresi Binomial Negatif dengan hipotesis berikut.

$H_0 : \beta_k(u_i, v_i) = \beta_k$

$H_1 : \beta_k(u_i, v_i) \neq \beta_k; i = 1, 2, \dots, n; k = 1, 2, \dots, p$

Statistik uji:

$$F_{hitung} = \frac{\text{devians Model A/df A}}{\text{devians Model B/df B}}$$

Dimisalkan model A adalah model binomial negatif dan model B adalah model GWNBR. Tolak H_0 jika $F_{hitung} > F_{(\alpha, df_A, df_B)}$ yang artinya bahwa ada perbedaan yang signifikan antara model binomial negatif dengan model GWNBR, sehingga perlu dilakukan pengujian parameter model GWNBR serentak dan parsial (Ricardo & Carvalho, 2013).

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

$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_p(u_i, v_i) = 0$

H_1 : minimal ada satu $\beta_k(u_i, v_i) \neq 0; k = 1, 2, \dots, p$

Statistik uji:

$$G^2 = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2(\ln L(\hat{\Omega}) - \ln L(\hat{\omega}))$$

Tolak H_0 jika nilai $G^2 > \chi^2_{(\alpha, p)}$

Pengujian signifikansi secara parsial untuk mengetahui parameter mana saja yang memberikan pengaruh yang signifikan terhadap model dengan hipotesis sebagai berikut.

$H_0 : \beta_k(u_i, v_i) = 0$

$H_1 : \beta_k(u_i, v_i) \neq 0; k = 1, 2, \dots, p$

Statistik Uji:

$$t_{hitung} = \frac{\hat{\beta}_k(u_i, v_i)}{SE(\hat{\beta}_k(u_i, v_i))}$$

H_0 ditolak jika statistik uji $|t_{hitung}| > t_{\alpha/2}$. Tolak H_0 artinya bahwa parameter tersebut berpengaruh signifikan terhadap variabel respon di tiap lokasi.

2.6 Multikolinieritas

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi adalah tidak ada kasus multikolinieritas. Pendeteksian kasus multikolinieritas dilihat melalui beberapa cara yaitu sebagai berikut.

1. Jika koefisien korelasi Pearson (r_{ij}) antara variabel prediktor lebih dari 0,95 maka terdapat korelasi antar variabel tersebut.
2. Jika nilai *Variance Inflation Factor* (VIF) lebih besar dari 10, maka menunjukkan adanya multikolinieritas antar variabel prediktor. Nilai VIF dinyatakan sebagai berikut.

$$VIF = \frac{1}{1 - R_j^2}$$

Dengan R_j^2 adalah koefisien determinasi antara satu variabel prediktor dengan variabel prediktor lainnya (Hocking, 1996).

2.7 Penentuan *Bandwidth* dan Pembobot Optimum

Fungsi kernel memberikan pembobotan sesuai bandwidth optimum yang nilainya bergantung pada kondisi data. Terdapat dua jenis fungsi kernel, yaitu fungsi kernel tetap (*fixed kernel*) dan fungsi kernel adaptif (*adaptive kernel*). Dalam penelitian ini bobot fungsi kernel yang digunakan adalah *fixed exponential kernel* dengan rumus sebagai berikut (Fotheringham, et al., 2002).

$$w_j(u_i, v_i) = \sqrt{\exp \left(- \left(\frac{d_{ij}}{b} \right)^2 \right)}$$

dengan $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak *Euclidean* antara lokasi (u_i, v_i) dan b adalah nilai bandwidth optimum pada tiap lokasi.

Pemilihan *bandwidth* optimum menjadi sangat penting karena akan mempengaruhi ketepatan model terhadap data, yaitu mengatur varians dan bias dari model. Penentuan *bandwidth* optimum dilakukan menggunakan metode *Cross Validation* (CV) sebagai berikut.

$$CV(b) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(b))^2$$

$\hat{y}_{\neq i}(b)$ merupakan nilai penaksir y_i dengan pengamatan lokasi (u_i, v_i) dihilangkan dari proses penaksiran.

2.8 Seleksi Model Terbaik

Kriteria yang sering digunakan untuk mendapatkan model terbaik adalah nilai Akaike Information Criterion (AIC). Nilai AIC dapat dirumuskan sebagai berikut (Collet, 1994).

$$AIC = -2 \ln \hat{L}(\hat{\theta}) + 2k$$

dimana :

$\hat{L}(\hat{\theta})$ adalah nilai *likelihood*

k adalah jumlah parameter β pada setiap model yang terentu.

2.9 Demam Berdarah Dengue (DBD)

Demam Berdarah *Dengue* (DBD) merupakan penyakit akut yang disebabkan oleh virus *dengue* dan ditularkan melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus* betina (Kemenkes, 2010). Di Indonesia DBD merupakan endemis baik di daerah perkotaan (urban) maupun di daerah pedesaan (rural). Di daerah perkotaan vektor utamanya adalah nyamuk *Aedes aegypti* sedangkan di daerah pedesaan oleh nyamuk *Aedes albopictus*. Namun sering terjadi bahwa kedua spesies nyamuk tersebut terdapat pada suatu daerah secara bersama-sama, misalnya di daerah yang bersifat semi-urban (Soedarto, 2012). Tempat perindukan yang sering dipilih *Aedes aegypti* adalah kawasan yang padat dengan sanitasi yang kurang memadai, terutama di genangan air dalam rumah, seperti pot, vas bunga, bak mandi atau tempat penyimpanan air lainnya seperti tempayan, drum, atau ember plastik (Tobing, 2011). Menurut Notoatmodjo (2003), 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.

3. METODOLOGI PENELITIAN

3.1. Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder pada tahun 2018 yang diperoleh dari Dinas Kesehatan Provinsi Jawa Tengah dan Badan Pusat Statistik (BPS). Unit observasi yang digunakan dalam penelitian ini adalah sebanyak 35 Kabupaten/Kota di Jawa Tengah. Mengingat data yang digunakan adalah data sekunder, diasumsikan bahwa alat ukur (kuesioner) yang digunakan telah tervalidasi dan petugas telah mengisi dengan benar.

3.2. Variabel Penelitian

Variabel dependen dalam penelitian ini adalah jumlah kasus DBD di Jawa Tengah pada tahun 2017 (Y). Untuk variabel independen terdiri dari 5 variabel yaitu, Persentase Rumah Sehat (X_1), Persentase Penduduk Miskin (X_2), Persentase Kualitas Air Bersih (X_3), Persentase Penduduk dengan Air Minum Layak (X_4), dan Rasio Tenaga Medis (X_5).

3.3. Langkah-langkah Penelitian

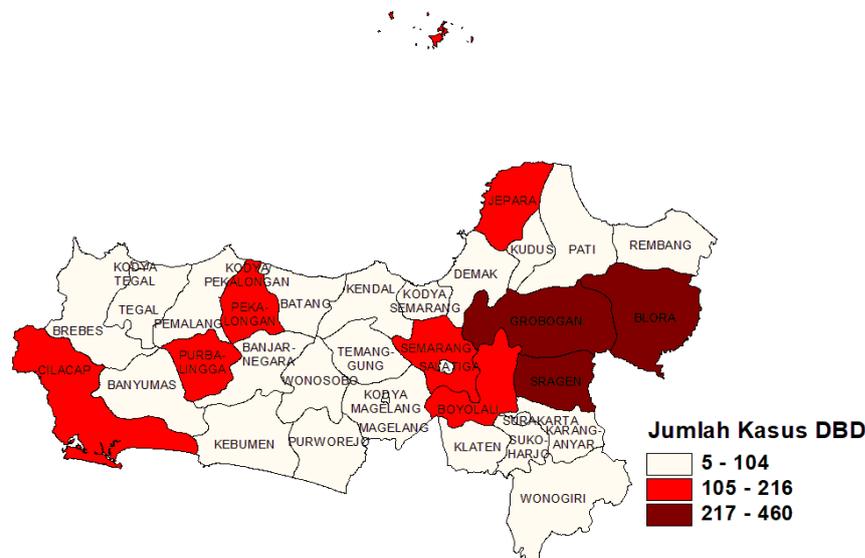
1. Mendeskripsikan karakteristik jumlah kasus DBD di Jawa Tengah tahun 2018 beserta faktor-faktor yang mempengaruhinya menggunakan pemetaan wilayah.
2. Pengujian kasus multikolinieritas berdasarkan kriteria koefisien korelasi dan VIF.
3. Menganalisis model regresi poisson dengan langkah sebagai berikut.
 - a. Penaksiran parameter model regresi poisson menggunakan metode *MLE*.
 - b. Menguji signifikansi parameter model regresi poisson secara serentak dan parsial.

- c. Menghitung nilai AIC model regresi poisson.
- d. Melakukan uji dispersi model regresi poisson.
4. Menganalisis model regresi binomial negatif dengan langkah sebagai berikut.
 - a. Penaksiran parameter model binomial negatif dengan metode MLE.
 - b. Menguji signifikansi parameter model binomial negatif secara serentak dan parsial.
 - c. Menghitung nilai AIC model binomial negatif.
5. Memodelkan GWNBR untuk kasus DBD di Provinsi Jawa Tengah tahun 2018, dengan langkah-langkah sebagai berikut.
 - a. Uji *Breusch-Pagan* untuk melihat heterogenitas spasial data dan uji *Moran's I* untuk menguji dependensi spasial data.
 - b. Menghitung jarak *Euclidean* antar lokasi pengamatan berdasarkan posisi geografis.
 - c. Mendapatkan *bandwidth* optimal untuk setiap lokasi pengamatan dengan menggunakan *Cross Validation (CV)*.
 - d. Menghitung matriks pembobot dengan menggunakan fungsi kernel *fixed exponential*.
 - e. Melakukan interpretasi model GWNBR yang didapatkan dan membentuk peta pengelompokkan.
 - f. Melakukan pemilihan model terbaik menggunakan AIC.

4. HASIL DAN PEMBAHASAN

4.1. Karakteristik Jumlah Kasus Demam Berdarah *Dengue* (DBD) di Jawa Tengah Tahun 2018

Kasus DBD di Jawa Tengah tahun 2018 mencapai 3519 kasus dengan rata-rata sebesar 100,543 kasus. Penyakit DBD ini telah menjangkit di seluruh wilayah Kabupaten/Kota di Jawa Tengah dengan jumlah tertinggi berada di Kabupaten Grobogan, yaitu sebanyak 460 kasus dan untuk jumlah kasus DBD terendah berada di Kota Tegal, yaitu 5 kasus. Pola persebaran jumlah kasus DBD di Jawa Tengah tahun 2018 dapat dilihat pada Gambar 1.



Gambar 1. Persebaran Kasus Jumlah DBD di Jawa Tengah tahun 2018

4.2. Permeriksaan Multikolinieritas

Pemeriksaan multikolinieritas dapat dilakukan dengan melihat nilai korelasi Pearson dan nilai VIF (Variance Inflation Factor). Berikut ini disajikan nilai korelasi Pearson.

Tabel 1. Koefisien Korelasi antar Variabel Prediktor

	X ₁	X ₂	X ₃	X ₄
X ₂	-0,308			
X ₃	0,156	-0,061		
X ₄	-0,187	0,005	-0,044	
X ₅	0,537	-0,456	0,030	-0,082

Berdasarkan Tabel 1, dapat diketahui bahwa semua variabel prediktor memiliki koefisien korelasi *Pearson* yang kurang dari 0,95 yang artinya tidak terdapat kasus multikolinieritas. Selain dari koefisien korelasi *Pearson*, kasus multikolinieritas dapat dilihat dari nilai VIF. Berikut ini nilai VIF dari setiap variabel prediktor.

Tabel 2. Nilai VIF dari Variabel Prediktor

Variabel	X ₁	X ₂	X ₃	X ₄	X ₅
VIF	1,50	1,28	1,03	1,04	1,63

Berdasarkan Tabel 2, nilai VIF dari semua variabel prediktor menunjukkan nilai yang kurang dari 10, maka dapat disimpulkan tidak terdapat kasus multikolinieritas, sehingga dapat dilanjutkan ke pemodelan regresi poisson dan binomial negatif.

4.3. Regresi Poisson

Setelah dilakukan pemeriksaan kasus multikolinieritas antara variabel prediktor dilanjutkan pada pemodelan regresi poisson. Data jumlah kasus DBD diasumsikan berdistribusi poisson karena merupakan data diskrit (*count*). Berikut merupakan estimasi parameter model regresi poisson.

Tabel 3. Penaksiran Parameter Model Regresi Poisson

Parameter	Estimasi	t-hitung	p-value
$\hat{\beta}_0$	3,508027	16,620	<2e-16*
$\hat{\beta}_1$	0,002087	1,300	0,19371
$\hat{\beta}_2$	-0,017719	-3,251	0,00115*
$\hat{\beta}_3$	0,017602	12,002	<2e-16*
$\hat{\beta}_4$	-0,000412	-0,398	0,69043
$\hat{\beta}_5$	-0,711769	-13,382	<2e-16*
Devians = 2345,4		Df = 29	
AIC=2568,4			

*) signifikansi dengan taraf nyata 5%

Berdasarkan hasil pengujian serentak dengan taraf signifikansi 5% diperoleh nilai $\chi^2_{(5)}$ sebesar 11,0705 yang artinya lebih kecil dari nilai devians (2345,4) maka tolak H_0 yang berarti minimal ada satu parameter yang berpengaruh terhadap model. Untuk pengujian secara individu dengan taraf signifikansi 5% didapatkan $t_{(0,05/2)}$ sebesar 2,045 maka diperoleh dari kelima variabel prediktor dalam model terdapat tiga variabel yang memberikan pengaruh secara signifikan terhadap jumlah kasus DBD di Jawa Tengah. Diantaranya adalah persentase penduduk miskin (X_2), persentase kualitas air bersih (X_3), dan rasio tenaga medis (X_5).

Berdasarkan Tabel 3 maka model regresi poisson yang terbentuk adalah sebagai berikut.

$$\ln(\hat{\mu}) = 3,508027 + 0,002087X_1 - 0,017719X_2 + 0,017602X_3 - 0,000412X_4 - 0,711769X_5$$

4.4. Deteksi Overdispersi

Berdasarkan hasil pengujian asumsi overdispersi, dapat diketahui bahwa parameter dispersi atau nilai rasio devians dengan derajat bebasnya adalah sebesar 81,80995 dan lebih besar dari 1, sehingga dapat disimpulkan bahwa terdapat overdispersi pada model. Kesimpulan tersebut diperkuat dengan nilai *z-score* model = 2,4728 dengan *p-value* = 0,006704. Sehingga pada taraf signifikansi 5% ($Z_{(0,05/2)}=1,96$) menghasilkan keputusan untuk menolak H_0 . Dengan demikian, untuk mengatasi kasus overdispersi dapat digunakan regresi binomial negatif untuk memodelkan jumlah kasus DBD di Jawa Tengah tahun 2018.

4.5. Regresi Binomial Negatif

Sebelum melakukan pemodelan dengan binomial negatif, terlebih dahulu mencari nilai initial θ yaitu didapatkan sebesar 1,7082. Untuk estimasi parameter model binomial negatif sebagai berikut.

Tabel 4. Penaksiran Parameter Model Regresi Binomial Negatif

Parameter	Estimasi	t-hitung	p-value
$\hat{\beta}_0$	2,732257	1,857	0,06332
$\hat{\beta}_1$	0,004494	0,354	0,72310
$\hat{\beta}_2$	-0,017080	-0,410	0,68217
$\hat{\beta}_3$	0,022124	2,620	0,00879*
$\hat{\beta}_4$	0,001965	0,243	0,80769
$\hat{\beta}_5$	-0,684362	-2,385	0,01710*
Devians = 38,004		Df = 29	
AIC=395,23			

*) signifikansi dengan taraf nyata 5%

Berdasarkan hasil pengujian serentak dengan taraf signifikansi 5% diperoleh nilai $\chi^2_{(5)}$ sebesar 11,0705 yang artinya lebih kecil dari nilai devians (38,004) maka tolak H_0 yang berarti minimal ada satu parameter yang berpengaruh terhadap model. Untuk pengujian secara individu dengan taraf signifikansi 5% didapatkan $t_{(0,05/2)}$ sebesar 2,045 maka diperoleh dari kelima variabel prediktor dalam model terdapat dua variabel yang memberikan pengaruh secara signifikan terhadap jumlah kasus DBD di Jawa Tengah, yaitu persentase kualitas air bersih (X_3) dan rasio tenaga medis (X_5).

Berdasarkan Tabel 4 maka model regresi poisson yang terbentuk adalah sebagai berikut.

$$\ln(\hat{\mu}) = 2,732257 + 0,004494X_1 - 0,017080X_2 + 0,022124X_3 - 0,001965X_4 - 0,684362X_5$$

4.6. Pengujian Aspek Data Spasial

Berdasarkan hasil pengujian heterogenitas diperoleh nilai statistik uji *Breusch-Pagan* (BP) sebesar 2,0821 dengan *p-value* sebesar 0,8377. Pada taraf signifikansi 5% didapatkan $\chi^2_{(5)}$ sebesar 11,0705, maka dapat disimpulkan bahwa variansi antar lokasi sama atau tidak terdapat perbedaan karakteristik antara satu lokasi pengamatan dengan yang lainnya. Berdasarkan hasil pengujian dependensi spasial diperoleh *p-value* sebesar 0,009268. Pada taraf signifikansi 5% maka dapat disimpulkan bahwa terdapat dependensi spasial atau pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang letaknya berdekatan.

Dari hasil pengujian heterogenitas spasial dan dependensi spasial, dapat disimpulkan bahwa data jumlah kasus DBD di Jawa Tengah hanya memenuhi salah satu asumsi yaitu

asumsi dependensi spasial. Walaupun demikian pada penelitian ini ingin melihat efek spasial sehingga dapat dilanjutkan pemodelan dengan menggunakan metode GWNBR.

4.7. Pemodelan GWNBR

Sebelum melakukan pemodelan GWNBR, terlebih dahulu menentukan *bandwidth* optimum menggunakan kriteria CV untuk menentukan pembobot dalam melakukan penaksiran parameter. Berdasarkan hasil pengujian diperoleh *best bandwidth size* dengan fungsi kernel *fixed exponential* sebesar 3,067126. Dalam melakukan perhitungan matriks pembobot membutuhkan perhitungan jarak (d_{ij}) antar Kabupaten/Kota di Provinsi Jawa Tengah. Selanjutnya dalam pemodelan GWNBR terdapat tiga pengujian parameter yaitu Pengujian Kesamaan Metode GWNBR dengan Binomial Negatif, Pengujian Serentak, dan Pengujian Parsial.

a) Pengujian Kesamaan Metode GWNBR dengan Binomial Negatif

Berdasarkan hasil pengujian kesamaan metode diperoleh nilai F_{hitung} sebesar 2,7008. Pada taraf signifikansi 5% didapatkan $F_{(0,05,29,29)}$ sebesar 1,8608 maka tolak H_0 . Sehingga dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara model binomial negatif dengan model GWNBR.

b) Pengujian Serentak

Berdasarkan hasil pengujian serentak diperoleh nilai devians model GWNBR sebesar 55,77588. Pada taraf signifikansi 5% didapatkan $\chi^2_{(5)}$ sebesar 11,0705 yang artinya bahwa paling tidak ada satu parameter model GWNBR yang berpengaruh signifikan terhadap model, maka perlu dilanjutkan dengan pengujian parsial.

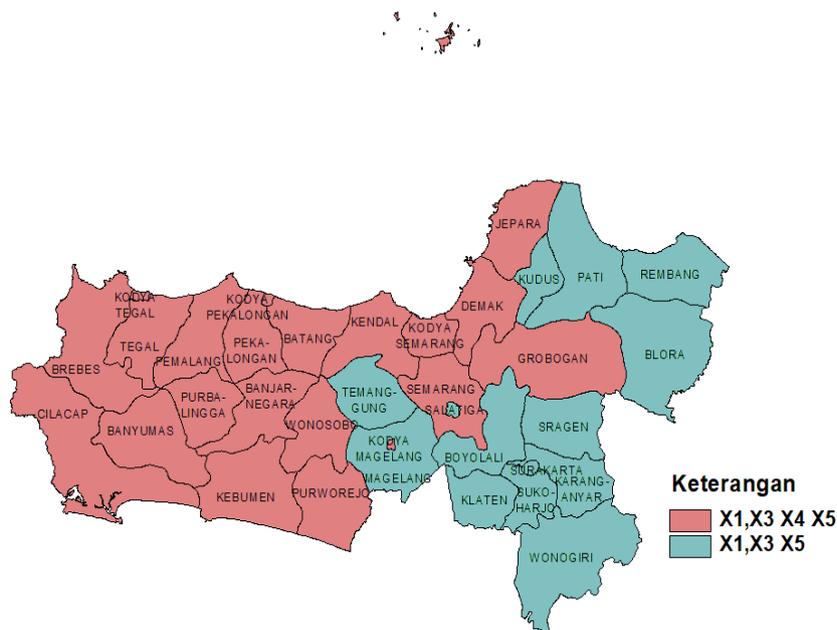
c) Pengujian Parsial

Berdasarkan hasil pengujian secara serentak diperoleh parameter yang signifikan berbeda-beda untuk tiap Kabupaten/Kota. Nilai $|t_{hit}|$ parameter setiap Kabupaten/Kota dibandingkan dengan nilai $Z_{(0,05/2)}$. Jika nilai $|t_{hit}| > 2,045$ maka tolak H_0 , yang artinya variabel tersebut memberikan pengaruh pada model. Pengelompokan variabel yang signifikan disajikan pada tabel berikut.

Tabel 5. Pengelompokan Kabupaten/Kota Berdasarkan Variabel yang Signifikan dalam Model GWNBR

Variabel yang Signifikan	Kabupaten/Kota	Jumlah
X_1, X_3, X_4, X_5	Kab. Cilacap, Kab. Banyumas, Kab. Purbalingga, Kab. Banjarnegara, Kab. Kebumen, Kab. Purworejo, Kab. Wonosobo, Kab. Grobogan, Kab. Jepara, Kab. Demak, Kab. Semarang, Kab. Kendal, Kab. Batang, Kab. Pekalongan, Kab. Pemalang, Kab. Tegal, Kab. Brebes, Kota Magelang, Kota Semarang, Kota Pekalongan, Kota Tegal	21
X_1, X_3, X_5	Kab. Magelang, Kab. Boyolali, Kab. Klaten, Kab. Sukoharjo, Kab. Wonogiri, Kab. Karanganyar, Kab. Sragen, Kab. Blora, Kab. Rembang, Kab. Pati, Kab. Kudus, Kab. Temanggung, Kota Surakarta, Kota Salatiga	14

Pengelompokan Kabupaten/Kota berdasarkan variabel yang signifikan di atas jika digambarkan dalam bentuk peta Jawa Tengah adalah sebagai berikut.



Gambar 2. Pengelompokan Kabupaten/Kota Berdasarkan Variabel yang Signifikan dalam Model GWNBR

Pada Gambar 8, Kabupaten dengan warna merah merupakan lokasi dengan variabel yang signifikan mempengaruhi jumlah kasus DBD yaitu X_1, X_3, X_4, X_5 dan Warna biru menunjukkan kelompok Kabupaten/Kota dengan variabel yang signifikan mempengaruhi jumlah kasus DBD yaitu X_1, X_3, X_4, X_5 . Variabel yang signifikan memberi pengaruh jumlah kasus DBD di semua Kabupaten/Kota adalah persentase rumah sehat (X_1), persentase kualitas air bersih (X_3) dan rasio tenaga medis (X_5).

Sebagai contoh akan disajikan pengujian parameter pada lokasi penelitian yang ke-33 (u_{33}, v_{33}) yaitu Kota Semarang.

Tabel 6. Pengujian Parameter Model GWNBR Di Kota Semarang

Parameter	Estimasi	Zhitung
$\hat{\beta}_0$	1,55512	4,6699*
$\hat{\beta}_1$	0,00020	13,2274*
$\hat{\beta}_2$	-0,00028	-1,7711
$\hat{\beta}_3$	0,00085	66,2712*
$\hat{\beta}_4$	-0,00002	-2,3918*
$\hat{\beta}_5$	-0,00027	-243,7268*

*) signifikansi dengan taraf nyata 5%

Suatu variabel memberikan pengaruh yang signifikan jika $|t_{hit}| > t_{(0,05/2)}$, dengan taraf nyata 5% maka $t_{(0,05/2)}$ adalah 2,045 sehingga dapat diketahui variabel yang signifikan adalah X_1, X_3, X_4 , dan X_5 . Maka model yang terbentuk adalah sebagai berikut.

$$\ln(\hat{\mu}) = 1,55512 + 0,00020X_1 - 0,00028X_2 + 0,00085X_3 - 0,00002X_4 - 0,00027X_5$$

Berdasarkan dari variabel yang signifikan dari model yang terbentuk di Kota Semarang dapat diperoleh kesimpulan bahwa setiap pertambahan 1 persen rumah sehat (X_1) maka akan menambah rata-rata jumlah kasus DBD sebesar $\exp(0,00020) = 1,00020 \approx 1$ kasus dengan asumsi variabel lain konstan. Pada setiap pertambahan 1 persen kualitas air bersih (X_3) maka akan menambah rata-rata jumlah kasus DBD sebesar $\exp(0,00085) = 1,00085 \approx 1$ kasus dengan asumsi variabel lain konstan. Untuk setiap pertambahan 1 persen

penduduk dengan akses air minum berkelanjutan yang layak (X_4) maka akan mengurangi rata-rata jumlah kasus DBD sebesar $\exp(0,00002) = 1,00002 \approx 1$ kasus dengan asumsi variabel lain konstan. Dan setiap kenaikan satu satuan rasio tenaga medis (X_5) maka akan mengurangi rata-rata jumlah kasus DBD sebesar $\exp(0,00027) = 1,00027 \approx 1$ kasus dengan asumsi variabel lain konstan.

4.8. Penentuan Model Terbaik dengan AIC

Pemilihan model terbaik berdasarkan kriteria AIC pada model regresi poisson, regresi binomial negatif, dan GWNBR adalah sebagai berikut.

Tabel 7. Pemilihan Model terbaik Terbaik dengan AIC

Model	AIC
Regresi Poisson	2568,40
Regresi Binomial Negatif	395,23
GWNBR	358,33

Tabel 7 menunjukkan bahwa dari ketiga model tersebut, GWNBR memiliki AIC paling kecil dibandingkan dengan regresi poisson dan regresi binomial negatif sehingga GWNBR lebih baik dalam memodelkan jumlah kasus DBD masing-masing Kabupaten/Kota di Jawa Tengah tahun 2018.

5. KESIMPULAN

Pada tahun 2018 persebaran kasus DBD di Jawa Tengah terjadi di seluruh Kabupaten/Kota. Jumlah Kasus DBD tertinggi adalah di Kabupaten Grobogan sebanyak 460 kasus yang kemudian diikuti oleh Kabupaten Blora dan Kabupaten Sragen dengan masing-masing kasus sebanyak 388 dan 316 kasus. Sedangkan jumlah kasus DBD terendah berada di beberapa Kabupaten/Kota di Jawa tengah dengan tingkat kasus 5-104 kasus. Pemodelan Jumlah Kasus DBD menggunakan regresi poisson didapatkan bahwa terdapat tiga variabel yang berpengaruh secara signifikan terhadap model yaitu X_2 , X_3 , dan X_5 . Sedangkan dengan menggunakan regresi binomial negatif hanya diperoleh dua variabel yang berpengaruh secara signifikan terhadap model yaitu X_3 , dan X_5 . Berdasarkan hasil pemodelan GWNBR dengan fungsi pembobot kernel *fixed exponential* didapatkan pengelompokkan sebanyak 2 kelompok berdasarkan variabel-variabel yang signifikan. Kelompok pertama beranggotakan 21 Kabupaten/Kota dengan variabel yang signifikan yaitu persentase rumah sehat (X_1), persentase kualitas air bersih (X_3), persentase penduduk dengan air minum layak (X_4), dan rasio tenaga medis (X_5). Kelompok kedua beranggotakan 14 Kabupaten/Kota dengan variabel yang signifikan yaitu persentase rumah sehat (X_1), persentase kualitas air bersih (X_3), dan rasio tenaga medis (X_5). Pemodelan menggunakan metode GWNBR diperoleh nilai AIC terkecil sehingga dapat dikatakan metode ini yang paling tepat untuk memodelkan jumlah kasus DBD di Jawa Tengah.

DAFTAR PUSTAKA

- Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Cameron, A. & Trivedi, P., 1998. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- Cameron, A. & Trivedi, P., 1990. Regression-Based Test For Overdispersion In The Poisson Model. *Journal of Econometrics*, 46(1), pp. 347-346.
- Collet, D., 1994. *Modelling Survival Data in Medical Research*. London: Chapman and Hall.
- Dainur, 1995. *Materi-Materi Pokok Ilmu Kesehatan Masyarakat*. Jakarta: Widya Medika.
- Depkes RI-Ditjen PPM dan PL, 2002. *Pedoman Teknis Penilaian Rumah Sehat*. Jakarta: Departemen Kesehatan Republik Indonesia.

- Dinas Kesehatan Provinsi Jawa Tengah, 2018. *Profil Kesehatan Provinsi Jawa Tengah Tahun 2018*. Semarang: Dinas Kesehatan Provinsi Jawa Tengah.
- Dinas Kesehatan Provinsi Riau, 2018. *Peningkatan Kemampuan Dokter Puskesmas dan RS tentang Tatalaksana DBD*, Pekanbaru: Dinas Kesehatan Provinsi Riau.
- Fotheringham, S., Brunson, C. & Charlton, M., 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. USA: Wiley.
- Greene, W., 2008. *Functional Forms for The Negative Binomial Model for Count Data, Foundation, and Trends in Econometrics*. New York: New York University.
- Hardin, J. & Hilbe, J., 2007. *Generalized Linear Models and Extensions Second Edition*. Texas: Stata Press.
- Hidayanti, U., 2015. Pemodelan dan Pemetaan Jumlah Kasus Demam Berdarah Dengue (DBD) di Surabaya dengan Geographically Weighted Negative Binomial Regression (GWNBR) dan Flexibly Shaped Spatial Scan Statistic. *Jurnal Sains dan Seni ITS*, 4(2), p. 120.
- Hilbe, J., 2011. *Negative Binomial Regression, Second Edition*. New York: Cambridge University Press.
- Hocking, R. R., 1996. *Methods and Applications of Linear Models: Regression and Analysis of Variance*. New York: John Wiley and Sons.
- Hosmer, D. & Lemeshow, S., 1995. *Applied Logistic Regression*. New York: John Wiley and Sons Inc.
- Kemkes, 2010. Demam Berdarah Dengue. *Buletin Jendela Epidemiologi*, Volume 2, p. 48.
- Kemkes, 2014. *Wujudkan Ketersediaan Air Minum dan Sanitasi Layak bagi Masyarakat*, Jakarta: Pusat Komunikasi Publik Sekretariat Jenderal Kementerian Kesehatan RI.
- McCullagh, P. & Nelder, J. A., 1989. *Generalized Linear Models*. London: Chapman and Hall.
- Myers, R. H., 1990. *Classical and Modern Regression with Application*. Boston: PWS-KENT Publishing Company.
- Notoatmodjo, S., 2003. *Ilmu Kesehatan Masyarakat, Prinsip-Prinsip Dasar*. Jakarta: Rineka Cipta.
- Ricardo, A. & Carvalho, T., 2013. *Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion*. New York: Springer Science.
- Soedarto, 2012. *Demam Berdarah Dengue*. Surabaya: Sagung Seto.
- Tobing, T., 2011. *Pemodelan Kasus Demam Berdarah Dengue (DBD) Di Jawa Timur Dengan Model Poisson Dan Binomial Negatif*. Bogor: Institut Pertanian Bogor.
- Walpole, R., 1995. *Pengantar Metode Statistika. Dalam Edisi Ketiga, Alih Bahasa : Bambang Sumantri*. Jakarta: PT Gramedia Pusaka Utama.