BAYESIAN GEOGRAPHICALLY WEIGHTED REGRESSION DALAM PEMODELAN ANGKA *INCIDENCE RATE*

I Gede Nyoman Mindra Jaya[1)], Neneng Sunengsih[2)]

[1)]*Departemen Statistika, Kampus Unpad Jatinangor, Sumedang; mindra@unpad.ac.id*

[2)]*Departemen Statistika, Kampus Unpad Jatinangor, Sumedang; nenks.stat@unpad.ac.id*

Abstrak

Dalam dunia kesehatan angka *incidence rate* dikenal sebagai angka yang menunjukkan jumlah penderita baru suatu penyakit dalam periode waktu tertentu dibandingkan atas populasi yang beresiko pada periode waktu tersebut. Angka ini umumnya dijadikan rujukan untuk memetakan resiko kesehatan suatu lokasi dan mengkaji lebih jauh faktor-faktor resiko yang berpengaruh pada tinggi rendahnya resiko suatu penyakit pada lokasi tertentu. Pemodelan faktor-faktor resiko terhadap angka *incidence rate* umumnya dilakukan menggunakan metode regresi linear namun pendekatan ini menemukan kendala untuk data yang mengandung heterogenitas spasial, yaitu kondisi dimana faktor resiko memiliki pengaruh yang berbeda untuk setiap lokasi. Sehingga dalam pemodelan yang dinilai lebih tepat adalah pemodelan *Geographically Weighted Regression* (GWR). Metode Bayesian diperkenalkan dalam menaksir parameter GWR untuk memberikan hasil taksiran yang lebih ajeg.

Kata Kunci. Bayesian, GWR, *Incidence Rate*

1. Pendahuluan

Pemetaan lokasi yang memiliki resiko tinggi akan terjadinya penyakit tertentu merupakan langkah bijak dalam upaya menanggulangi penyebaran penyakit dan menekan resiko buruk yang dapat ditimbulkannya (Pringle, 1996). Salah satu upaya yang dapat dilakukan adalah menghitung ukuran-ukuran epidemiology suatu penyakit pada suatu lokasi. Salah satu ukuran yang paling umum digunakan adalah angka *incidence rate*. Dalam dunia kesehatan angka *incidence rate* dikenal sebagai angka yang menunjukkan jumlah penderita baru suatu penyakit dalam periode waktu tertentu dibandingkan atas populasi yang beresiko pada periode waktu tersebut (Rothman dkk, 2008). Angka ini umumnya dijadikan rujukan untuk memetakan resiko kesehatan suatu lokasi dan mengkaji lebih jauh faktor-faktor resiko yang berpengaruh pada tinggi rendahnya resiko suatu penyakit pada lokasi tertentu (Tango, 2010). Pemodelan faktor-faktor resiko terhadap angka *incidence rate* umumnya dilakukan menggunakan metode regresi linear namun pendekatan ini menemukan kendala untuk data yang mengandung heterogenitas spasial, yaitu kondisi dimana faktor resiko memiliki pengaruh yang berbeda untuk setiap lokasi. Oleh karena itu, pemodelan yang dinilai lebih tepat adalah pemodelan *Geographically Weighted Regression* (GWR) (Jaya[a] dkk, (2016), Fotheringham dkk, (2002)). Akan tetapi Model GWR ini kurang baik jika diterapkan untuk kasus dengan adanya data *outlier* dan vairans *error* yang tidak konstan, maka pemodelan yang lebih tepat adalah Bayesian GWR (BGWR).

Metode Bayesian diperkenalkan dalam menaksir parameter GWR untuk memberikan hasil taksiran yang lebih ajeg dibandingkan metode klasik (misalnya: *maximum likelihood, Ordinary Least Square*) karena metode Bayesian tidak memerlukan asumsi asimtotik, yang berarti bahwa ukuran sampel yang besar tidak diperlukan untuk menggambar statistik yang valid (Chumney,2012). Metode Bayesian juga memiliki keunggulan dibandingkan pendekatan klasik karena tidak membutuhkan asumsi normalitas dan mengurangi pengaruh dari heteroskedastisitas spasial melalui substitusi informasi prior.

Sebagai contoh penerapan, metode Bayesian GWR ini diaplikasikan pada pemodelan angka *incidence rate* penyakit demam berdarah (DB) di Kota Bogor. Penyakit DB adalah adalah salah satu penyakit menular yang penularannya melibatkan vektor nyamuk. Bogor merupakan kota Hujan dengan angka kejadian kasus DB sangat tinggi setiap tahunnya. Pada Tahun 2009 tercatat 1540 kasus dari total penduduk 984.780 sehingga angka *incidence* nya mencapai 154 per 100.000 penduduk.

Dinas Kesehatan Pemerintah Kota Bogor telah melakukan berbagai upaya dalam menekan laju peningkatan kejadian Kasus DB seperti pergerakan masyarakat dalam kegiatan Pemberantasan Sarang Nyamuk (PSN) melalui 3M Plus, yaitu menguras dan menyikat bak mandi, menutup tempat penampungan air, mengubur dan memusnahkan barang-barang bekas. Selain itu juga dilakukan program sosialisasi penanganan, pembinaan kader di setiap daerah dan program *fogging*. Untuk meningkatkan efektifitas dari upaya tersebut perlu diketaui faktor-faktor yang paling berpengaruh terhadap tingginya angka kejadian DB di setiap kecamatan di Kota Bogor. Informasi ini penting untuk peringatan dini kepada pemerintah agar dapat mengambil tindakan preventif yang tepat (Jaya[b] dkk, 2017). Tinggi rendahnya angka kejadian kasus DB di suatu wilayah diyakini dipengaruhi oleh banyak faktor diantaranya adalah Angka Bebas Jentik dan jumlah penduduk. Adanya heterogenitas spasial menyebabkan bahwa pengaruh faktor-faktor tersebut mungkin berbeda untuk setiap wilayah.

2. Tinjauan Pustaka

2.1. Angka *Incidence Rate* (IR)

Angka *Incidence Rate* umumnya dijadikan rujukan dalam dunia kesehatan untuk mengetahui situasi dan kondisi kesehatan masyarakat di suatu lokasi. Angka ini merupakan rasio jumlah penderita baru atas total populasi yang beresiko pada periode waktu tertentu.

$$IR = \frac{Total\ Penderita\ Baru}{Total\ Populasi\ Beresiko} \quad (1)$$

Untuk menghindari nilai yang terlalu kecil, umumnya ukuran ini dikalikan dengan 100.000 penduduk.

2.2. *Geographically Weighted Regression* (GWR)

Metode GWR merupakan pengembangan dari metode WLS dengan memperhatikan struktur spasial dalam data (Matthews dan Yang (2012), Nakaya dkk (2005)). Struktur spasial yang dimaksud adalah heteroskedastisitas spasial yang dikenal juga non-stasionaristas spasial. Non-stasionaritas spasial dapat dimaknai sebagai perbedaan hubungan antara variabel dependen dengan independen untuk setiap lokasi yang berbeda. Model GWR juga dikenal sebagai model lokal, yaitu model regresi dengan koefisien regresi yang bervariasi untuk setiap lokasi.

2.2. Model GWR

Misalkan suatu lokasi ditandai oleh suatu penciri (umumnya berupa koordinat *latitude* dan *longitude*) (u_i, v_i) maka model GWR dapat dituliskan sebagai berikut (Fotheringham, dkk, 2002):

$$y(u_i,v_i) = \beta_0(u_i,v_i)+\sum_{k=1}^{K}\beta_k(u_i,v_i)X_{ik}(u_i,v_i)+\varepsilon(u_i,v_i), \quad \varepsilon(u_i,v_i)\sim N(\mathbf{0},\sigma^2\mathbf{I}) \quad (2)$$

dengan $y(u_i,v_i)$ adalah variabel dependen, $X(u_i,v_i)$ menyatakan variabel independen, dan $\beta_0(u_i,v_i)$ dan $\beta_1(u_i,v_i)$ menyatakan parameter regresi GWR pada lokasi ke-*i*.

2.3. Penaksiran Parameter GWR

Penaksiran parameter lokal dalam GWR menggunakan pendekatan *Weighted Least Square* (WLS). Ide dasar dari penaksiran parameter GWR adalah memberikan bobot w_i untuk setiap titik lokasi dengan mempertimbangkan kedekatan jarak antar titik lokasi sehingga dengan koefisien GWR ditaksir secara terpisah untuk setiap lokasi (Fotheringham, dkk, 2002). Taksiran parameter GWR diperoleh dengan prosedur WLS sebagai berikut :

$$\sum_{i=1}^{n} w_i \ \varepsilon_i{}^2 = \sum_{i=1}^{n} w_i \left[Y_i - \beta_0 (u_i, v_i) - \sum_{k=1}^{p} \beta_k (u_i, v_i) X_{ik}\right]^2 \quad (3)$$

dengan taksiran parameter lokalnya adalah sebagai berikut :

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X}'\boldsymbol{W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X}'\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} \quad (4)$$

dengan

$$\boldsymbol{W}(u_i, v_i) = \begin{bmatrix} w_{i1} & 0 & 0 & 0 \\ 0 & w_{i2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_{in} \end{bmatrix}$$

Dengan w_{ij} merupakan pembobot jarak antara lokasi ke-i dengan lokasi lainnya (lokasi ke-j) yang dinormalkan sehingga jumlah vektor baris ($w_{i1}, w_{i2}, \ldots, w_{in}$) = 1, dengan $w_{ii} = 0$. $\hat{\beta}$ (u_i, v_i) memenuhi *sifat Best Linear Unbiased Estimation.*

2.5. Pembobot Model GWR

Fungsi Kernel adalah fungsi pembobot yang umumnya digunakan dalam GWR. Fungsi Kernel diusulkan untuk memberikan pembobot sesuai *bandwidth* optimum yang nilainya bergantung pada variabilitas data. Dua fungsi kernel yang umumnya digunakan yaitu (Fotheringham dkk (2002), Oktavia (2011)):

1. **Fungsi Kernel *Gaussian***

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

2. **Fungsi Kernel *Bi-square***

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

2.6. *Bandwidth* pada GWR

Bandwidth adalah radius atau jumlah pengamatan di sekitar masing-masing titik lokasi dan mengendalikan jarak dalam fungsi pembobotan (Guo dkk, 2008). *Bandwidth* digunakan sebagai dasar menentukan bobot setiap pengamatan terhadap model regresi pada lokasi yang diamati. Penentuan *Bandwidth* yang optimal menemukan dua masalah yaitu Bias dan *Variance*. Semakin kecil *Bandwidth* akan meghasilkan *variance* yang besar dan semakin besar *Bandwidth* akan semakin besar pula bias taksirannya (Fotheringham dkk (2002), Dwinata (2009)).

Fungsi pembobot Kernel mengandung parameter *bandwidth* yang nilainya tidak diketahui. Oleh karena itu maka perlu dilakukan penaksiran terhadap parameter *bandwidth*. *Bandwidth* dapat dimisalkan sebagai radius (b) suatu lingkaran, sehingga sebuah titik lokasi pengamatan yang ada dalam radius lingkaran masih dianggap berpengaruh membentuk parameter di titik lokasi pengamatan ke-*i*. Pemilihan *bandwidth* optimum dalam GWR penting karena akan mempengaruhi ketepatan model terhadap data.

Untuk mendapatkan nilai *bandwidth* optimum dapat dilakukan perhitungan dengan menggunakan metode validasi silang atau *Cross Validation* (CV) sebagai berikut :

$$CV = \sum_{i=1}^{n}[y_i - \hat{y}_{-i}(b)]^2 \tag{7}$$

keterangan:

i = lokasi ke-i

b = *bandwidth*

$\hat{y}_{-i}(b)$= nilai prediksi y_i dengan pengamatan di titik lokasi ke-*I* dihilangkan dari proses penaksiran

Kriteria yang digunakan adalah memilih *bandwidth* dengan *CV* minimum

2.7. Model BGWR

Model GWR memberikan taksiran yang BLUE jika homogenitas *variance error* terpenuhi. Jika terjadi heteroskedastisitas dalam kekeliruan maka penaksir GWR tidak lagi BLUE. Metode alternatif Bayesian dapat digunakan sebagai solusi jika terjadinya masalah heteroskedastisitas dalam kekeliruan (LeSage, 2004). BGWR merupakan pengembangan dari GWR dengan memasukan parameter pemulusan pada penaksiran parameter GWR sehingga parameter regresi BGWR dapat dinyatakan sebagai berikut:

$$\boldsymbol{\beta_i} = (w_{i1} \otimes \boldsymbol{I_k} \cdots w_{in} \otimes \boldsymbol{I_k})\begin{pmatrix}\boldsymbol{\beta_1}\\ \vdots \\ \boldsymbol{\beta_n}\end{pmatrix} + \boldsymbol{u_i} \tag{8}$$

Kekeliruan pada model BGWR diasumsikan mengikuti distribusi normal $\varepsilon_i \sim [0,\sigma^2\mathbf{V}_i]$ dan $\boldsymbol{u_i} \sim N\left[0, \sigma^2\delta^2\left(\boldsymbol{X'W_i^2X}\right)^{-1}\right]$, dengan $\boldsymbol{V}_i$ = diag[v_1,v_2,…,v_n] serta $\frac{r}{v_i} {\sim} X^2{}_{(r)}/r$, σ^2 adalah varian kekeliruan dan $\boldsymbol{V}_i$ adalah matriks diagonal berukuran $n \times n$ yang menunjukkan heteroskedastisitas varian. Sebaran prior $\boldsymbol{V}_i \sim X^2{}_{(r)}$, dimana *r* adalah *hyperparameter* yang mengontrol variabilitas dari $\boldsymbol{V}_i$. Semakin kecil nilai *hyperparameter r* maka keragaman varian antar lokasi semakin besar. Kekeliruan acak $\boldsymbol{u}_i$ sebagai parameter pemulusan diasumsikan menyebar normal dengan rataan nol dan varians $\sigma^2\delta^2\left(\boldsymbol{X'W_i^2X}\right)^{-1}$dengan δ^2 sebagai faktor pembobot yang mengatur nilai $\boldsymbol{\beta_i}$. Prior ini digunakan untuk menunjukkan keragaman parameter penghalus hubungan $\boldsymbol{\beta_i}$ (LeSage, 2004). Jika $n \rightarrow \infty$ ($\boldsymbol{V}_i = \boldsymbol{I}_n$), maka pendugaan BGWR akan menghasilkan pendugaan yang sama dengan GWR. LeSage menunjukkannya pada bentuk persamaan berikut :

$$\boldsymbol{\tilde{y}_i} = \boldsymbol{\tilde{X}_i\beta_i} + \boldsymbol{\varepsilon_i} \tag{9}$$

$$\boldsymbol{\beta_i} = \boldsymbol{J_i\gamma} + \boldsymbol{u_i} \tag{10}$$

dengan:

$\boldsymbol{\tilde{y}_i} = \boldsymbol{W_i y}$

$\boldsymbol{\tilde{X}_i} = \boldsymbol{W_i X}$

$\boldsymbol{J_i} = (w_{i1} \otimes \boldsymbol{I_k} \ \cdots \ w_{in} \otimes \boldsymbol{I_k})$

$$\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\beta_1} \\ \vdots \\ \boldsymbol{\beta_n} \end{pmatrix}$$

Persamaan (10) dapat disusun sebagai berikut:

$$\begin{pmatrix} \widetilde{\boldsymbol{y}}_{\boldsymbol{i}} \\ \boldsymbol{J_i\gamma} \end{pmatrix} = \begin{pmatrix} \widetilde{X}_i \\ -I_k \end{pmatrix} \boldsymbol{\beta_i} + \begin{pmatrix} \boldsymbol{\varepsilon_i} \\ \boldsymbol{u_i} \end{pmatrix} \tag{11}$$

Jika $\boldsymbol{V_i} = \boldsymbol{I_n}$, maka $\widehat{\boldsymbol{\beta}}_i$ dinyatakan sebagai berikut:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} = \boldsymbol{R}\left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{y}}_{\boldsymbol{i}} + \widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}\boldsymbol{J_i\gamma}/\delta^2\right) \tag{12}$$

$$\boldsymbol{R} = \left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}} + \widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}/\delta^2\right)^{-\mathbf{1}} \tag{13}$$

dan jika $\delta \rightarrow \infty$, maka pendugaan BGWR sama dengan pendugaan GWR yang ditunjukkan oleh persamaan berikut:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} = \left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}\right)^{-\mathbf{1}}\left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{y}}_{\boldsymbol{i}}\right) \tag{14}$$

Pendugaan Parameter BGWR

Pendugaan parameter BGWR dapat dilakukan dengan metode *Gibbs Sampling*. Distirbusi posterior gabungan didekati oleh distribusi bersyarat marginal dari setiap parameter yang akan ditaksir. Parameter yang akan ditaksir dalam BGWR adalah $\boldsymbol{\beta_i}$, $\sigma, \delta, \boldsymbol{V_i}$dan $\boldsymbol{\gamma}$.

Distribusi perluang bersyarat dari posterior $\boldsymbol{\beta_i} \mid \sigma_{i,} \delta, \boldsymbol{\gamma}, \boldsymbol{V_i}$ adalah:

$$p\left(\boldsymbol{\beta_i} \mid \sigma_{i,} \delta, \boldsymbol{\gamma}, \boldsymbol{V_i}\right) \propto N\left(\widehat{\boldsymbol{\beta}}_i, \sigma_i^2 \boldsymbol{R}\right) \tag{15}$$

dengan :

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}} = \boldsymbol{R}\left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\boldsymbol{V}_{\boldsymbol{i}}^{-\mathbf{1}}\widetilde{\boldsymbol{y}}_{\boldsymbol{i}} + \widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}\boldsymbol{J_i\gamma}/\delta^2\right)$$

$$\boldsymbol{R} = \left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\boldsymbol{V}_{\boldsymbol{i}}^{-\mathbf{1}}\widetilde{\boldsymbol{X}}_{\boldsymbol{i}} + \widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}/\delta^2\right)^{-\mathbf{1}}$$

Distribusi peluang bersyarat dari $\sigma \mid \boldsymbol{\beta_i}\ \delta, \boldsymbol{\gamma}, \boldsymbol{V_i}$ adalah:

$$p(\sigma \mid \boldsymbol{\beta_i}\ \delta, \boldsymbol{\gamma}, \boldsymbol{V_i}) \propto \sigma_i^{-(m+1)} exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\varepsilon_i' V_i^{-1}\varepsilon_i\right)\right\} \tag{16}$$

dengan m menunjukkan jumlah pengamatan dengan pembobot yang berarti atau tidak bernilai nol.

Distribusi perluang bersyarat dari $\boldsymbol{V_i} \mid \boldsymbol{\beta_i}\ \delta, \sigma, \boldsymbol{\gamma}$ adalah:

Sebaran posterior bersyarat untuk adalah:

$$p\left\{\left[\left(\frac{e_i^2}{\sigma^2}\right) + r\right] / \boldsymbol{V_i} \mid \boldsymbol{\beta_i}\ \delta, \sigma, \boldsymbol{\gamma}\right\} \propto \chi^2{}_{r+1} \tag{17}$$

Distribusi perluang bersyarat dari $\delta \mid \boldsymbol{\beta_i}, \sigma, \boldsymbol{\gamma}$ adalah $\chi^2_{(nk)}$ yang ditunjukkan pada persamaan di atas .

$$p\left(\delta \,\middle|\, \boldsymbol{\beta_i}, \sigma, \boldsymbol{\gamma}\right) \propto \delta^{-nk}\ exp\left\{-\sum_{i}^{n} = 1\, \frac{(\boldsymbol{\beta_i} - \boldsymbol{J_i\gamma})'\left(\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}'\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}\right)^{-1}(\boldsymbol{\beta_i} - \boldsymbol{J_i\gamma})}{2\sigma^2\delta^2}\right\} \tag{18}$$

Tahapan proses *Gibbs Sampling* dalam penaksiran parmeter Bayesian GWR adalah sebagai berikut (LeSage, 2004) :

1. Menentukan nilai secara acak untuk parameter $\boldsymbol{\beta}_i^0, \sigma^0, \delta^0, \boldsymbol{V_i}^0, \boldsymbol{\gamma}^0$
2. Tiap observasi i = 1,…, n,
   a. Bangkitkan $\boldsymbol{\beta}_i{}^1$dari $P(\boldsymbol{\beta}_i|\sigma^0, \delta^0, \boldsymbol{V}_i^0, \boldsymbol{\gamma^0})$
   b. Bangkitkan σ^1dari $P(\sigma|\delta^0, \boldsymbol{\beta}_i^1, \boldsymbol{V}_i^0, \boldsymbol{\gamma^0})$
   c. Bangkitkan$\boldsymbol{V}_i^1$dari $P(\boldsymbol{V}_i|\boldsymbol{\beta}_i^1, \sigma^1, \delta^0, \boldsymbol{\gamma^0})$
3. Gunakan nilai$\boldsymbol{\beta}_i^1$, $i = 1, \cdots, n$ untuk memperbaharui$\boldsymbol{\gamma}^0$ menjadi $\boldsymbol{\gamma}^1$
4. Nilai δ^1 diperoleh dari$P(\delta|\boldsymbol{\beta_i^1}, \sigma^1, \boldsymbol{V_i}^{\boldsymbol{1}}, \boldsymbol{\gamma^1})$
5. Ganti nilai $\boldsymbol{\beta}_i^0, \sigma^0, \delta^0, \boldsymbol{V_i}^0, \boldsymbol{\gamma}^0$ pada langkah 1 dengan$\boldsymbol{\beta}_i^1, \sigma^1, \delta^1, \boldsymbol{V_i}^1, \boldsymbol{\gamma}^1$
6. Ulangi langkah 1-5 sebanyak q bangkitan hingga mendekati konvergen.

Dugaan parameter diperoleh dari rataan sampel posterior.

3. Hasil dan Pembahasan

3.1. Deskripsi Variabel Penelitian

Variabel yang dipergunakan dalam penelitian ini meliputi tingkat kejadian kasus (IR) per 100.000 penduduk, Angka Bebas Jentik (%). Angka tingkat kejadian kasus DB pada tahun 2009 rata-rata nya mencapai 157 kasus dengan nilai maksimum mencapai 79. Ini artinya bahwa untuk 100.000 penduduk di kota Bogor secara rata-rata terjadi sebanyak 157 kasus, angka Ini masuk kategori yang tinggi. Angka bebas jentik cukup tinggi mencapai 92.24%.

3.2. Pemodelan Angka *Incidence Rate*

3.2.1. Model Regresi Linier Multipel

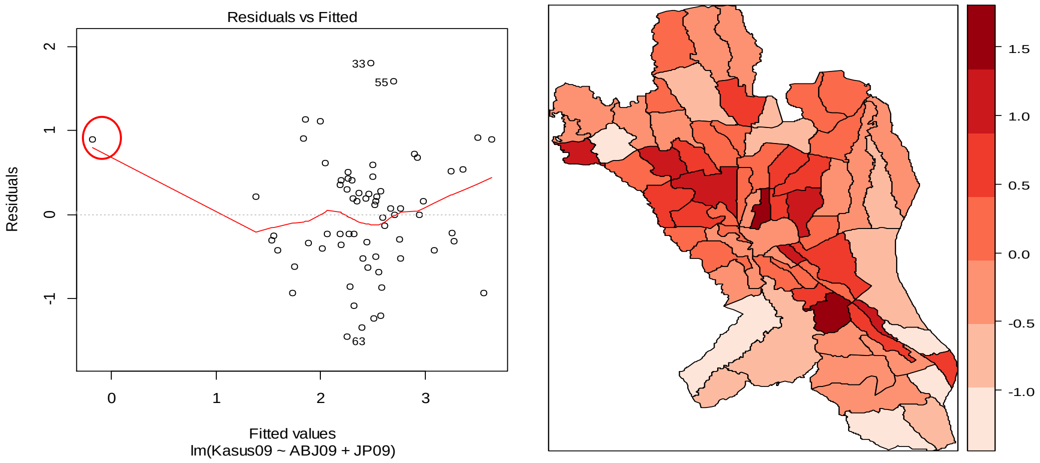
Tahapan sebelum dilakukan pemodelan Bayesian GWR adalah perhitungan model regresi linear klasik untuk mendapatkan taksiran parameter Global dari GWR dan nilai residual yang akan diuji apakah ada pelanggaran asumsi baik *outlier* ataupun heteroskedastisitas. Hasilnya disajikan dalam Tabel berikut :

Tabel 1. Taksiran Parameter Model Regresi Linier Multipel

Variabel		Estimate	Std. Error	t value	Pr(>|t|)
Intersep		2.69E+01	3.56E+00	7.544	1.89E-10
Angka Bebas Jentik		-2.60E-01	3.81E-02	-6.825	3.53E-09
Jumlah Penduduk		-3.40E-05	1.17E-05	-2.914	0.00488
R^2	: 0.4233	F	: 23.86		
		Df	: 2 & 65		
		p-value	: 1.699e-08		

*Outpu*t di atas menjelaskan bahwa angka bebas jentik dengan koefisien -0.260 menunjukkan bahwa semakin tinggi angka bebas jentik, maka tingkat kejadian kaus DB akan semakin turun. *Effect* dari Angka bebas jentik singifikan pada level 5%.

Jumlah penduduk juga memberikan pengaruh negative. Semakin banyak penduduk maka jumlah kasus semakin rendah. Nilai R^2 dari model sebesar 42.33%, artinya bahwa model ini menjelaskan sebesar 42.33% variansi dari angka kejadian kasus dijelaskan oleh angka bebas jentik dan jumlah penduduk. Namun sebelum diambil kesimpulan dari hasil ini dilakukan analisis residual untuk menilai apakah model ini tepat atau tidak.



(a) Residual dengan Fitted Plot (b) Spasial Residual Plot

Gambar 1. Plot Residual

Plot residual di atas pada bagian (a) menunjukkan adanya pola heteroskedastisitas dan diperkuat dengan gambar bagian (b) yang menyatakan adanya pola spasial. Dimana nilai data yang tinggi cenderung mengelompok begitu juga untuk data residual yang rendah. Sehingga untuk pemodelan regresi dinilai lebih tepat menggunakan model regresi yang memperhatikan adanga heterogenitas antara lokasi yaitu model *Geographically Weighted Regression*.

Tabel 2. Taksiran Parameter GWR

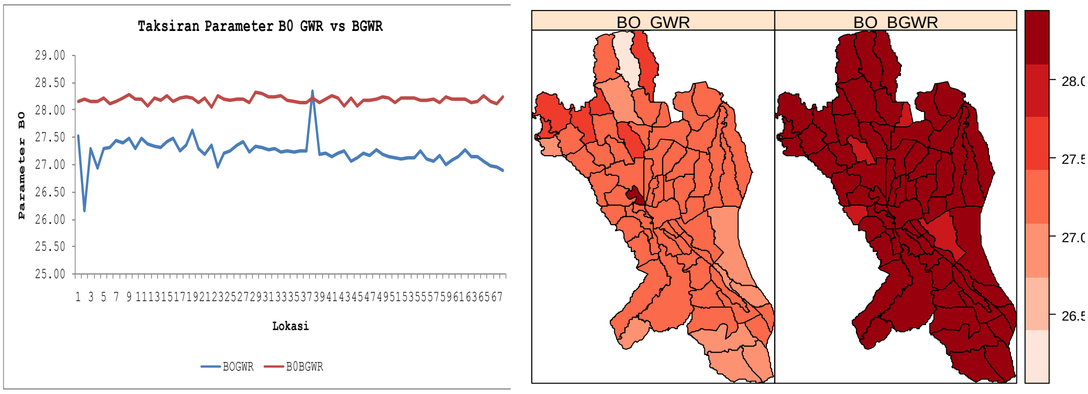
```
***************************************************************************
 << Geographically varying (Local) coefficients >>
***************************************************************************


Summary statistics for varying (Local) coefficients
Variable                      Mean             STD
-------------------- --------------- ---------------
Intercept                26.825358        0.460766
AB                       -0.259429        0.004495
JP                       -0.000035        0.000001

Variable                       Min             Max           Range
-------------------- --------------- --------------- ---------------
Intercept                26.156400       28.343638        2.187238
AB                       -0.274776       -0.252709        0.022066
JP                       -0.000039       -0.000031        0.000008

Variable               Lwr Quartile          Median    Upr Quartile
-------------------- --------------- --------------- ---------------
```

```
Intercept                 27.133738      27.239049      27.318746
AB                        -0.264251      -0.263428      -0.262369
JP                        -0.000035      -0.000035      -0.000035

Variable            Interquartile R     Robust STD
-------------------- --------------- ---------------
Intercept                  0.185007       0.137144
AB                         0.001881       0.001395
JP                         0.000001       0.000001
 (Note: Robust STD is given by (interquartile range / 1.349))

****************************************************************************
 GWR ANOVA Table
****************************************************************************
Source                          SS          DF            MS           F
----------------- ------------------- ---------- --------------- ---
Global Residuals              30.249     65.000
GWR Improvement                0.223      0.534          0.418
GWR Residuals                 30.027     64.466          0.466   0.896602

****************************************************************************
```

Untuk taksiran Global dari GWR sama dengan taksiran regresi dengan metode OLS. Terlihat bahwa nilai R^2 model regresi GWR lebih baik dibandingkan dengan model OLS. Model GWR memberikan nilai R^2 sebesar 42.75. Hasil ini menginformasikan bahwa sebesar 42.75% keragaman pada tingkat kejadian kasus DB di Kota Bogor dijelaskan oleh variabel angka bebas jentik dan jumlah penduduk. Berikut ini disajikan grafik koefisien regresi untuk masing-masing lokasi. Berdasarkan plot (a) Gambar 1, menunjukkan adanya *outlier* sehingga dilakukan pemodelan dengan Bayesian GWR. Selain itu, pemodelan spasial dengan variansi lokasi dengan jumlah penduduk berbeda membutuhkan suatu model pemulusan Bayesian. Hasilnya adalah sebagai berikut :

Tabel 3. Taksiran Parameter Bayesian GWR

```
***************************************************************
Bayesian geographically weighted regression model
***************************************************************
R-squared                = 0.4212
Nobs, Nvars              = 68, 3
ndraws,nomit             = 10000, 100
r-value                  = 4.0000
delta-value              = 50.0111
gam(m,k) d-prior         = 50, 2
Bandwidth                = 4.4718
Decay type               = exponential
prior type               = distance
```

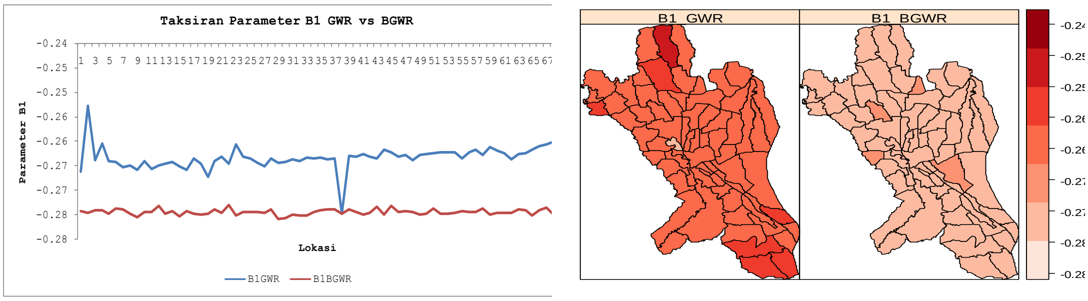


(a) Perbandingan Line Plot GWR dengan BGWR untuk Taksiran Parameter β_0

(b) Perbandingan Spasial Plot GWR dengan NGWR untuk Taksiran Parameter β_0

Gambar 2. Perbandingan Taksiran Parameter β_0

Bayesian GWR secara umum memberikan taksiran parameter intersep yang lebih besar daripada GWR, namun dengna variansi yang lebih minimum. Seanjutnya **Gambar 3** menunjukkan perbandingan taksiran parameter β_1 model GWR dengan Bayesian GWR.

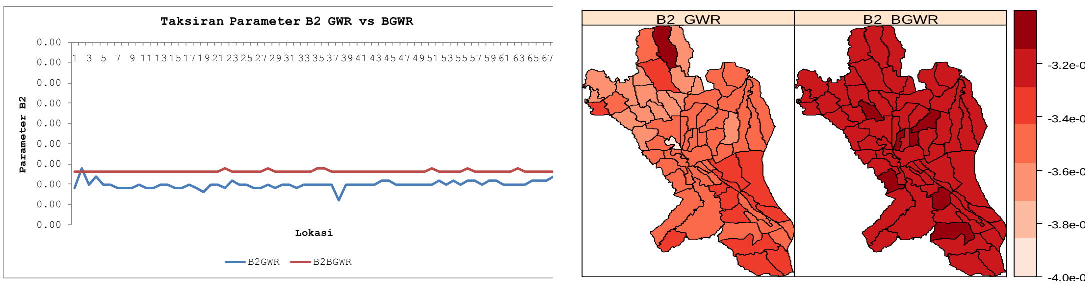


(a) Perbandingan Line Plot GWR dengan BGWR untuk Taksiran Parameter β_1

(b) Perbandingan Spasial Plot GWR dengan NGWR untuk Taksiran Parameter β_1

Gambar 3. Perbandingan Taksiran Parameter β_1

Terlihat secara umum Bayesian GWR memberikan taksiran parameter slop untuk angka bebas jentik yang lebih kecil dibandingkan GWR, namun dengan variansi yang lebih minimum.



(a) Perbandingan Line Plot GWR dengan BGWR untuk Taksiran Parameter β_2

(b) Perbandingan Spasial Plot GWR dengan NGWR untuk Taksiran Parameter β_2

Gambar 4. Perbandingan Taksiran Parameter β_2

Bayesian GWR secara umum memberikan taksiran parameter slop untuk jumlah penduduk yang lebih besar daripada GWR, namun dengan variansi yang lebih minimum
Berdasarkan peta - peta di atas dapat dicermati bahwa secara umum ada hubungan signifikan antara angka bebas jentik dan jumlah penduduk dengan angka kejadian kasus DB di Kota Bogor.

4. Simpulan dan Saran

Metode Bayesian GWR memberikan hasil yang terlihat berbeda dibandingkan dengan GWR. Bayesian GWR memuluskan taksiran parameter regresi sehingga memiliki variansi yang lebih kecil dibandingkan metode GWR. Kondisi ini bisa dipandang dari dua sudut pandang yang bebeda. Pertama kondisi ini sangat tepat jika prior yang digunakan dalam pemodelan Bayesian merefleksikan informasi yang akurat yang dimiliki oleh peneliti yang diperoleh dari sumber informasi yang valid. Kedua, model yang dihasilkan mungkin saja terlalu mulus (*over smooth*). Pada penelitian ini peneliti menggunakan keyakinan bahwa andaikan efek dari variabel Angka Bebas Jentik dan Jumlah Penduduk berbeda untuk setiap lokasi, namun perbedaan tersebut seharusnya tidak terlalu besar mengingat karakteristik kecamatan di Kota Bogor tidak terlalu berbeda dari sisi Jumlah penduduk dan kemampuan mengelola angka bebas jentik.

Berdasarkan pemodelan secara spasial dapat diketahui lebih jelas pola spasial yang terjadi pada data. Analisis regresi BGWR membantu dalam mencermati pola spasial pengaruh dari masing-masing variabel secara lokal, Tinggi Rendahanya Kasus DB di Kota Bogor secara signifikan dipengaruhi oleh Angka Bebas Jentik dan Juga Jumlah penduduk, Tingginya angka kasus DB di suatu kecamatan tidak berarti bahwa penderita terinfeksi DB di kecamatan tersebut namun dapat terinfeksi di Kecamatan tempat bekerja atau sekolah. Pengembangan model masih perlu dilakukan untuk kondisi data dengan *outlier* juga untuk kasus beberapa lokasi yang nilai *incidence rate* nya nol. Karena pemodelan standar kurang tepat pada kondisi data seperti tersebut.

Daftar Pustaka

Chumney, Frances L. 2012. *Comparison of Maximum Likelihood, Bayesian, Partial Least Squares, and Generalized Structured Component Analysis Methods for Estimation of Structural Equation Models with Small Samples: An Exploratory Study*. Public Access Theses and Dissertations from the College of Education and Human Sciences. 145. http://digitalcommons.unl.edu/cehsdiss/145

Fotheringham, A stewart *dkk*. 2002. *Geographically weighted Regression The Analysis Of Spasially Verying Relationship*. United Kingdom : University of Newcastle.

Fotheringham, A., Brunsdon, C., & Charlton, M. 2002. *Geographically weighted regression: the analysis of spatially*. New York: John Wiley and Sons.

Guo, L., Ma, Z., & Zhang, L. 2008. Comparison of bandwidth selection in application of geographically weighted regression: a case study. *Can. J. For. Res* , 38, 2526–2534.

Jaya, I. G., Abdullah, A. S., Hermawan, E., & Ruchjana, B. N. (2016). Bayesian Spatial Modeling and Mapping of Dengue Fever: A Case Study of Dengue Fever in The City of Bandung, Indonesia. *International Journal of Applied Mathematics and Statistics* , 54 (3), 94-103.

Jaya, I. G., Folmer, H., Ruchjana, B. N., Kristiani, F., & Andriyana, Y. 2017. *Modeling of Infectious Diseases: A Core Research Topic for the Next Hundred Years*. Dalam R. Jackson, & P. Schaeffer, Regional Research Frontiers-Vol 2 (hal. 239-255). United State: Springer International Publishing.

LeSage, J. P. (2004). A Family of Geographically Weighted Regression Models. Dalam L. Anselin, R. Florax, & S. J. Rey, *Advances in Spatial Econometrics: Methodology, Tools and Applications* (hal. 241-263). Berlin: Springer.

Matthews, S. A., & Yang, T. C. 2012. Mapping the results of local statistics: Using geographically weighted regression. *Demographic Research* , 26 (6), 151-166.

Nakaya, T., Fotheringham, A., Brunsdon, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statist. Med.* , 24, 2695–2717.

Oktavia, Mega. 2011. *Estimasi Model Linier Spasial Dengan Geographically Weighted Weighed Poisson Regression (GWPR).* Tugas Akhir Matematika-Fakultas SAINS dan Teknologi UIN, Malang.

Pringle, D. 1996. Mapping Disease Risk Estimates Based on Small Numbers: An Assessment of Empirical Bayes Techniques. *The Economic and Social Review* , 27 (4), 341-363.

Rothman, K. J., Greenland, S., & Lash, T. L. 2008. *Modern Epidemiology, 3rd Edition*. California: Lippincott Williams & Wilkins.

Schabenberger, O. & Gotway, C.A. 2005. *Statistical Methods for Spasial Data Analysis*, Chapman & Hall/CRC.

Tango, T. 2010. *Statistical Methods for Disease Clustering*. Japan: Springer.