

Analisis Faktor-Faktor Risiko yang Mempengaruhi Angka Kesakitan Malaria di Provinsi Jawa Timur dengan Metode *Geographically Weighted Regression*

Theresia Widiastuti dan Puhadi

Departemen Statistika, FMKSD, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: purhadi@statistika.its.ac.id

Abstrak— Sebagai kelanjutan dari program nasional untuk Indonesia Bebas Malaria di tahun 2030, Dinas Kesehatan Republik Indonesia memiliki target bahwa Pulau Jawa dan Bali sebagai daerah yang bukan endemik malaria akan bebas dari penyakit malaria terlebih dahulu pada tahun 2015. Jawa Timur sebagai salah satu provinsi di Pulau Jawa dengan jumlah penduduk tertinggi kedua di Indonesia setelah Jawa Barat (BPS, 2010), walaupun dengan jumlah kasus malaria yang tergolong rendah namun masih tetap memiliki beberapa kasus malaria bahkan masih ditemukan terdapat korban jiwa. Berdasarkan data dari Dinas Kesehatan Provinsi Jawa Timur, pada tahun 2009 ditemukan sebanyak 1489 kasus, 2010 ditemukan 947 kasus, pada tahun 2011 meningkat menjadi 1222 kasus, dan 2012 terjadi peningkatan, sehingga mencapai 1320 kasus. Disebutkan juga bahwa pada tahun 2013 penyakit malaria telah menimbulkan 7 kematian (Dinas Kesehatan Provinsi Jawa Timur 2013). Berdasarkan hasil analisis pada aspek spasial data tidak terdapat heterogenitas spasial menggunakan uji *Breusch Pagan* dengan *P-value* sebesar sebesar 0,0003589. Fungsi pembobot GWR yang digunakan adalah fungsi pembobot Fixed Gaussian yaitu dengan CV minimum sebesar 135,738 dan *bandwidth* 0,500. Faktor signifikan yang mempengaruhi Angka Kesakitan Malaria adalah rasio konfirmasi laboratorium (mikroskop), rasio penderita diobati dengan ACT + Primaquine, rasio hasil pengobatan penderita (*follow up* tidak lengkap), rasio jenis parasit Pf, dan rasio malaria tanpa komplikasi (rawat jalan).

Kata Kunci—Angka Kesakitan Malaria, Dependensi Spasial, Heterogenitas Spasial.

I. PENDAHULUAN

Malaria merupakan salah satu penyakit parasit yang tersebar luas dan hampir ditemukan di seluruh bagian dunia termasuk Indonesia. Persebaran penyakit malaria banyak ditemukan di negara-negara yang beriklim tropis dan subtropis. Penduduk yang berisiko terkena berjumlah sekitar 2,3 milyar orang atau 41% dari jumlah penduduk dunia. Kasus malaria yang ditemukan berjumlah 300-500 juta kasus dan mengakibatkan 1,5 – 2,7 juta kematian, terutama di negara-negara yang terdapat di Benua Asia dan Afrika. Penyakit malaria disebabkan oleh parasit *Plasmodium*, yaitu makhluk hidup bersel satu yang termasuk dalam kelompok *protozoa*. Malaria ditularkan melalui gigitan nyamuk *Anopheles* betina yang mengandung *plasmodium* di dalamnya, di mana banyak spesiesnya ditemukan di Indonesia. *Plasmodium* tersebut kemudian hidup dan berkembang biak dalam sel darah manusia. Jumlah spesies *Anopheles* di permukaan bumi telah ditemukan kurang lebih 424 spesies, dan 70 di antaranya diketahui sebagai

vektor malaria. Sebagai kelanjutan dari program nasional untuk Indonesia Bebas Malaria di tahun 2030, Dinas Kesehatan Republik Indonesia memiliki target bahwa Pulau Jawa dan Bali sebagai daerah yang bukan endemik malaria akan bebas dari penyakit malaria terlebih dahulu pada tahun 2015. Jawa Timur sebagai salah satu provinsi di Pulau Jawa dengan jumlah penduduk tertinggi kedua di Indonesia setelah Jawa Barat, walaupun dengan jumlah kasus malaria yang tergolong rendah namun masih tetap memiliki beberapa kasus malaria bahkan masih ditemukan terdapat korban jiwa. Berdasarkan kondisi tersebut, tentu perlu dilakukan kebijakan-kebijakan atau upaya demi membasmi atau setidaknya mengurangi kasus malaria di Jawa Timur seperti dengan melakukan pemeriksaan maupun penyuluhan tentang kondisi lingkungan serta berbagai usaha lain pada masing-masing daerah atau kabupaten pada kasus malaria dengan angka yang tinggi. Maka dari itu dalam penelitian ini dilakukanlah pemetaan kabupaten dan kota berdasarkan kasus faktor-faktor penyebab angka kesakitan malaria. Penelitian mengenai malaria telah banyak dilakukan sebelumnya terutama pada daerah Indonesia yang endemik terhadap malaria. Penelitian tentang kasus malaria telah dilakukan oleh beberapa pihak. Menurut Haryati [1], dengan menggunakan metode *Geographically Weighted Negative Binomial Regression* (GWNBR), variabel global yang berpengaruh pada kasus malaria yaitu persentase rumah sehat, persentase rumah tangga dengan pembuangan limbah sehat, persentase puskesmas, persentase penyuluhan kesehatan, persentase rumah tangga dengan perilaku hidup bersih dan sehat (PHBS), persentase rumah tangga dengan sarana air bersih, dan persentase TUPM sehat. Sedangkan menurut Lestari [2], jumlah ibu hamil yang berada di Provinsi Nusa Tenggara Barat, Nusa Tenggara Timur, Maluku, Maluku Utara, Papua, dan Papua Barat adalah sebanyak 274 orang dengan 31% (84 orang) diantaranya terjangkit penyakit malaria. Faktor-faktor yang memiliki pengaruh signifikan terhadap penyakit malaria pada ibu hamil adalah rumah panggung, atap ijuk/rumbia, atap seng, dan lantai semen plesteran retak.

II. URAIAN PENELITIAN

A. Statistika Deskriptif

Statistika deskriptif adalah analisis yang berkaitan dengan pengumpulan dan penyajian data sehingga dapat memberikan informasi yang berguna. Analisis ini bertujuan

menguraikan tentang sifat-sifat atau karakteristik dari suatu keadaan dan untuk membuat deskripsi atau gambaran yang sistematis dan akurat mengenai fakta-fakta, sifat-sifat dari fenomena yang diselidiki. Contoh dari statistika deskriptif adalah tabel frekuensi, tabulasi silang [3]. Untuk penelitian ini digunakan ukuran minimum, maksimum, *mean*, dan varians.

B. *VIF (Variance Inflation Factor)*

VIF (Variance Inflation Factor) adalah salah satu statistik yang dapat digunakan untuk mendeteksi gejala multikolonieritas (*multicollinearity, collinearity*) pada analisis regresi yang akan dibentuk. *VIF* tidak lain adalah mengukur keeratan hubungan antar variabel bebas (*X*). Jika dalam perhitungan *VIF* didapat hasil di atas 10 maka menunjukkan adanya multikolonieritas antar variabel predictor dengan rumus *VIF* sebagai persamaan (2.2).

$$VIF_j = \frac{1}{1 - R_k^2}$$

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

C. *Pengujian Asumsi Residual Berdistribusi Normal*

Pengujian asumsi residual memiliki tujuan untuk mengetahui kelayakan suatu model. Dalam menggunakan metode *Geographically Weighted Regression* diperlukan data yang memenuhi asumsi residual berdistribusi Normal [4].

Untuk mengetahui hal ini dilakukan pengujian dengan uji statistik Kolmogorov-Smirnov dengan hipotesis sebagai berikut.

$$H_0: F_n(y) = F_0(y) \text{ (residual berdistribusi Normal)}$$

$$H_0: F_n(y) \neq F_0(y) \text{ (residual tidak berdistribusi Normal)} \tag{2.3}$$

Statistik uji

$$D = \text{Sup} |F_n(x) - F_0(x)|$$

dimana

$F_n(x)$: nilai distribusi kumulatif sampel

$F_0(x)$: nilai distribusi kumulatif bawah x untuk distribusi Normal ($P(Z < Z_i)$)

Daerah penolakan :

Tolak H_0 jika $D > D_\alpha$ dimana D_α merupakan nilai kritis untuk uji Kolmogorov-Smirnov satu sampel atau jika $P\text{-value} < \alpha$.

D. *Heterogenitas Spasial*

Pengujian ini dilakukan untuk mengetahui apakah terdapat karakteristik atau keunikan sendiri di setiap lokasi pengamatan. Adanya heterogenitas spasial dapat menghasilkan parameter regresi yang berbeda-beda di setiap lokasi pengamatan. Heterogenitas spasial dapat diuji dengan menggunakan statistik uji Breusch-Pagan yang mempunyai hipotesis sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma^2$ (tidak terdapat heterogenitas spasial)

$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma^2$ (ada heterogenitas spasial)

Statistik uji :

$$BP = \left(\frac{1}{2}\right) f^T Z(Z^T Z)^{-1} Z^T f$$

Dengan elemen vektor f adalah :

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$$

Dimana e_i^2 merupakan residual *least square* untuk observasi ke- i dan Z merupakan matriks berukuran $n \times (p+1)$ yang berisi vektor yang sudah dinormalstandartkan untuk tiap observasi.

Daerah kritis : Tolak H_0 jika $BP > X_{\alpha;p}^2$

Dalam menggunakan metode *Geographically Weighted Regression* terlebih dahulu diperlukan data yang memenuhi asumsi aspek spasial. Model *Geographically Weighted Regression* dapat ditulis sebagai persamaan (2.7) [4].

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon_i$$

Dimana : $i=1,2,\dots,n$

y_i : Nilai observasi variabel respon untuk lokasi ke- i

x_{ik} : Nilai observasi variabel prediktor ke- k pada lokasi pengamatan ke- i , $k=1,2,\dots,p$

$\beta_0(u_i, v_i)$: Nilai *intercept* model regresi GWR

$\beta_k(u_i, v_i)$: Koefisien regresi variabel prediktor ke- k pada lokasi pengamatan ke- i

(u_i, v_i) : Koordinat letak geografis (lintang, bujur) dari lokasi pengamatan ke- i

ε_i : error pengamatan ke- i yang diasumsikan identik, independen dan berdistribusi normal dengan mean nol dan varian konstan σ^2 .

E. *Pembobot Model Geographically Weighted Regression*

Peran pembobot pada model *Geographically Weighted Regression* sangat penting karena nilai pembobot yang mewakili letak data observasi satu dengan lainnya. Besarnya pembobot untuk tiap lokasi yang berbeda dapat ditentukan salah satunya dengan menggunakan fungsi kernel (*kernel function*). Fungsi kernel digunakan untuk mengestimasi parameter dalam model *Geographically Weighted Regression* jika fungsi jarak (w_j) adalah fungsi yang kontinu. Pembobot yang terbentuk dari fungsi kernel dapat ditulis sebagai berikut [5].

1. Fungsi *Gaussian*

$$w_{ij}(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right)$$

2. Fungsi *Bisquare*

$$w_{ij}(u_i, v_i) = \begin{cases} 1 - \left(\frac{d_{ij}}{b}\right)^2 & , \text{ untuk } d_{ij} \leq b \\ 0 & , \text{ untuk } d_{ij} > b \end{cases}$$

dengan

b : Parameter non negatif yang disebut dengan parameter penghalus (*bandwidth*)

$b_{(p)}$: *Bandwidth* adaptif yang menetapkan p sebagai jarak terdekat (*nearest neighbour*) dari lokasi i

d_{ij} : jarak *Euclidean* antara lokasi (u_i, v_i)

$$\text{dimana } d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$$

Pemilihan *bandwidth* optimum menjadi sangat penting karena akan mempengaruhi ketepatan model terhadap data, yaitu mengatur varians dan bias dari model. Oleh karena itu, digunakan metode *Cross Validation* (CV) untuk menentukan *bandwidth* optimum, yang dirumuskan sebagai persamaan (2.12) [4].

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

Dimana

$\hat{y}_{\neq i}(b)$ = nilai estimasi y_i dimana pengamatan lokasi (u_i, v_i) dihilangkan dari proses estimasi

n = jumlah sampel

Bandwidth yang optimal ditunjukkan dengan nilai CV minimum.

F. Pengujian Model *Geographically Weighted Regression*

Pengujian hipotesis pada model *Geographically Weighted Regression* terdiri dari uji kesesuaian antara model regresi linier dengan model *Geographically Weighted Regression* dan uji parsial model *Geographically Weighted Regression*.

a. Uji Kesesuaian Model *Geographically Weighted Regression*

Pengujian kesesuaian model dilakukan dengan menguji kesesuaian dari koefisien parameter secara serentak, yaitu dengan mengkombinasikan uji regresi linier dengan model untuk data spasial. Sehingga diperoleh model yang paling sesuai untuk menggambarkan data yang diperoleh. Berikut adalah langkah analisisnya [4].

Hipotesis:

H_0 : $\beta_k(u_i, v_i) = \beta_k, k = 1, 2, \dots, p, i = 1, 2, \dots, n$ (tidak ada perbedaan signifikan antara model regresi linier dengan model *Geographically Weighted Regression*)

H_1 : minimal ada satu $\beta_k(u_i, v_i) \neq \beta_k$ (ada perbedaan signifikan antara regresi linier dan *Geographically Weighted Regression*)

Penentuan statistik uji dengan cara menurunkan rumus *Sum Square Error* (SSE) yang diperoleh masing-masing dibawah H_0 dan H_1 . Jika dibawah kondisi H_0 maka

menggunakan metode OLS untuk mendapatkan nilai SSE seperti berikut.

$$SSE(H_0) = \hat{\varepsilon}^T \hat{\varepsilon} = (y - \hat{y})^T (y - \hat{y}) = y^T (I - H)y$$

Dengan $H = X(X^T X)^{-1} X^T$ yang bersifat idempotent, yaitu matriks bujur sangkar di mana berlaku $A^2 = A$ atau $A^n = A$ untuk suatu n . Dibawah kondisi H_1 , koefisien regresi yang bervariasi secara spasial ditentukan dengan metode *Geographically Weighted Regression*, sehingga diperoleh SSE sebagai berikut

$$SSE(H_0) = \hat{\varepsilon}^T \hat{\varepsilon} = (y - \hat{y})^T (y - \hat{y}) = y^T (I - L)^T (I - L)y$$

Dimana

$$L = \begin{pmatrix} x_1^T (X^T W(u_1, v_1) X)^{-1} X^T W(u_1, v_1) \\ x_2^T (X^T W(u_2, v_2) X)^{-1} X^T W(u_2, v_2) \\ \vdots \\ x_n^T (X^T W(u_n, v_n) X)^{-1} X^T W(u_n, v_n) \end{pmatrix}$$

L adalah matriks berukuran $n \times n$ dan I merupakan matriks identitas yang berukuran $n \times n$.

$$R_0 = (1 - H)^T (1 - H)$$

$$R_1 = (1 - L)^T (1 - L)$$

Sehingga diperoleh statistik uji sebagai berikut [4]

$$F = \frac{(SSE(H_0) - SSE(H_1))}{\frac{SSE(H_1)}{\delta_1}}$$

Daerah penolakan H_0 adalah apabila $F_{hitung} > F_{(1-\alpha; df_1; df_2)}$, yang artinya ada perbedaan signifikan antara regresi linier dan GWR.

Dengan

$$V = tr(R_0 - R_1); \delta_1 = tr(R_1); \delta_2 = tr(R_1^2)$$

$$df_1 = \frac{v^2}{v^*}; df_2 = \frac{\delta_1^2}{\delta_1}$$

$$v^* = tr[(R_0 - R_1)^2]$$

b. Uji Parsial Model *Geographically Weighted Regression*

Pengujian parameter model *Geographically Weighted Regression* dilakukan untuk mengetahui parameter yang mempengaruhi secara signifikan terhadap variabel respon secara parsial. Langkah pengujiannya adalah sebagai berikut.

Hipotesis:

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

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

Estimasi parameter $\beta(u_i, v_i)$ akan mengikuti distribusi normal dengan rata-rata $\beta(u_i, v_i)$ dan matriks kovarians $CC^T \sigma^2$ dengan $C = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i)$, sehingga didapatkan $\frac{\hat{\beta}_k(u_i, v_i) - \beta_k(u_i, v_i)}{\sigma \sqrt{c_{kk}}} \sim N(0, 1)$, dimana c_{kk} adalah elemen diagonal ke- k dari matriks CC^T

Statistik uji:

$$t = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma} \sqrt{c_{kk}}}$$

Daerah penolakan H_0 adalah jika $|t_{hitung}| > t_{\alpha/2, (\delta_1^2/\delta_2)}$ atau jika $Pvalue < \alpha$ yang artinya parameter variabel prediktor berpengaruh signifikan terhadap variabel respon.

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data yang digunakan diperoleh dari Dinas Kesehatan Provinsi Jawa Timur.

B. Variabel Penelitian

Dalam metode *Geographically Weighted Regression*, setiap wilayah pengamatan model yang berbeda-beda. Oleh karena itu, di dalam penelitian ini variabel yang digunakan berasal dari 38 kabupaten/kota di Provinsi Jawa Timur. Macam-macam variabel yang digunakan di setiap kabupaten/kota di Jawa Timur serta skala variabel yang digunakan dapat dilihat pada Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Kode	Variabel
Y	Angka Kesakitan Malaria
X ₁	Rasio penderita malaria positif
X ₂	Rasio Konfirmasi Laboratorium (Mikroskop)
X ₃	Rasio Klasifikasi Asal Penularan (Impor)
X ₄	Rasio Penderita diobati dengan ACT + Primaquine
X ₅	Rasio Hasil pengobatan penderita (Sembuh)
X ₆	Rasio Hasil pengobatan penderita (<i>follow up</i>)
X ₇	Rasio Jenis parasit Pf
X ₈	Rasio Jenis Parasit Pv
X ₉	Rasio Malaria tanpa komplikasi (rawat inap)
X ₁₀	Rasio Malaria tanpa komplikasi (rawat jalan)

C. Struktur Data

Pada penelitian ini digunakan metode GWR untuk memodelkan faktor-faktor yang mempengaruhi nilai Angka Kesakitan Malaria di Provinsi Jawa Timur. Terdapat satu variabel respon (Y) dan 10 variabel prediktor. Struktur data dalam penelitian ini ditampilkan pada Tabel 3.2 sebagai berikut.

Tabel 3.2 Struktur Data

Kabupaten/Kota	u_i	v_i	Y	X ₁	X ₂	...	X ₁₀
Pacitan	u_1	v_1	Y ₁	X ₁₁	X ₂₁	...	X _{10;1}
Ponorogo	u_2	v_2	Y ₂	X ₁₂	X ₂₂	...	X _{10;2}
Trenggalek	u_3	v_3	Y ₃	X ₁₃	X ₂₃	...	X _{10;3}
...
Batu	u_{38}	v_{38}	Y ₃₈	X _{1;38}	X _{2;38}	...	X _{10;38}

D. Langkah Analisis

Metode yang digunakan pada penelitian ini adalah *Geographically Weighted Regression*. Berikut adalah langkah-langkah analisis yang digunakan dalam melakukan penelitian ini sesuai dengan tujuan penelitian pada Bab 1.

1. Mendeskripsikan Angka Kesakitan Malaria dan faktor yang mempengaruhinya di setiap kabupaten/kota di Provinsi Jawa Timur

2. Memodelkan faktor yang mempengaruhi Angka Kesakitan Malaria di Provinsi Jawa Timur dengan menggunakan metode *Geographically Weighted Regression*.
 - a. Menguji multikolinieritas pada setiap variabel dengan menggunakan nilai VIF.
 - b. Memodelkan menggunakan regresi *Ordinary Least Square*
 - c. Memeriksa heterogenitas spasial dengan menggunakan statistik uji *Breusch-Pagan*
 - d. Melakukan pemilihan pembobot terbaik berdasarkan *bandwith*
 - e. Melakukan pemodelan GWR
 - f. Mendapatkan estimasi parameter untuk model GWR pada masing-masing lokasi
 - g. Melakukan perbandingan model GWR dan OLS.
3. Menginterpretasikan hasil dan memberikan kesimpulan

IV. ANALISIS DAN PEMBAHASAN

Sebelum melakukan pemodelan dengan melakukan metode Regresi Linier dengan pendekatan *Ordinary Least Square* (OLS) dan *Geographically Weighted Regression* (GWR), terlebih dahulu dilakukan analisis secara deskriptif untuk mengetahui karakteristik variabel yang digunakan dalam analisis dapat dilihat di Tabel 1

Tabel 1. Statistika Deskriptif Variabel dalam Analisis

Variabel	Rata-rata	Varians	Koefisien Varians	Minimum	Maksimum
Y	10,06	267,72	162,64	0	66,37
X ₁	0,1006	0,0268	162,64	0	0,6637
X ₂	2796	1955570	178,51	0	0,6220
X ₃	0,0926	0,0205	154,44	0	0,6342
X ₄	0,0768	0,0182	175,63	0	0,6637
X ₅	0,0545	0,0086	170,41	0	0,4144
X ₆	0,0357	0,0107	289,10	0	0,5063
X ₇	0,0360	0,0037	168,13	0	0,2762
X ₈	0,0535	0,0097	183,65	0	0,4484
X ₉	0,0387	0,0054	189,16	0	0,3245
X ₁₀	0,0462	0,0082	196,51	0	0,4050

Tabel 1 menunjukkan nilai rata-rata, koefisien varians, minimum dan maksimum setiap variabel yang digunakan dalam penelitian. Berdasarkan Tabel 4.1, dapat diketahui nilai rata-rata angka kesakitan malaria adalah 10,06. Hal ini berarti pada tahun 2016, setiap 1.000.000 penduduk ada jiwa \approx 11 jiwa.

Untuk mengetahui pengelompokan kabupaten/kota di Jawa Timur berdasarkan angka kesakitan malaria tahun 2016 dapat menggunakan peta tematik. Pada penelitian ini dilakukan klasifikasi dalam peta tematik menjadi 4 kelompok metode *Natural Break*.



Gambar 4.1 Persebaran Angka Kesakitan Malaria di Jawa Timur 2016

Daerah yang memiliki Angka Kesakitan Malaria paling tinggi ditandai dengan warna merah tua dan Angka Kesakitan Malaria paling rendah ditandai dengan warna merah muda

seperti terlihat pada Gambar 4.1. kabupaten/kota yang terletak dibagian timur termasuk didalam daerah Angka Kesakitan Malaria yang berkategori tinggi, sedangkan bagian selatan didominasi dengan Angka Kesakitan Malaria kategori rendah dan sangat rendah. Kabupaten Madiun dan Kabupaten Trenggalek menjadi salah satu kabupaten dengan Angka Kesakitan Malaria kategori sangat tinggi.

Sebelum melakukan pemodelan dengan menggunakan metode regresi linier, dilakukan pengujian asumsi multikolinieritas dengan melihat nilai VIF (*Variance Inflation Factors*).

Tabel 2. Nilai VIF Variabel Prediktor

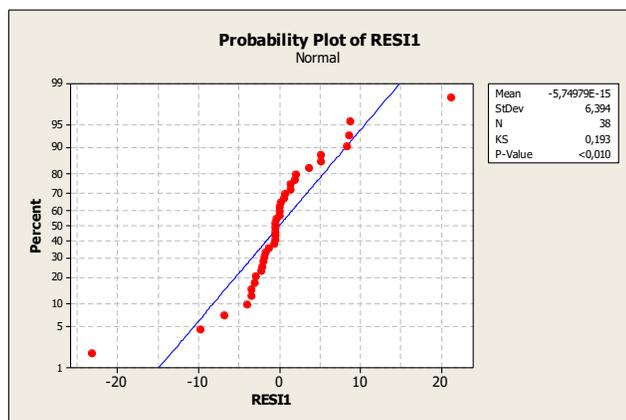
Kode	Variabel	Nilai VIF
x ₂	Rasio Konfirmasi Lab (Mikroskop)	4,202
x ₄	Rasio Penderita diobati dengan ACT + Primaquine	2,208
x ₆	Rasio Hasil pengobatan penderita (<i>follow up</i> tidak lengkap)	5,376
x ₇	Rasio Jenis Parasit Pf	2,717
x ₁₀	Rasio Malaria tanpa komplikasi (rawat jalan)	6,329

Berdasarkan Tabel 2, setelah menghilangkan beberapa variabel yang tidak signifikan dapat terlihat bahwa nilai VIF seluruh variabel kurang dari 10, sehingga penelitian dapat dilanjutkan pada pemodelan dengan menggunakan metode regresi linier.

Salah satu persyaratan dalam pemodelan dengan menggunakan metode *Geographically Weighted Regression* adalah terpenuhinya asumsi residual normal. Hipotesis yang digunakan dalam pengujian asumsi residual berdistribusi normal ini adalah sebagai berikut.

$$H_0 : F_n(y) = F_0(y) \text{ (residual berdistribusi normal)}$$

$$H_1 : F_n(y) \neq F_0(y) \text{ (residual tidak berdistribusi normal)}$$



Berdasarkan Gambar 4.2 dapat terlihat bahwa data tersebut mengikuti garis regresi yang menunjukkan bahwa data tersebut tidak mengikuti asumsi residual normal, namun pada penelitian ini data diasumsikan normal, karena hasilnya akan lebih baik apabila tidak ditransformasikan. Setelah salah satu syarat pemodelan *Geographically Weighted Regression* terpenuhi, selanjutnya dilakukan pemodelan dengan menggunakan Regresi Linier dengan pendekatan *Ordinary Least Square* untuk mengetahui pengaruh variabel prediktor

terhadap variabel respon sebelum dilakukan pemodelan dengan menggunakan metode *Geographically Weighted Regression* (GWR).

Uji parameter secara serentak merupakan uji untuk mengetahui apakah semua variabel prediktor yang dimasukkan ke dalam model memberikan pengaruh signifikan secara bersama-sama terhadap model. Digunakan hipotesis untuk pengujian parameter secara serentak sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{11} = 0$$

$$H_1 : \text{minimal ada } \beta_k \neq 0; k = 1, 2, \dots, 11$$

Tabel 3. Analisis Varians Model Regresi Linier Berganda

Sumber Variasi	Derajat Bebas	Jumlah Kuadrat	Rata-rata Kuadrat	F	P-Value
Regresi	5	9785,0	1957,0	519,63	0,000
Error	32	120,5	3,80		
Total	37	9905,5			

Berdasarkan Tabel 3, diperoleh informasi untuk melakukan pengujian model regresi linear berganda secara serentak. Diperoleh P-value sebesar 0,000 sehingga disimpulkan tolak H_0 yang artinya minimal ada satu variabel prediktor yang memiliki pengaruh signifikan terhadap variabel respon. Selanjutnya dilakukan pengujian signifikansi parameter secara parsial untuk mengetahui variabel mana saja yang berpengaruh secara signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0; k = 1, 2, \dots, 11$$

Tabel 4. Uji Signifikansi Parameter secara Parsial

Variabel	Koefisien	Koefisien SE	P-Value
Rasio Konfirmasi Lab (Mikroskop) (x_2)	54,715	3,894	0,000
Rasio Penderita diobati dengan ACT + Primaquine (x_4)	24,756	2,605	0,000
Rasio Hasil pengobatan penderita (<i>follow up</i> tidak lengkap) (x_6)	25,178	5,118	0,000
Rasio Jenis Parasit Pf (x_7)	51,898	5,955	0,000
Rasio Malaria tanpa komplikasi (rawat jalan) (x_{10})	15,279	6,111	0,043

Setelah dilakukan uji signifikansi secara serentak dan secara parsial, langkah berikutnya adalah mendapatkan model terbaik dalam regresi linier dengan melakukan regresi linier dengan variabel prediktor yang signifikan, sehingga model terbaik untuk regresi linier dapat ditulis sebagai berikut.

$$\hat{Y} = 0,237 + 0,00130x_2 + 33,7x_4 + 35,8x_6 + 54,1x_7 + 18,6x_{10}$$

Pengujian heterogenitas spasial dilakukan untuk mengetahui adanya keberagaman dalam hubungan secara kewilayahan. Heterogenitas spasial dapat diidentifikasi dengan menggunakan pengujian *Breusch Pagan*. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{38}^2 \text{ (tidak terdapat heterogenitas spasial)}$$

$$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma^2 \text{ (ada heterogenitas spasial)}$$

Pada penelitian ini diperoleh p-value sebesar 0,0003589 Dengan menggunakan taraf signifikansi sebesar 0,05, maka

diputuskan tolak H_0 atau terdapat heterogenitas spasial pada data yang diamati.

Setelah aspek heterogenitas spasial terpenuhi, selanjutnya dilakukan pemilihan pembobot optimum yang akan dimasukkan dalam pemodelan GWR.

Tabel 5. Pemilihan Pembobot Optimum

Fungsi Pembobot	CV minimum	Bandwidth
Fixed Gaussian	135,738	0,500
Fixed Bisquare	139,686	1,244

Salah satu kriteria pembobot terbaik adalah dengan meminimumkan nilai CV. Berdasarkan Tabel 4.5, dapat diketahui bahwa nilai CV paling minimum diperoleh dengan menggunakan fungsi pembobot Fixed Gaussian yaitu dengan CV minimum sebesar 135,738 dan *bandwidth* 0,500. Langkah selanjutnya yaitu membentuk matriks pembobot yang digunakan untuk penaksiran parameter di tiap pengamatan. Matriks pembobot spasial yang diperoleh untuk tiap-tiap lokasi kemudian digunakan untuk membentuk model GWR, sehingga Rangkuman hasil estimasi parameter model GWR dapat dilihat pada Tabel 6

Tabel 6. Rangkuman Hasil Estimator Model GWR

Bandwidth	0,500
R ²	0,998
SSE	18,334

Berdasarkan Tabel 4.6 dapat terlihat nilai *bandwidth* optimum yang diperoleh dengan menggunakan fungsi pembobot Fixed Gaussian adalah 0,500 yang artinya titik yang berada di radius 0,500 dianggap berpengaruh secara optimal dalam membentuk parameter model lokasi. Informasi lain yang didapatkan yaitu nilai koefisien determinasi (R^2) model GWR sebesar 0,998 yang berarti model yang terbentuk dapat menjelaskan variabel angka kesakitan sebesar 0,998 dan sisanya dijelaskan oleh variabel lain yang tidak diamati dalam model.

Pemodelan angka kesakitan malaria dengan menggunakan metode GWR diharapkan memperoleh hasil yang lebih baik daripada pemodelan dengan menggunakan regresi OLS dengan hipotesis sebagai berikut.

$H_0 : \beta_k(u_i, v_i) = \beta_k; i = 1, 2, \dots, 38; k = 1, 2, \dots, 4$ (tidak ada perbedaan signifikan antara model regresi linier dengan model GWR)

$H_1 : \text{minimal ada satu } \beta_k(u_i, v_i) \neq \beta_k; i = 1, 2, \dots, 38; k = 1, 2, \dots, 4$ (ada perbedaan yang signifikan antara model regresi linier dengan model GWR)

Diperoleh nilai F_{hitung} sebesar 9,776 atau lebih besar dibanding $F_{Tabel(0,05;3;32)} = 2,38$ sehingga dapat disimpulkan tolak H_0 artinya ada perbedaan signifikan antara model regresi linier dengan model GWR.

Pengujian signifikansi parameter model GWR secara parsial dilakukan untuk mengetahui parameter mana saja yang berpengaruh secara signifikan terhadap angka kesakitan malaria di setiap kabupaten/kota di Jawa Timur. Hipotesis yang

digunakan dalam pengujian model GWR secara parsial adalah sebagai berikut.

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

$H_1 : \beta_k(u_i, v_i) \neq 0; k = 1, 2, \dots, 4$

Berdasarkan pengujian signifikansi parameter model GWR, diperoleh nilai t_{hitung} yang selanjutnya dibandingkan dengan nilai t_{Tabel} untuk mengetahui variabel mana saja yang secara signifikan mempengaruhi angka kesakitan malaria di Jawa Timur. Selanjutnya diperoleh $t_{Tabel(0,05;20;528)} = 2,086$. Diketahui bahwa jika nilai $|t_{hitung}| > t_{tabel}$ maka parameter signifikan pada lokasi pengamatan.

Selanjutnya nilai t_{hitung} yang diperoleh dibandingkan dengan nilai t_{Tabel} sebesar 2,086 untuk mendapatkan variabel signifikan.

Untuk taraf signifikansi sebesar 0,05, diperoleh model GWR Kota Surabaya sebagai berikut.

$$\hat{Y} = 0,094383 + 0,000391x_2 + 48,972x_4 + 60,821x_6 + 55,169x_7 + 24,145x_{10}$$

Model GWR masing-masing kabupaten/kota di Jawa Timur sebagai berikut :

Tabel 7. Variabel Signifikan pada Kabupaten/Kota di Jawa Timur

Kabupaten/Kota	Variabel yang Signifikan
Kab. Pacitan	x_2, x_4, x_6, x_7
Kab. Ponorogo	x_2, x_4, x_6, x_7
Kab. Trenggalek	x_2, x_4, x_6, x_7
Kab. Tulungagung	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Blitar	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Kediri	x_2, x_4, x_6, x_7
Kab. Malang	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Lumajang	x_2, x_4, x_{10}
Kab. Jember	x_2, x_4
Kab. Banyuwangi	x_2, x_4
Kab. Bondowoso	x_2, x_4
Kab. Situbondo	x_2, x_4
Kab. Probolinggo	x_2, x_4, x_6, x_7
Kab. Pasuruan	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Sidoarjo	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Mojokerto	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Jombang	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Nganjuk	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Madiun	x_2, x_4, x_6, x_7
Kab. Magetan	x_2, x_4, x_6, x_7
Kab. Ngawi	x_2, x_4, x_6, x_7
Kab. Bojonegoro	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Tuban	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Lamongan	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Gresik	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Bangkalan	$x_2, x_4, x_6, x_7, x_{10}$
Kab. Sampang	x_4, x_6
Kab. Pamekasan	x_4
Kab. Sumenep	x_2, x_4
Kota Kediri	$x_2, x_4, x_6, x_7, x_{10}$
Kota Blitar	$x_2, x_4, x_6, x_7, x_{10}$
Kota Malang	$x_2, x_4, x_6, x_7, x_{10}$
Kota Probolinggo	x_2, x_4, x_6

Tabel 7. Variabel Signifikan pada Kabupaten/Kota di Jawa Timur (Lanjutan)

Kabupaten/Kota	Variabel yang Signifikan
Kota Pasuruan	$x_2, x_4, x_6, x_7, x_{10}$
Kota Mojokerto	$x_2, x_4, x_6, x_7, x_{10}$
Kota Madiun	x_2, x_4, x_6, x_7
Kota Surabaya	$x_2, x_4, x_6, x_7, x_{10}$
Kota Batu	$x_2, x_4, x_6, x_7, x_{10}$
Kota Kediri	$x_2, x_4, x_6, x_7, x_{10}$

Tabel 8. Variabel Signifikan pada Kabupaten/Kota di Jawa Timur

Variabel yang Signifikan	Kabupaten/Kota
x_4	Kab. Pamekasan
x_2, x_4	Kab. Jember, Kab. Banyuwangi, Kab. Bondowoso, Kab. Situbondo, Kab. Sumenep
x_4, x_6	Kab. Sampang
x_2, x_4, x_6	Kota Probolinggo
x_2, x_4, x_{10}	Kab. Lumajang
x_2, x_4, x_6, x_7	Kab. Pacitan, Kab. Ponorogo, Kab. Trenggalek, Kab. Kediri
	Kab. Probolinggo, Kab. Madiun, Kab. Magetan, Kab. Ngawi, Kota Madiun.
$x_2, x_4, x_6, x_7, x_{10}$	Kab. Tulungagung, Kab. Blitar, Kab. Malang



Gambar 4.13 Pemetaan Variabel Signifikan di Kupatan/Kota di Jawa Timur

Pemilihan model terbaik merupakan proses evaluasi dari model untuk mengetahui seberapa besar peluang masing-masing model yang terbentuk sudah sesuai dengan data. Hasil perbandingan kedua metode tersebut dapat dilihat pada Tabel 9

Tabel 9. Perbandingan Model OLS dan GWR

Kriteria	Regresi Linier	GWR
R^2	98,8%	99,8%
SSE	120,5	17,979

Berdasarkan Tabel 9 diperoleh informasi bahwa berdasarkan nilai R^2 dan nilai SSE, model GWR lebih baik dibandingkan model regresi linier dengan pendekatan OLS. Model GWR terbukti mampu meningkatkan nilai R^2 dan menurunkan nilai *Sum Square Error*.

DAFTAR PUSTAKA

[1] N. Haryati, Analisis Faktor-Faktor Risiko yang Mempengaruhi Jumlah Kasus Malaria di Jawa Timur Tahun 2013 dengan Geographically Weighted Negative Binomial Regression (GWNBR), Surabaya: Jurusan Statistika FMIPA ITS, 2015.

[2] L. Firdial, 2011, Surabaya: Tugas Akhir Jurusan Statistika FMIPA ITS, Pemodelan Angka Harapan Hidup di Jawa Timur dan Jawa Tengah dengan Metode Geographically Weighted Regression .

[3] R. Walpole, Intoduction to Statistics, New York: Macmillan Publishing Co. Inc, 1995.

[4] A. S. Fotheringham, C. Brunsdon and M. E. Charlton, Geographically Weighted Regression: The Analysis of Spatially Varying Relationship, England: John Wiley and Sons LTD, 2002.

[5] J. Lesage and R. K. Pace, Introduction to Spatial Econometrics, New York: CRC Press, 2001.