BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan metode statistika yang digunakan untuk membentuk model hubungan antara peubah terikat atau dependen (Y) dengan satu atau lebih peubah bebas atau independen (X). Analisis regresi mempunyai dua tujuan utama yaitu untuk memprediksi dan untuk menganalisis hubungan kausal (Latan, 2014). Analisis regresi yang digunakan apabila terdapat satu peubah bebas dan satu peubah terikat, maka disebut sebagai regresi linear sederhana. Sedangkan apabila terdapat lebih dari satu peubah bebas dan satu peubah terikat, maka disebut regresi linear berganda (Kurniawan, 2008).

2.2 Analisis Regresi Linear Berganda

Pada umumnya persoalan penelitian yang menggunakan analisis regresi memerlukan lebih dari satu variabel bebas dalam model regresinya. Model yang dapat digunakan untuk persoalan seperti ini adalah model regresi linear berganda. Model regresi linear berganda secara umum dinyatakan sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon \tag{2.1}$$

dimana $\beta_1, \beta_2, \ldots, \beta_p$ adalah parameter (koefisien regresi) yang harus ditaksir berdasarkan data ε adalah galat [ε ~ N (0, σ^2)].

Menggunakan notasi matriks, maka persamaan (2.1) dapat ditulis sebagai berikut (Walpole & Myers, 1995):

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.2)$$

Dengan bentuk sederhana, persamaan (2.10) dapat dinotasikan sebagai berikut:

$$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon} \quad (2.3)$$

dimana:

Y : matriks variabel tak bebas berukuran $n \times 1$.

X : matriks variabel bebas berukuran $n \times (p + 1)$.

$\boldsymbol{\beta}$: matriks koefisien regresi berukuran $(p + 1) \times 1$.

$\boldsymbol{\varepsilon}$: matriks galat berukuran $n \times 1$.

2.2.1 Estimasi Model Regresi Linear Berganda

Koefisien-koefisien pada model regresi merupakan nilai estimasi atau taksiran parameter didalam model regresi untuk kondisi yang sebenarnya. Metode yang sering digunakan dalam menaksir koefisien parameter model regresi adalah Metode *Ordinary Least Square* (OLS). Metode ini meminimumkan jumlah kuadrat galat. Secara sederhana penaksiran parameter model regresi menggunakan OLS dapat dituliskan dalam bentuk matriks sebagai berikut:

$$(\mathbf{X}^{\mathrm{T}}\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathrm{T}}\mathbf{Y} \quad (2.4)$$

jika matriks $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ tak singular, maka jawaban untuk koefisien regresi dapat ditulis sebagai berikut (Walpole & Myers, 1995) :

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \tag{2.5}$$

dengan $\hat{\beta}$ merupakan vektor estimasi dari koefisien regresi $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$, dari persamaan (2.4) dapat dinyatakan dengan matriks (Kurniawan & Yuniarto, 2016):

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} \quad \mathbf{X}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} n & \sum x_1 & \cdots & \sum x_p \\ \sum x_1 & \sum x_1^2 & \cdots & \sum x_1 x_p \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_p & \sum x_1 x_p & \cdots & \sum x_p^2 \end{bmatrix} \quad \mathbf{X}^{\mathrm{T}}\mathbf{Y} = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \vdots \\ \sum x_p y \end{bmatrix} \tag{2.6}$$

2.2.2 Pengujian Parameter Model

Pengujian parameter dalam model regresi bertujuan untuk mengetahui apakah parameter tersebut telah menunjukkan hubungan yang nyata antara peubah bebas dengan peubah terikat. Disamping itu pengujian tersebut juga digunakan untuk mengetahui kelayakan parameter dalam menerangkan model. Menurut Agresti (1996), pengujian terhadap keberartian model dapat dilakukan dengan dua tahap pengujian yaitu pengujian secara simultan dengan uji F dan pengujian secara parsial dengan uji t.

2.2.2.1 Pengujian Simultan

Pengujian parameter secara simultan bertujuan untuk melihat apakah ada pengaruh yang nyata antara peubah bebas terhadap peubah terikat secara serentak atau menyeluruh dalam model dengan menggunakan analisis varians (Anova), yang disajikan pada tabel 2.1 sebagai berikut:

Tabel 2.1 Tabel Anova (*Analisis Of Variance*)

Sumber Variansi	DF	SS	MS	F_{hitung}
Regresi	p	$SSR = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$	$MSR = \frac{SSR}{p}$	
Error	$n-p-1$	$SSE = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$	$MSE = \frac{SSE}{n-p-1}$	$F_{Hitung} = \frac{MSR}{MSE}$
Total	$n-1$	$SST = \sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2$		

Hipotesis untuk uji signifikansi model secara simultan:

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_j = 0$ (Tidak ada pengaruh peubah bebas terhadap peubah terikat)

H_1 : paling sedikit ada satu $\beta_j \neq 0$ dengan $j = 1, 2, \ldots, k$ (minimal ada satu peubah bebas yang berpengaruh terhadap peubah terikat)

Kriteria Uji:

Tolak H_0 jika $F_{hitung} > F_{tabel}$ atau nilai $P \leq \alpha$. Dalam hal ini dapat dikatakan bahwa minimal ada satu peubah bebas yang mempengaruhi peubah terikat (Sembiring, 2003).

2.2.2.2 Pengujian Parsial

Pengujian parsial digunakan untuk menguji signifikansi parameter model secara masing-masing antara variabel bebas terhadap variabel terikat.

Hipotesis untuk uji signifikansi model secara parsial:

H_0 : $\beta_i = 0$ (Tidak ada pengaruh variabel bebas ke-i terhadap variabel terikat untuk $i = 1, 2, \ldots, p$)

H_1 : $\beta_i \neq 0$ (Ada pengaruh variabel bebas ke-i terhadap variabel terikat)

Statistik uji *t* dinyatakan dalam rumus sebagai berikut:

$$t_{Hitung} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \tag{2.7}$$

dengan $SE(\hat{\beta}_i) = \sqrt{\frac{MSE}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$

dimana

x_i : nilai variabel bebas pada pengamatan ke-*i*

$\bar{x}$: nilai rata-rata variabel bebas

Kriteria uji:

Tolak H_0 jika $|t_{hitung}| > t_{tabel}$ atau nilai $P \leq \alpha$.

2.3 Data Spasial

Menurut Fotheringham A S *et al* (2000) data spasial terdiri atas observasi beberapa fenomena yang memiliki beberapa kecenderungan spasial. Data spasial disebut juga data dependen, karena berasal dari lokasi spasial yang berbeda dimana hal ini mengindikasikan ketergantungan antara pengukuran dan lokasi. Posisi lokasi dari suatu pengamatan memungkinkan adanya hubungan dengan pengamatan lain yang berdekatan. Hubungan antar pengamatan tersebut dapat berupa kedekatan jarak antar pengamatan maupun antar persingggungan pengamatan. Adanya efek spasial merupakan hal yang sering terjadi antara suatu wilayah dengan wilayah lainnya.

2.4 Pengujian Efek Spasial

Efek spasial adalah ketergantungan yang terjadi akibat adanya korelasi antar wilayah. Adanya efek spasial merupakan hal yang sering terjadi antara suatu wilayah dengan wilayah lainnya. Menurut Anselin (1988) efek spasial yang dihasilkan oleh informasi antar lokasi dapat dibagi menjadi dua jenis, yaitu dependensi spasial dan heterogenitas spasial.

2.4.1 Dependensi spasial

Dependensia spasial atau disebut juga aoutokorelasi spasial merupakan korelasi antara variabel independen dengan dirinya sendiri berdasarkan ruang, atau dapat diartikan sebagai suatu ukuran kemiripan dari objek di dalam suatu ruang (jarak, waktu, dan wilayah) (Karim, 2012). Pengujian autokorelasi spasial dapat dilakukan dengan metode *Moran's I*. Uji *Moran's I* merupakan sebuah uji statistik lokal yang digunakan untuk mengidentifikasi adanya autokorelasi spasial pada suatu lokasi. Perhitungan autokorelasi spasial menggunakan Indeks Moran dengan matriks pembobot W berdasarkan perkalian silang adalah sebagai berikut:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.8}$$

Hipotesis yang digunakan dalam menguji dependensi spasial adalah:

H_0 : I=0 (tidak ada dependensi spasial antar lokasi)

H_1 : I≠0 (ada dependensi spasial antar lokasi)

Statistik Uji:

$$Z_{hitung} = \frac{I - \mathrm{I}_o}{\sqrt{\mathrm{var(I)}}} \sim N(0,1) \qquad (2.9)$$

Dengan nilai harapan:

$$\mathrm{E}(I) = \mathrm{I}_o = -\frac{1}{\mathrm{n}-1} \qquad (2.10)$$

Ragam untuk pendekatan normal:

$$\mathrm{var}(I) = \frac{\mathrm{n}^2 S_1 - \mathrm{n}S_2 + 3S_o^2}{(\mathrm{n}^2-1)S_o^2}; \qquad (2.11)$$

$$S_1 = \frac{1}{2}\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{n}}(W_{\mathrm{ij}} + W_{\mathrm{ji}})^2; \qquad (2.12)$$

$$S_2 = \sum_{\mathrm{i}=1}^{\mathrm{n}}\left(\sum_{\mathrm{j}=1}^{\mathrm{n}} W_{\mathrm{ij}} + \sum_{\mathrm{j}=1}^{\mathrm{n}} W_{\mathrm{ji}}\right)^2; \qquad (2.13)$$

$$S_0 = \sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{n}} W_{\mathrm{ij}}; \qquad (2.14)$$

Keterangan:

x_i = data variable lokasi ke-*i* (*i* = 1,2,…,*n*)

x_j = data variable lokasi ke-*j* (*j* = 1,2,…,*n*)

$\bar{x}$ = rata-rata data

W = matriks pembobot Queen

n = data jumlah lokasi penelitian

$Var(I)$ = Variansi Moran'I

$E(I)$ = Nilai harapan Moran'I

Kriteria Uji:

Pengambilan keputusan H_0 ditolak yaitu jika $|Z_i| > Z_{\alpha/2}$ atau P-*value* < α.

Nilai Indeks Moran adalah -1 hingga 1. Nilai Indeks Moran (I > 0), menandakan adanya autokorelasi spasial yang positif, sedangkan nilai Indeks Moran (I < 0) menendakan bahwa adanya autokorelasi spasial yang negatif. Adapun jika nilai Indeks Moran (I = 0) menandakan tidak adanya autokorelasi spasial pada data.

2.4.2 Heterogenitas Spasial

Heterogenitas spasial disebabkan karena adanya perbedaan karakteristik antar titik lokasi pengamatan. Menurut Anselin (1988), heterogenitas spasial tercermin dari galat dalam pengukuran yang mengakibatkan heteroskedastisitas artinya variansi galat yang dihasilkan tidak konstan. Untuk mendeteksi ada atau tidaknya heterogenitas spasial dalam model dilakukan uji *Breusch-Pagan*.

Hipotesis:

$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$ (tidak terjadi heterogenitas spasial antar wilayah)

H_1 : Minimal ada satu $\sigma_i^2 \neq \sigma^2$ (terjadi heterogenitas spasial antar wilayah)

Statistik Uji:

$$BP = \frac{1}{2} b^T Z \left(Z^T Z\right)^{-1} Z^T b \tag{2.15}$$

dimana, elemen vektor b dirumuskan $b = \frac{e_i^2}{\sigma^2} - 1$ dengan e_i merupakan galat atau error untuk pengamatan ke-i dengan asumsi $e \sim IIDN(0,\sigma 2)$, σ^2 adalah ragam dari e_i, dan Matriks $\boldsymbol{Z}$ adalah matriks yang berukuran $n \times (p+1)$ dan telah distandarisasi untuk setiap pengamatan.

Kriteria Uji:

Tolak H_0 jika $BP > \chi_p^2$ atau P-*value* < α. Dalam hal ini dapat dikatakan bahwa terjadi heterogenitas spasial antar wilayah.

2.5 Model GWR

Geographically Weighted Regression (GWR) merupakan pengembangan dari regresi linear, dimana setiap parameter dihitung pada setiap lokasi pengamatan sehingga penduga parameter yang dihasilkan sesuai dengan jumlah lokasi yang digunakan atau dengan kata lain penduga parameter memiliki nilai yang berbeda-beda dengan lokasi lainnya (Fotheringham *et al*., 2002). Adapun Model untuk *Geographically Wighted Regression* (GWR) sebagai berikut:

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) X_{ik} + \varepsilon_i \qquad (2.16)$$

dimana

Y_i : Nilai variabel respon pada titik lokasi pengamatan ke-*i*

$\beta_0(u_i, v_i)$: Konstanta/*intercept* GWR

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

u_i, v_i : Titik koordinat lintang dan bujur pada lokasi pengamatan ke-*i*

X_{ik} : Nilai variabel prediktor ke-*k* pada titik lokasi pengamatan ke-*i*

ε_i : Error pada titik lokasi ke-*i*

2.5.1 Pembobot GWR

Peran pembobot pada model GWR sangat penting karena nilai pembobot ini mewakili letak data observasi satu dengan lainnya. Pembobot berupa matriks diagonal dimana elemen-elemen diagonalnya merupakan sebuah fungsi pembobot dari setiap titik lokasi pengamatan. Matriks pembobot pada GWR merupakan matriks pembobot yang berbasis pada kedekatan titik lokasi pengamatan ke-*i* dengan titik lokasi pengamatan lainnya. Pengamatan terdekat ke titik lokasi pengamatan ke-*i* umumnya diasumsikan memiliki pengaruh paling besar terhadap penaksiran parameter di titik lokasi pengamatan ke-*i*. Oleh karena itu, matriks pembobot $W(u_i, v_i)$ akan semakin besar seperti jarak yang semakin dekat.

Menurut Chasco *et al*. (2007), pembobotan sendiri dapat dilakukan dengan metode yang berbeda-beda, diantaranya dengan menggunakan fungsi kernel (*kernel function*). Fungsi kernel digunakan untuk mengestimasi paramater dalam model GWR jika fungsi jarak (w_j) adalah fungsi yang kontinu dan monoton turun. Menurut Wheeler dan Antonio (2010), ada dua jenis fungsi kernel dalam GWR yaitu fungsi kernel tetap (*fixed kernel*) dan fungsi kernel adaptif (*Adaptive Kernel*). Fungsi *fixed kernel* adalah fungsi kernel yang memiliki *bandwidth* yang sama untuk setiap lokasi pengamatan. Sedangkan fungsi *adaptive kernel* adalah fungsi kernel yang memiliki *bandwidth* yang berbeda untuk setiap lokasi. Hal ini disebabkan oleh fungsi *kernel adaptive* yang dapat menyesuaikan dengan titik-titik pengamatan (Pamungkas *et al*, 2016).

Perbedaan paling mendasar antara fungsi *Fixed Kernel* dan *Adaptive Kernel* adalah dalam penentuan *bandwidth* optimumnya. Fungsi *fixed kernel*, *bandwidth* optimumnya yang berupa jarak akan sama dimanapun lokasinya berada. Sedangkan *adaptive kernel* akan menggunakan jarak tertangga terdekat (*Nearest Neighborhood*) untuk menentukan berapa titik yang memiliki karakteristik yang sama dengan bidang yang akan dicari modelnya (Fotheringham,dkk., 2002). Terdapat kelemahan yang berpotensi menjadi masalah potensial yang terjadi pada penggunaan fungsi *Fixed Kernel*, dimana untuk beberapa lokasi pada area penelitian yang hanya terdiri dari beberapa titik data yang tersedia untuk kalibrasi model atau titik data yang berjauhan disekitar pusat lokasinya akan menjadi masalah "weak data". Sebagai alternatif, fungsi *adaptive kernel* dapat digunakan untuk mengurangi kelemahan data tersebut. Dimana fungsi *adaptive Kernel* akan menyesuaikan dengan sendirinya ukuran variansi sesuai dengan kerapatan datanya. Nilai *bandwidth* yang lebar jika titik datanya jarang, dan akan menghasikan *bandwidth* yang kecil jika titik datanya lebih padat (Fotheringham, dkk., 2002).

Menurut Fotheringham (2020) terdapat tiga jenis fungsi *kernal adaptive* yang dijadikan sebagai pembobot spasial dalam analisis dengan GWR yaitu *adaptive gaussian kernel, adaptive bisquare kernel* dan *adaptive tricube kernel*.

1. *Adaptive Gaussian Kernel*

   Matriks pembobot fungsi *adaptive gaussian kernel* dinyatakan dengan formula sebagai berikut.

$$\textit{Adaptive Gaussian}: W_{ij} = \exp\left[-\frac{1}{2}\left({d_{ij}}/{h_i}\right)^2\right] \quad (2.17)$$

2. *Adaptive Bisquare Kernel*

Dalam menghitung fungsi pembobot *adaptive bisquare kernel* dapat dilakukan dengan perhitungan berikut (Chasco, *et al.*, 2007)

$$\text{Adaptive Bisquare} : W_{ij} = \begin{cases} \left(1 - \left(\dfrac{d_{ij}}{h_i}\right)^2\right)^2 & , untuk\ d_{ij} \le h \\ 0 & , untuk\ d_{ij} > h \end{cases} \tag{2.18}$$

3. *Adaptive Tricube Kernel*

Dalam menghitung fungsi pembobot kernel *tricube* dapat dilakukan dengan perhitungan berikut:

$$\text{Adaptive Bisquare} : W_{ij} = \begin{cases} \left(1 - \left(\dfrac{d_{ij}}{h_i}\right)^3\right)^3 & , untuk\ d_{ij} \le h \\ 0 & , untuk\ d_{ij} > h \end{cases} \tag{2.19}$$

dimana konstan *b* adalah parameter penghalus (*bandwidth*) yang mengontrol seberapa jauh radius yang masih mempengaruhi lokasi ke-*i*. d_{ij} adalah fungsi jarak *euclidean* diantara lokasi ke-*i* dank ke-*j* dimana $i, j = 1, 2, \ldots, n$ yang didefenisikan dengan:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \tag{2.20}$$

Jika pembobot yang digunakan adalah fungsi kernel maka pemilihan *bandwidth* ini sangatlah penting. Menurut Bravendi (2018), ketika *bandwidth* terlalu besar, pembobot akan menjadi sangat kecil. Ketika *bandwidth* kecil maka, pembobot akan menjadi sangat besar. Oleh karena itu pemilihan *bandwidth* optimum menjadi penting karna akan mempengaruhi ketetapan model terhadap data. Salah satu metode yang digunakan untuk menentukan *bandwidth* optimum adalah dengan menggunakan *Cross Validation* (CV) yang secara sistematis didefinisikan sebagai berikut (Fotheringham, et al., 2002):

$$CV = \sum \left[y_i - \hat{y}_{\neq i}(h)\right]^2 \qquad (2.21)$$

dengan $\hat{y}_{\neq i}(b)$ adalah nilai estimasi $\hat{y}_i$ dimana pengamatan dilokasi i dihilangkan dari proses penaksiran dan n adalah banyaknya sampel. *Bandwidth* yang optimal ditunjukkan dengan nilai CV minimum.

2.5.2 Estimasi Model GWR

Estimasi parameter model GWR dilakukan dengan metode *Weighted Least Square* (WLS) yaitu dengan memberikan penimbang atau pembobot yang berbeda pada setiap lokasi pengamatan. Pemberian bobot ini sesuai dengan Hukum I Tobbler "*Everything is related to everything else, but near thin are more related than distant things*/ segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang dekat lebih mempunyai pengaruh dari pada sesuatu yang jauh" (Miller dalam Yasin, 2011). Menurut Carlton (2009) Unit observasi yang memiliki kedekatan lokasi, akan memiliki bobot lebih besar dibandingakn unit observasi yang lokasinya lebih jauh. Penaksir parameter dengan GWR untuk setiap variabel ke-k pada lokasi pengamatan, dinyatakan sebagai berikut:

$$\hat{\beta}(u_i, v_i) = \left(X^T W(u_i, v_i) X\right)^{-1} X^T W(u_i, v_i) Y \qquad (2.22)$$

dengan $W(u_i, v_i)$ adalah matriks pembobot spasial untuk lokasi i, Y adalah vector kolom data variabel respon (Y) dan X adalah matriks rancangan data variabel bebas atau variabel penjelas. Matrik rancangan dapat dinyatakan sebagai

$$X=\begin{bmatrix}1 & X_{1;1} & \cdots & X_{p;1}\\ 1 & X_{1;2} & \cdots & X_{p;2}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & X_{1;m} & \cdots & X_{p;m}\end{bmatrix}=\begin{bmatrix}X_1^T\\ X_2^T\\ \vdots\\ X_{m1}^T\end{bmatrix}$$

, X (Fortheringham *et al*., 2002).

2.5.3 Pengujian Parameter Model GWR

Pengujian model GWR terdiri dari dua macam, yaitu uji kesesuaian antara model model GWR (*Goodness of Fit*) dan uji parsial model GWR.

1. Uji Kesesuaian Model (*Goodness of Fit*)

Pengujian kesesuaian model bertujuan untuk menjelaskan apakah model GWR dapat menjelaskan lebih baik dibandingkan model regresi linier atau tidak. Adapun hipotesis yang digunakan dalam pengujian kesesuaian model GWR adalah sebagai berikut:

H_0 : $\beta_k(u_i,v_i)=\beta_k$ dengan $k=1,2,\ldots,p$ dan $i=1,2,\ldots,p$ (Tidak ada perbedaan yang signifikansi antara model regresi OLS dan GWR)

H_1 : paling sedikit ada satu $\beta_k(u_i,v_i)\neq k$ dengan $k=1,2,\ldots,p$ dan $i=1,2,\ldots,p$ (Ada perbedaan yang signifikansi antara model regresi OLS dan GWR)

Statistik Uji:

$$F^*=\frac{SSE(H_1)/df_1}{SSE(H_2)/df_2} \qquad (2.23)$$

dengan

$$SSS(H_1)=y^T(I-H)\,\text{dimana } H=X(X^TX)^{-1}X^T \qquad (2.24)$$

$$SS(H_0)=y^T(I-L)^T(I-L)y \qquad (2.25)$$

$$df_1 = \frac{\partial_1^2}{\partial_2} \text{ dimana } \partial_i = tr\left(\left[(I-L)^T (I-L)\right]^i\right), i = 1,2 \tag{2.26}$$

$$df_1 = n - p - 1 \tag{2.27}$$

dengan I merupakan matriks identitas berukuran $n \times n$ serta L adalah matriks proyeksi dari model GWR, berikut adalah matriks proyeksinya:

$$L = \begin{bmatrix} X_1^T \left(X^T W(u_1,v_1) X\right)^{-1} X^T W(u_1,v_1) \\ X_2^T \left(X^T W(u_2,v_2) X\right)^{-1} X^T W(u_2,v_2) \\ \vdots \\ X_n^T \left(X^T W(u_n,v_n) X\right)^{-1} X^T W(u_n,v_n) \end{bmatrix}$$

Kriteria uji:

Tolak H_0 jika $F^* \geq F_{\alpha;df_1;df_2}$ dan sebaliknya.

2. Uji Parsial Model GWR

Pengujian parameter model GWR dilakukan dengan menguji parameter secara parsial. Pengujian ini bertujuan untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Bentuk hipotesisisnya sebagai berikut:

H_0 : $\beta_k(u_i,v_i) = 0$ dengan $k, i = 1,2,\ldots,p$ (Tidak ada pengaruh antara variabel bebas terhadap variabel tak bebas)

H_1 : paling sedikit ada satu $\beta_k(u_i,v_i) \neq 0$ dengan $k, i = 1,2,\ldots,p$ (Ada pengaruh antara variabel bebas terhadap variabel tak bebas)

Statistik Uji:

$$\hat{\beta}_k = \frac{(u_i,v_i)}{\hat{\sigma}\sqrt{c_{kk}}} \tag{2.28}$$

dengan c_{kk} adalah elemen diagonal ke-*k* dari matriks $C_i C_i^T$, dimana nilai C_i dinyatakan dalam persamaan (2.26).

$$C_i = \left(X^T W(u_1, v_1) X\right)^{-1} X^T W(u_1, v_1) \tag{2.29}$$

Kriteria uji:

Tolak H_0 jika $\left|t_{hitung}\right| \geq t_{\alpha/2;df}$. Dalam hal ini dapat dikatakan bahwa ada pengaruh antara variabel bebas terhadap variabel tak bebas.

2.6 Pemilihan Model Terbaik

2.6.1 Koefisien Determinansi (R^2)

Koefisien determinasi (R^2) merupakan Sumpangan pengaruh yandiberikan variabel bebas terhadap variabel terikat. Menurut Gujarati (1993), besaran koefisien determinasi (R^2) merupakan besaran yang paling lazim digunakan untuk mengukur kecocokan suatu model (*goodness of fit*) garis regresi. Sedangkan menurut Ghozali (2012: 97) koefisien determinasi merupakan alat untuk mengukur seberapa jauh kemampuan model dalam menenrangkan variansi variabel dependen. Nilai R^2 yang kecil atau mendekati nol berarti kemampuan variabel bebas dalam menjelaskan variabel tak bebas sangat terbatas, sedangkan nilai R^2 mendekati satu berarti kemampuan dari variabel bebas dalam menjelaskan variabel tak bebas sangat kuat, sehingga mengidentifikasikan bahwa model mampu menjelaskan variabilitas suatu data (Putri, 2013). berikut rumus dari koefisien determinasi adalah sebagai berikut:

$$R^2 = 1 - \frac{SSE}{SST} \tag{2.30}$$

2.6.2 Akaike Iformation Criterion (AIC)

AIC dalam *Acquah* (2013) adalah suatu ukuran informasi yang berisi pengukuran terbaik dalam uji kelayakan estimasi model. AIC digunakan untuk memilih model terbaik diantara model-model yang diperoleh. Pemilihan model didasarkan pada kesalahan hasil ekspektasi yang terkecil yang membentuk data observasi baru (*error*) yang berdistribusi sama dari data yang digunakan, lebih lanjut AIC mampu mengukur kococokan model dari estimasi menggunakan estimasi *maximum likelihood* dari data yang sama, didefinisikan:

$$AIC = -2\log(L) + 2p \tag{2.31}$$

dimana p adalah jumlah parameter model dan L adalah nilai *maximum likelihood* hasil estimasi model. Evaluasi dilakukan dengan cara membandingkan nilai AIC model yang diperoleh, model dengan nilai AIC paling kecil adalah model yang terbaik.

2.7 Demam Berdarah *Dengue* (DBD)

Demam berdarah *dengue* atau DBD adalah penyakit febril akut yang ditemukan di daerah tropis, dengan penyebaran geografis yang mirip dengan malaria. Penyakit ini disebabkan oleh salah satu dari empat *serotipe* virus dari genus *Flavivirus, family Flaviviridae* (Gama dan Betty, 2010). Penyakit DBD sering muncul sebagai kejadian luar biasa (KLB) dengan angka kematian relatif tinggi (Utami, 2013).

DBD adalah penyakit yang ditandai dengan beberapa gejala klinis seperti: demam tinggi mendadak tanpa sebab yang jelas dan berlangsung terus menerus selama 2-7 hari, terjadi manifestasi perdarahan (petekie, purpura, pendarahan konjungtiva, epistkasis, ekimosis, melena dan hematuria), uji Tourniqet positif, Trombositopeni (100.000/ µ atau kurang), terjadi peningkatan hematokrit 20% atau lebih, bila status lanjut dapat disertai pembesaran hati (Kemenkes RI, 2011).

Pada modul pengendalian DBD apabila keluarga/masyarakat menemukan gejala DBD, maka pertolongan pertama oleh keluarga adalah sebagai berikut:

1. Tirah baring selama demam
2. Memberi parasetamol
3. Kompres hangat
4. Memperbanyak minum, kecuali susu coklat dan sirup merah

2.8 Faktor-faktor Penyebab Demam Berdarah

Penularan penyakit DBD memiliki tiga faktor yang memegang peranan pada penularan infeksi virus, yaitu manusia, virus dan vektor perantara (Hadinegoro *et al*, 2001). Banyak faktor yang menjadi penyebab demam berdarah, beberapa contohnya adalah sebagai berikut:

1. Kepadatan Penduduk

Kepadatan penduduk adalah banyaknya penduduk per satuan luas. Kepadatan penduduk kasar atau *Crude Population Density* (CPD) menunjukkan jumlah penduduk untuk setiap kilometer persegi luas wilayah. Luas wilayah yang dimaksud adalah luas seluruh daratan pada suatu wilayah administrasi. Kepadatan penduduk merupakan indikator dari tekanan penduduk di suatu daerah (BPS, 2016).

Menurut WHO (2009) kepadatan penduduk yang tinggi di Indonesia merupakan salah satu faktor risiko penularan penyakit DBD. Semakin padat penduduk, nyamuk *Aedes aegypti* semakin mudah menularkan virus dengue dari satu orang ke orang lainnya. Pertumbuhan penduduk yang tidak memiliki pola tertentu dan urbanisasi yang tidak terkontrol menjadi faktor yang juga berperan dalam munculnya kejadian luar biasa penyakit DBD.

Salah satu faktor meningkatnya kasus DBD adalah kepadatan penduduk. Dalam penelitian Sari (2005) kepadatan penduduk yang tinggi akan mempermudah terjadinya infeksi virus *dengue*. Daerah yang berpenduduk padat akan meningkatkan jumlah kejadian DBD, hal ini disebabkan oleh kemampuan jarak terbang nyamuk betina kurang dari 100 meter sehingga memungkinkan terjadinya penularan.

2. Suhu atau Iklim

Iklim merupakan kondisi rata-rata cuaca pada suatu wilayah dalam periode waktu yang sangat lama. Menurut Sutamihardja (2004) sistem iklim sangat kompleks dan interaktif, terdiri dari atmosfir, permukaan tanah, salju dan es, lautan dan badan air lainnya, serta makhluk hidup. Iklim sering didefinisikan sebagai cuaca rata-rata dan biasanya dijelaskan dengan suhu rata-rata, variabilitas suhu presipitasi dan angina selama suatu periode waktu, yang berkisar dari bulan ke jutaan tahun (periode yang biasa digunakan adalah 30 tahun).

Energi pada system iklim diperoleh dari radiasi matahari (Sutamihardja, 2009). Secara langsung maupun tidak langsung, angin dan awan di permukaan bumi terkait dengan matahari. Panas dari matahari menghasilkan perubahan suhu bumi yang mengarah pada perbedaan suhu dan tekanan akibat siklus siang dan malam. Perbedaan suhu ini juga menyebabkan pergerakan angin yang selalu bergerak dari tekanan tinggi ke tekanan rendah (Numberi, 2009). Iklim dan cuaca merupakan dua hal yang sangat berhubungan. Perubahan-perubahan statistik pada cuaca dalam beberapa waktu menunjukkan terjadinya perubahan iklim (Kementerian Negara Lingkungan Hidup, 2007).

Faktor-faktor iklim dan lingkungan yang paling banyak berperan dalam penyakit DBD adalah suhu udara, kelembapan nisbi dan ketersediaan air. Menurut Agoes (2005) apabila pemanasan bumi secara bertahap meningkat, maka pengaruhnya adalah percepatan pertumbuhan nyamuk. Siklus perkawinan dan pertumbuhan nyamuk dari telur menjadi larva dan nyamuk dewasa, yang sangat dipengaruhi oleh faktor suhu dan kelembapan, akan dipersingkat sehingga populasi nyamuk semakin meningkat.

3. Rumah Layak Huni

Menurut BPS rumah layak huni adalah rumah yang memiliki akses terhadap sumber air minum layak dan sanitasi yang layak. Kemudian kondisi atap, lantai dan dinding juga memenuhi standar layak serta memiliki pencahayaan yang cukup. Berbagai penelitian penyakit menular membutikan bahwa kondisi perumahan yang berdesak-desakan dan kumuh mempunyai kemungkinan lebih besar terserang penyakit (Mukono 2006). Menurut Kinansi dan Martiningsih (2015) kondisi rumah yang layak mempunyai kemungkinan sangat kecil untuk terjadinya DBD.

4. Jumlah Puskesmas

Pusat kesehatan masyarakat adalah fasilitas kesehatan masyarakat yang menyelenggarakan upaya kesehatan masyarakat dan upaya kesehatan perseorangan tingkat pertama dengan lebih mengutamakan upaya promotif dan preventif, untuk mencapai derajat kesehatan masyarakat yang setinggi-tingginya diwilayah kerjanya (Permenkes, No.75 Tahun 2014). Menurut WHO (2009) adanya akeses yang baik untuk mecapai pelayanan kesehatan dan penanganan yang tepat, baik sejak awal maupun perawatan lanjutan serta peningkatan pengetahuan tetntang DBD dapat menurunkan tingkat kematiannya hingga di bawah 1%.