

## ROBUST SPATIAL AUTOREGRESSIVE UNTUK PEMODELAN ANGKA HARAPAN HIDUP PROVINSI JAWA TIMUR

Hidayatul Musyarofah<sup>1</sup>, Hasbi Yasin<sup>2</sup>, Tarno<sup>3</sup>

<sup>1,2,3</sup>Departemen Statistika FSM Universitas Diponegoro

Email: [hasbiyasin@live.undip.ac.id](mailto:hasbiyasin@live.undip.ac.id)

### ABSTRACT

Spatial regression analysis is regression method used for type of data has a spatial effect. Spatial regression showing the presence of spatial effects on the response variable (Y) is a Spatial Autoregressive (SAR). Outlier often found in research spatial data. The outlier is called the spatial outliers. The analysis can be used to handle outliers in general is Robust Regression. There are several estimator that can be used in which the estimator Robust Regression S, M, MM and LTS. Meanwhile, Robust Regression were used to handle spatial outlier is a combination of SAR and Regression Robust method to form a new method that is Robust Spatial Autoregressive (Robust SAR). Type estimator used in this study is the S-Estimator. This study was conducted to determine the best model on a case study Life Expectancy of East Java Province. The best model is analyzed by comparing the methods of SAR and SAR Robust method. Based on the analysis results obtained MSE and Adjusted R<sup>2</sup> values for the SAR method are 1.7521 and 55.54% while for the Robust SAR method are 0.7456 and 62.30%. The Robust SAR model has a lower MSE value and a higher Adjusted R<sup>2</sup> when compared to the SAR model. Thus the best model for modeling the life expectancy in East Java is Robust SAR models.

**Keywords:** *Spatial Autoregressive (SAR), Robust SAR, Life expectancy*

### 1. PENDAHULUAN

Angka Harapan Hidup adalah suatu indikator penting untuk mengukur tingkat kesehatan suatu negara. Angka Harapan Hidup saat lahir didefinisikan sebagai rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir<sup>[3]</sup>. Angka Harapan Hidup dapat naik atau turun karena berbagai faktor yang mempengaruhi. Metode statistika yang dapat digunakan untuk mengetahui hubungan antar variabel yang saling berhubungan adalah Analisis Regresi. Pada saat observasi seringkali ditemukan data-data yang mengandung efek lokasi (*spatial effect*). Metode yang dapat digunakan untuk mengatasi data-data spasial adalah Analisis Regresi Spasial. Pemodelan *Spatial Autoregressive* (SAR) adalah salah satu model regresi spasial yang dapat digunakan untuk data yang menunjukkan adanya efek spasial pada variabel dependen<sup>[1]</sup>. Penculan juga sering ditemukan dalam penelitian data spasial. Regresi *Robust* adalah alat penting yang dapat digunakan untuk menganalisis data yang terkontaminasi dengan penculan. Ada beberapa jenis estimator dalam Regresi *Robust* diantaranya yaitu: *M-estimator*, *S-estimator*, *MM-estimator* dan *LTS-estimator*. *S-estimator* adalah metode *Robust* yang memiliki nilai *breakdown point* paling tinggi<sup>[4]</sup>. Sedangkan regresi *robust* yang digunakan untuk mengatasi penculan yang muncul pada data spasial merupakan kombinasi dari metode *robust* dan regresi spasial.

Pemodelan regresi spasial dan *robust* dapat dilakukan dengan bantuan software. Software yang akan digunakan pada penelitian ini ada Matlab R2013b. Salah satu fasilitas yang terdapat pada matlab adalah GUI (*Graphical User Interface*). GUI merupakan tampilan grafis yang memudahkan user berinteraksi dengan perintah teks. Dengan GUI, program yang dibuat menjadi lebih user *friendly*, sehingga user mudah menjalankan suatu

aplikasi program<sup>[2]</sup>. Pada tulisan ini peneliti tertarik untuk mengambil judul “*Robust Spatial Autoregressive Berbasis GUI Matlab untuk Pemodelan Angka Harapan Hidup Provinsi Jawa Timur*”. Studi kasus yang digunakan adalah Angka Harapan Hidup setiap Kota/ Kabupaten Provinsi Jawa Timur pada Tahun 2017. Estimator *Robust* yang digunakan adalah estimator S.

## 2. TINJAUAN PUSTAKA

### 2.1 Regresi Spasial

Model umum regresi spasial dapat dituliskan sebagai berikut<sup>[10]</sup> :

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (1)$$

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_n) \quad (2)$$

Keterangan :

- $\mathbf{y}$  = Vektor variabel dependen berukuran  $n \times 1$ ,
- $\rho$  = Koefisien parameter spasial lag dari variabel dependen,
- $\mathbf{W}$  = Matriks pembobot spasial yang berukuran  $n \times n$ ,
- $\mathbf{X}$  = Matriks variabel independen berukuran  $n \times p$ ,  $p = k+1$
- $\boldsymbol{\beta}$  = Vektor koefisien parameter regresi berukuran  $p \times 1$ ,  $p = k+1$
- $\lambda$  = Koefisien parameter spasial error,
- $\mathbf{u}$  = Vektor error yang mempunyai efek spasial dengan ukuran  $n \times 1$ ,
- $\boldsymbol{\varepsilon}$  = Vektor error dengan ukuran  $n \times 1$ ,
- $\mathbf{I}$  = Matriks identitas berukuran  $n \times n$ ,
- $n$  = banyaknya amatan
- $k$  = banyaknya variabel independen

### 2.2 Matriks Pembobot Spasial

Matriks pembobot spasial merupakan matriks yang nilainya menggambarkan hubungan ketetanggaan antara dua lokasi/ wilayah<sup>[5]</sup>. Terdapat beberapa jenis matriks pembobot spasial, yaitu sebagai berikut:

1. *Rook contiguity* (Persinggungan Sisi) adalah wilayah yang bersinggungan sisi dengan wilayah yang sedang diamati maka diberi nilai  $W_{ij} = 1$ , sedangkan wilayah yang lainnya diberi nilai  $W_{ij} = 0$ .
2. *Bishop contiguity* (Persinggungan Sudut) adalah wilayah yang bersinggungan titik sudut dengan wilayah yang sedang diamati maka diberi nilai  $W_{ij} = 1$ , sedangkan wilayah yang lainnya diberi nilai  $W_{ij} = 0$ .
3. *Queen contiguity* (Persinggungan Sisi dan Sudut) adalah wilayah yang bersinggungan sisi dan titik sudut dengan wilayah yang sedang diamati maka diberi nilai  $W_{ij} = 1$ , sedangkan wilayah yang lainnya diberi nilai  $W_{ij} = 0$ .

### 2.3 Uji Moran's I

*Uji Moran's I* merupakan salah satu metode yang dapat digunakan untuk mendeteksi adanya Autokorelasi spasial<sup>[7]</sup>. Indeks *Moran's I* nilainya berada pada rentang  $-1 \leq I \leq 1$ . Indeks Moran's I bernilai positif apabila daerah yang berdekatan memiliki kemiripan atribut, bernilai negatif apabila daerah yang saling berdekatan cenderung berbeda dan bernilai nol apabila nilai atribut tersusun secara acak dan independen di dalam ruang.

Hipotesis:

$H_0$ : Tidak ada autokorelasi spasial antar lokasi

$H_1$ : Ada autokorelasi spasial antar lokasi

Statistik Uji:

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

dimana:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} c_{ij}}{s^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

$$E(I) = I_0 = -\frac{1}{n-1} \quad S_1 = \frac{1}{2} \sum_{i \neq j}^n (w_{ij} + w_{ji})^2$$

$$var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - [E(I)]^2 \quad S_2 = \sum_{i \neq j}^n (w_{i0} + w_{0i})^2$$

$$c_{ij} = (x_i - \bar{x})(x_j - \bar{x}) \quad w_{i0} = \sum_{j=1}^n w_{ij}$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad w_{0i} = \sum_{j=1}^n w_{ji}$$

$H_0$  ditolak jika  $|Z_{hitung}| > Z_{\alpha/2}$

## 2.4 Spatial Autoregressive

Model *Spatial Autoregressive* adalah model yang mengkombinasikan model regresi sederhana dengan lag spasial pada variabel dependen (Y). Model ini digunakan untuk data spasial *cross sectional*<sup>[10]</sup>. Model SAR memiliki bentuk persamaan sebagai berikut<sup>[11]</sup>:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \boldsymbol{\alpha} \mathbf{1}_n + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3)$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

Atau dapat dituliskan sebagai berikut:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{Z} \boldsymbol{\delta} + \boldsymbol{\varepsilon} \quad (4)$$

dengan:

$$\mathbf{Z} = [\mathbf{1}_n \quad \mathbf{X}]$$

$$\boldsymbol{\delta} = \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}$$

$\boldsymbol{\alpha}$  = Vektor parameter konstan berukuran  $n \times 1$

$\mathbf{1}_n$  = Vektor yang berisi angka 1 berukuran  $n \times 1$

### 2.4.1 Estimasi Parameter

Parameter SAR dapat diestimasi dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Fungsi likelihood dibentuk melalui persamaan error ( $\boldsymbol{\varepsilon}$ ) yang berdistribusi normal .

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{Z} \boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} = \mathbf{y} - \rho \mathbf{W} \mathbf{y} - \mathbf{Z} \boldsymbol{\delta}$$

$$\boldsymbol{\varepsilon} = (\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y} - \mathbf{Z} \boldsymbol{\delta}$$

Fungsi kepadatan peluang dari  $\varepsilon_i$  dengan  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$  adalah sebagai berikut:

$$f(\varepsilon_i | \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{(-\frac{\varepsilon_i^2}{2\sigma^2})}$$

didapatkan fungsi logaritma natural untuk mengestimasi  $\rho$ :

$$l(\rho | \mathbf{y}) = C - \frac{n}{2} \ln([e_0 - \rho e_d]^T [e_0 - \rho e_d]) + \ln |\mathbf{I}_n - \rho \mathbf{W}|$$

Dimana C adalah konstanta yang tidak bergantung pada parameter  $\rho$ . Besar parameter  $\rho$  adalah pada rentang  $\frac{1}{\lambda_{max}} < \rho < \frac{1}{\lambda_{min}}$  dengan  $\lambda$  adalah nilai eigen dari matriks  $\mathbf{W}$  yang terstandarisasi.

Estimasi parameter  $\boldsymbol{\delta}$  menjadi  $\widehat{\boldsymbol{\delta}} = \widehat{\boldsymbol{\delta}}_0 - \widehat{\rho} \widehat{\boldsymbol{\delta}}_d$

Estimasi parameter untuk  $\sigma^2$ , yaitu  $\widehat{\sigma^2} = \frac{(e_0 - \widehat{\rho} e_d)^T (e_0 - \widehat{\rho} e_d)}{n}$

## 2.5. Uji Hipotesis dalam Regresi Linier Berganda

### 2.5.1. Uji F

Uji F dapat digunakan untuk menentukan apakah terdapat hubungan secara bersama-sama variabel X terhadap Y [13], dengan hipotesis sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$  (tidak ada pengaruh variabel X secara bersama-sama terhadap variabel Y)

$H_1$ : terdapat  $\beta_j \neq 0$ , dengan  $j = 1, 2, \dots, k$  (paling sedikit ada satu variabel  $X_j$  yang berpengaruh terhadap variabel Y)

Statistik Uji:

$$F_{hitung} = \frac{SSR/k}{SSE/(n-k-1)}$$

$H_0$  ditolak jika nilai  $F_{hitung} > F_{\alpha, k, n-k-1}$

### 2.5.1 Uji Wald

Uji yang paling relevan diterapkan didalam data spasial ada Uji Wald. Uji Wald ini dapat digunakan untuk menguji signifikansi koefisien model secara individu [1], dengan hipotesis sebagai berikut:

$H_0: \beta_j = 0$  (Parameter tidak signifikan)

$H_1: \beta_j \neq 0$  (Parameter signifikan)

Statistik Uji Wald (W):

$$W = \left( \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2 \quad SE(\hat{\beta}_j) = \sqrt{(\sigma^2(\hat{\beta}_j))}$$

Keterangan :

$SE(\hat{\beta}_j)$  = dugaan galat baku untuk koefisien  $\beta_j$

$\hat{\beta}_j$  = nilai dugaan untuk parameter ( $\beta_j$ )

Keputusan:

$H_0$  ditolak jika nilai  $W > \chi^2_{(\alpha, 1)}$

## 2.6 Uji Asumsi

### 2.6.1 Asumsi Normalitas Residual

Salah satu metode yang dapat digunakan adalah Uji Lilliefors [5] dengan hipotesis sebagai berikut:

$H_0$ : Residual berdistribusi normal

$H_1$ : Residual tidak berdistribusi normal

Statistik Uji:

$$T_2 = \text{Sup}|F^*(x) - S(x)|$$

dengan :

$T_2$  : Nilai statistik uji Lilliefors

$F^*(x)$ : Probabilitas kumulatif normal standar

$S(x)$  : Probabilitas kumulatif empiris dari  $Z_i$

Dimana  $Z_i = \frac{X_i - \bar{X}}{s}$ ,  $i = 1, 2, \dots, n$

$H_0$  ditolak bila nilai  $T_2 > T_{2(\alpha; NT)}$  atau  $p\text{-value} < \alpha$ .  $T_{2(\alpha; NT)}$  diperoleh dari nilai tabel Lilliefors.

### 2.6.2 Asumsi Heteroskedastisitas Residual

Untuk mendeteksi adanya efek heteroskedastisitas dapat menggunakan uji Glejser [8], dengan hipotesis sebagai berikut :

- H<sub>0</sub> : Tidak ada gejala heteroskedastisitas  
H<sub>1</sub> : Ada gejala heteroskedastisitas

Statistik Uji:

Langkah-langkah menguji asumsi homoskedastisitas dengan menggunakan uji Glejser adalah sebagai berikut:

1. Dapatkan nilai error dari regresi spasial

$$\varepsilon_i = y_i - \hat{y}_i$$

2. Hitung nilai absolut dari nilai error yang diperoleh ( $|\varepsilon_i|$ ).
3. Regresikan  $|\varepsilon_i|$  dengan variabel independen  $X_1, X_2, \dots, X_k$ .

$$|\varepsilon_i| = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k$$

4. Lakukan pengujian signifikansi parameter individu terhadap model regresi pada langkah 3.

Jika  $\alpha_j$  dengan  $j = 1, 2, \dots, k$  tidak signifikan maka dapat disimpulkan tidak terjadi heteroskedastisitas (asumsi homoskedastisitas terpenuhi), begitu pula sebaliknya.

### 2.6.3 Asumsi Autokorelasi Residual

Untuk mendeteksi adanya autokorelasi dapat menggunakan metode *Durbin-Watson*<sup>[8]</sup> dengan hipotesis sebagai berikut :

H<sub>0</sub>: Tidak ada autokorelasi

H<sub>1</sub>: Ada autokorelasi antar positif atau negatif

Statistik Uji:

$$d = \frac{\sum_{t=2}^{t=N} (e_t - e_{t-1})^2}{\sum_{t=1}^{t=N} e_t^2}$$

dengan:

$d$  = nilai *Durbin – Watson*

$e_t$  = residual pada waktu tertentu (t)

$e_{t-1}$  = residual pada waktu sebelumnya (t-1)

**Tabel 1.** Aturan Keputusan Uji *Durbin – Watson*

Jika	Hasil Keputusan
$0 < d < d_L$	Menolak H <sub>0</sub> ; ada autokorelasi positif
$d_L \leq d \leq d_U$	Daerah keragu-raguan; Pengujian tidak meyakinkan
$d_U \leq d \leq 4 - d_U$	Menerima H <sub>0</sub> dan H <sub>0</sub> <sup>*</sup> ; Tidak ada autokorelasi
$4 - d_U \leq d \leq 4 - d_L$	Daerah keragu-raguan; Pengujian tidak meyakinkan
$4 - d_L \leq d \leq 4$	Menolak H <sub>0</sub> <sup>*</sup> ; ada autokorelasi negatif

### 2.6.4 Asumsi Multikolinearitas

Cara yang digunakan untuk mengetahui adanya multikolinieritas yaitu dengan menggunakan *Variance Inflation Faktor* (VIF). Nilai VIF diperoleh dengan rumus:

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

$R_j^2$  merupakan koefisien determinasi variabel independen ke -  $j$  terhadap variabel independen lainnya di mana  $j = 1, 2, \dots, k$ . Dikatakan terjadi multikolinearitas pada variabel independen apabila nilai VIF <sub>$j$</sub>  lebih besar dari 10<sup>[8]</sup>.

## 2.7. Ukuran Kebaikan Model

### 2.7.1. Koefisien Determinasi (Adjusted R<sup>2</sup>)

Koefisien determinasi (R<sup>2</sup>) adalah besaran yang digunakan untuk menilai kebaikan model regresi<sup>[13]</sup>. Kebanyakan peneliti lebih menyukai menggunakan nilai *Adjusted R<sup>2</sup>* sebagai ukuran kecocokan model. Karena *Adjusted R<sup>2</sup>* hanya akan

meningkat ketika variabel ditambahkan ke model kemudian nilai MSE dari model yang baru berkurang. *Adjusted R<sup>2</sup>* dapat ditentukan dengan rumus sebagai berikut :

$$R^2_{adj} = \frac{SSR}{SST} = 1 - \frac{SSE/(n - k - 1)}{SST/(n - 1)}$$

Dengan SSR = *Regression sum of squared* model SAR

SST = *Total sum of squared* model SAR

SSE = *Sum of squared of the residual error* model SAR

### 2.7.2. Mean Squared Error (MSE)

MSE diperoleh berdasarkan rumus<sup>[13]</sup>:

$$MSE = \frac{SSE}{n-k-1}$$

## 2.8 Pencilan Spasial

*Spatial Outlier* adalah objek yang tereferensi secara spasial dimana atribut non spasialnya relatif sangat berbeda dengan lingkungannya, meskipun mereka tidak berbeda secara signifikan dari populasi lainnya<sup>[15]</sup>. Salah satu metode yang dapat digunakan untuk mendeteksi *Spatial Outlier* adalah dengan *Moran Scatterplot*. Nilai tersebut dapat diidentifikasi berdasarkan rumus sebagai berikut :

$$(Z[f(i)] \times (\sum_j (W_{ij} Z[f(j)])) < 0 \tag{5}$$

$$Z_i = (f(i) - \mu_f) / \sigma_f \tag{6}$$

$$I_i = (\sum_j (W_{ij} Z[f(j)])) \tag{7}$$

Keterangan :

$Z[f(i)]$  = nilai atribut yang distandarisasi,

$Z[f(j)]$  = transpose nilai atribut yang distandarisasi,

$I_i$  = nilai atribut yang distandarisasi dikalikan dengan matriks bobot standarisasi,

$W_{ij}$  = bobot terstandarisasi berukuran n x n,

$\mu_f$  = rata – rata dari nilai atribut,

$\sigma_f$  = standar deviasi dari nilai atribut.

Apabila nilai  $(Z[f(i)] \times (\sum_j (W_{ij} Z[f(j)]))$  lebih kecil dari 0 maka dapat dikatakan sebagai *Spatial Outlier*.

## 2.9 Regresi Robust

Regresi *robust* adalah alat penting yang dapat digunakan untuk menganalisis data yang terkontaminasi dengan pencilan. Estimator di dalam regresi *robust* diantaranya adalah *estimator M*, *estimator LTS*, *estimator S* dan estimator MM. Penggunaan Regresi *robust* memerlukan fungsi pembobot. Ada beberapa jenis bobot yang dapat digunakan untuk Regresi *robust*, diantaranya adalah Tukey Bisquare dan Huber.

## 2.10 Fungsi Objektif

Fungsi objektif adalah fungsi yang digunakan untuk mencari fungsi pembobot pada regresi robust<sup>[6]</sup>. Fungsi objektif Tukey Bisquare adalah sebagai berikut :

$$\rho(u_i) = \begin{cases} \frac{c_2^2}{6} \left\{ 1 - \left[ 1 - \left( \frac{u_i}{c_2} \right)^2 \right]^3 \right\} & , \text{ untuk } |u_i| \leq c_2 \\ \frac{c_2^2}{6} & , \text{ untuk } |u_i| > c_2 \text{ dengan } c_2 = 1.547 \end{cases}$$

## 2.11 Robust Spatial Autoregressive S-Estimator

Estimator S adalah metode *Robust* yang memiliki nilai *breakdown point* paling<sup>[4]</sup>. *Breakdown point* adalah satu nilai yang dapat digunakan untuk mengukur tingkat kekekaratan suatu estimator. Semakin tinggi nilai *breakdown point* maka dapat

dikatakan bahwa suatu estimator memiliki tingkat kekekaran yang semakin tinggi. *Robust Spatial Autoregressive* adalah metode Robust S-Estimator yang digunakan untuk data spasial.

*S-estimator* didefinisikan dengan :

$$\hat{\beta} = \min_{\beta} s(e_1(\beta), e_2(\beta) \dots, e_n(\beta)) \quad (8)$$

$e_i$  merupakan residual ke- $i$  dari  $\beta$  dan  $s(e_1, e_2 \dots, e_n)$  didefinisikan sebagai solusi dari :

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{e_i}{s}\right) = \frac{1}{1} \sum_{i=1}^n \rho\left(\frac{y_i - \sum_{j=0}^k x_{ij}\beta_j}{s}\right) = K \quad (9)$$

Agar *breakdown point* 50%, maka  $K = E_{\Phi}\rho(u_i) = 0.1995^{[14]}$ . Nilai  $s$  dapat dicari dengan<sup>[13]</sup> :

$$s = \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2} \quad (10)$$

Dan pembobot  $w_i = w(u_i) = \frac{\rho(u_i)}{u_i^2}$  untuk iterasi berikutnya.  $S$  untuk iterasi pertama :

$$S = \frac{MAD}{0,6475} = \frac{\text{median}|e_i - \text{median}(e_i)|}{0,6475}$$

Penyelesaian persamaan (8) dapat dilakukan dengan mencari turunan parsial pertama dari  $\rho$  terhadap  $\beta_j$  ( $j = 0, 1, \dots, k$ ) kemudian disama dengankan 0, sehingga didapatkan :

$$\sum_{i=1}^n x_{ij} \psi\left(\frac{y_i - \sum_{j=0}^k x_{ij}\beta_j}{s}\right) = 0, \quad j = 0, 1, \dots, k \quad (11)$$

Taksiran model *Robust Spatial Autoregressive* diasumsikan sebagai  $\theta$ , sehingga dapat diselesaikan sebagai berikut :

$$\theta \boldsymbol{\varepsilon} = \theta (\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y} - \theta \mathbf{Z} \boldsymbol{\delta}$$

Dengan menggunakan metode kuadrat terkecil maka didapatkan :

$$\hat{\boldsymbol{\delta}}_{OLS} = (\mathbf{Z}^T \theta \mathbf{Z})^{-1} \mathbf{Z}^T \theta (\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y} \quad (12)$$

Setelah estimator  $\boldsymbol{\delta}_{OLS}$  diperoleh, maka residual awal dapat ditulis kembali sebagai berikut:

$$\boldsymbol{\varepsilon} = (\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y} - \mathbf{Z} \hat{\boldsymbol{\delta}}_{OLS} \quad (13)$$

Dari fungsi pengaruh dapat diperoleh fungsi pembobot :

$$w(u_i) = \frac{\psi(u_i)}{u_i}$$

Maka fungsi pembobot dapat diubah menjadi :

$$w(u_i) = \frac{\psi\left(\frac{(\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y} - \mathbf{Z} \hat{\boldsymbol{\delta}}_{OLS}}{\frac{MAD}{0,6475}}\right)}{\frac{(\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y} - \mathbf{Z} \hat{\boldsymbol{\delta}}_{OLS}}{\frac{MAD}{0,6475}}}$$

Sehingga diperoleh :

$$\hat{\boldsymbol{\delta}}^m = (\mathbf{Z}^T \mathbf{w}_i^{m-1} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{w}_i^{m-1} (\mathbf{I}_n - \rho \mathbf{W}) \mathbf{y}$$

Dimana  $\mathbf{w}_i = \mathbf{w}(u_i)$  adalah matriks pembobot yang berukuran  $n \times n$  dimana diagonal utamanya berisi pembobot  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ .

Apabila menggunakan pembobot Tukey Bisquare maka nilai  $\mathbf{w}_i$  adalah sebagai berikut :

$$\mathbf{w}(u_i) = \mathbf{w}_i = \frac{\psi(u_i)}{u_i} = \begin{cases} \left[1 - \left(\frac{u_i}{c_2}\right)^2\right]^2 & , \text{ untuk } |u_i| \leq c_2 \\ 0 & , \text{ untuk } |u_i| > c_2 \end{cases}$$



Fungsi  $\psi$  tidak linear, maka estimasi parameter diselesaikan dengan metode iterasi kuadrat terkecil terboboti atau IRLS (*Iteratively Reweighted Least Square*). Dilakukan iterasi sampai diperoleh estimator yang konvergen, yakni selisih nilai  $\hat{\delta}^m$  dan  $\hat{\delta}^{m+1}$  mendekati 0, dan  $m$  merupakan banyaknya iterasi.

## 2.12 Angka Harapan Hidup

Angka Harapan Hidup (AHH) saat lahir sebagai rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. Berdasarkan sebuah penelitian yang telah dilakukan terhadap faktor-faktor yang mempengaruhi Angka Harapan Hidup di Turki pada tahun 1965-2005 bahwa faktor sosial, ekonomi, pelayanan kesehatan dan lingkungan dapat mempengaruhi naik atau turunnya Angka Harapan Hidup<sup>[9]</sup>.

## 2.13 Faktor-Faktor yang Mempengaruhi Angka Harapan Hidup

Ada beberapa faktor yang diduga mempengaruhi Angka Harapan Hidup, yaitu Angka Harapan Sekolah, Jumlah Penduduk Usia > 15 Tahun yang sudah bekerja, Gizi Buruk, Jumlah Penduduk Miskin, Bayi Lahir dengan Berat Badan Rendah dan Persentase Rumah Sehat.

### 2.1.1 Matlab dan *Graphical User Interface* (GUI)

Matlab(Matrix Laboratory) merupakan salah satu bahasa pemrograman yang dikembangkan oleh MathWorks. GUI merupakan tampilan grafis yang memudahkan user berinteraksi dengan perintah teks. Dengan GUI, program yang dibuat menjadi lebih user *friendly*, sehingga user mudah menjalankan suatu aplikasi program

## 3. METODE PENELITIAN

### 3.1. Jenis, Sumber Data dan Variabel Penelitian

Data yang digunakan oleh penulis dalam penulisan penelitian ini adalah data sekunder, yaitu data Angka Harapan Hidup (AHH) sebagai variabel Y, Angka Harapan Sekolah( $X_1$ ), Jumlah Penduduk Usia > 15 Tahun yang Bekerja( $X_2$ ), Gizi Buruk( $X_3$ ), Jumlah Penduduk Miskin( $X_4$ ), Bayi Berat Badan Lahir Rendah( $X_5$ ), dan Rumah Sehat( $X_6$ ) setiap kabupaten/kota di Jawa Timur pada tahun 2017. Data tersebut dapat diakses pada website Badan Pusat Statistik(BPS) Provinsi Jawa Timur <https://jatim.bps.go.id/> dan buku Profil Kesehatan Provinsi Jawa Timur Tahun 2017 yang telah diterbitkan oleh Dinas Kesehatan Provinsi Jawa Timur.

### 3.2. Software yang digunakan

Software statistik yang digunakan dalam penelitian ini adalah Matlab R2013b.

### 3.3. Langkah-langkah Analisis Data

Langkah- langkah analisis data yang akan dilakukan untuk menentukan model *Robust Spasial Autoregressive* adalah sebagai berikut :

1. Mendapatkan data yang akan dianalisis.
2. Menentukan matriks pembobot spasial menggunakan *Queen Contiguity*.
3. Melakukan deteksi autokorelasi spasial dengan uji *Moran's I*.
4. Mengestimasi koefisien regresi menggunakan *Spatial Autoregressive (SAR)*
5. Melakukan uji F dan Uji Wald.
6. Menghitung ukuran kebaikan model.
7. Melakukan uji asumsi pada model *Spatial Autoregressive (SAR)*.
8. Mendeteksi pencilan spasial dengan *Moran's Scatterplot*.
9. Mengestimasi parameter dengan *Robust Spatial Autoregressive S-estimator*.

Prosedur estimasinya adalah sebagai berikut :

- a. Menghitung nilai residual
- b. Menghitung nilai S
- c. Menghitung nilai residual yang distandarisasi,  $u_i = \frac{e_i}{s}$
- d. Menentukan nilai pembobot  $w_i$

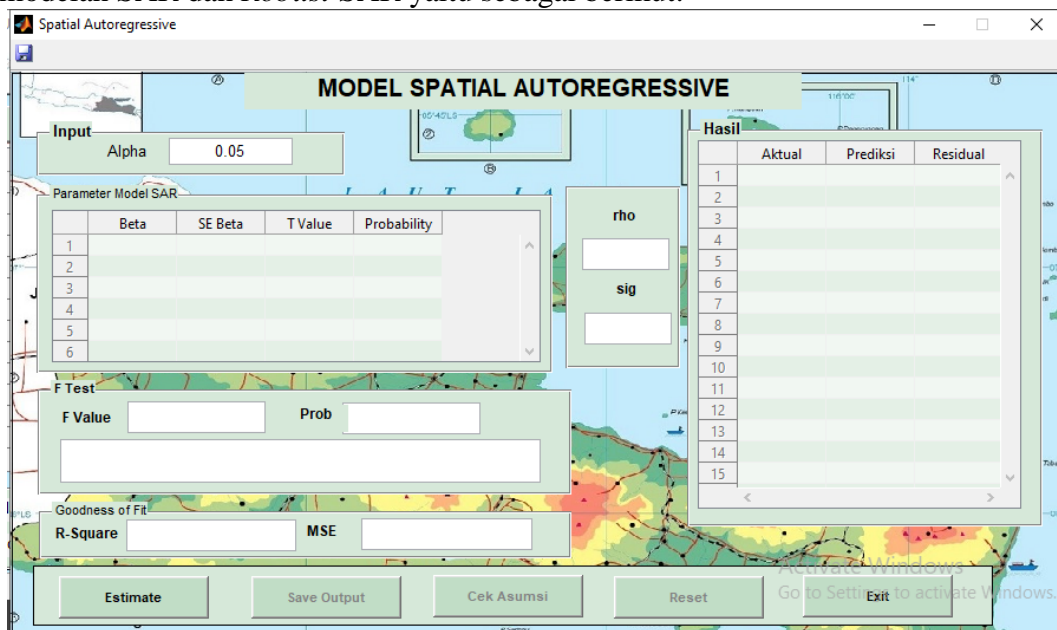


- e. Menghitung estimasi parameter regresi baru ( $\hat{\delta}_j^m$ ) dengan menggunakan nilai pembobot  $w_i$ .
  - f. Mengulang langkah a sampai e untuk mendapatkan nilai  $\hat{\beta}_j$  konvergen.
10. Mendapatkan model *Robust Spatial Autoregressive S-estimator*
  11. Melakukan uji F dan Uji Wald.
  12. Menghitung ukuran kebaikan model.
  13. Menentukan Model regresi terbaik.

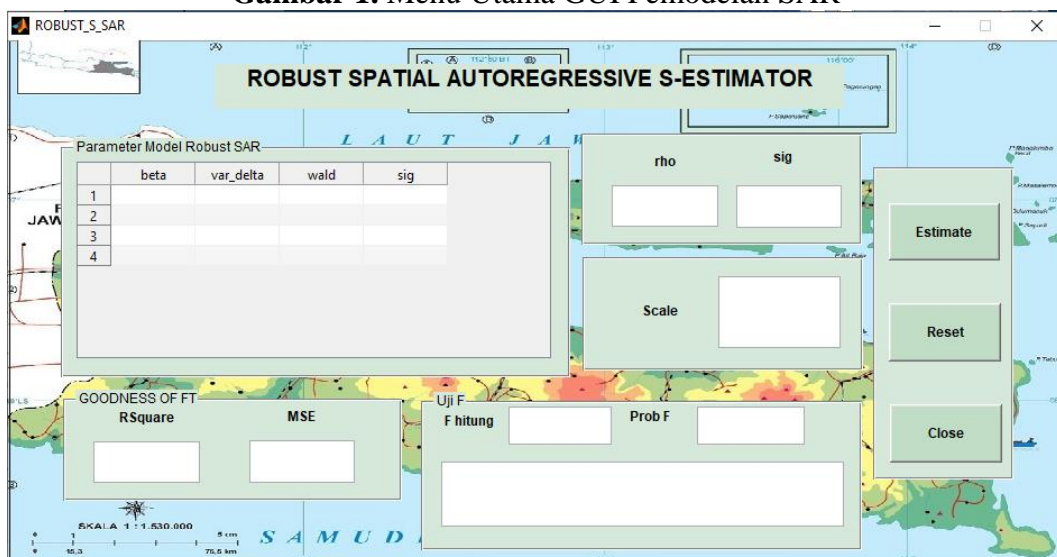
#### 4. HASIL DAN PEMBAHASAN

##### 4.1. Proses Pembuatan GUI Pemodelan SAR dan *Robust SAR*

GUI pemodelan SAR dan *Robust SAR* dibuat melalui empat tahap yaitu mendesain *figure* atau tampilan, mengatur *layout* komponen untuk masing-masing *property inspector*, menyusun perintah (sintaks), serta *finishing*. Hasil pembuatan GUI pemodelan SAR dan *Robust SAR* yaitu sebagai berikut:



Gambar 1. Menu Utama GUI Pemodelan SAR



Gambar 1. Menu Utama GUI Pemodelan *Robust SAR*

## 4.2. Uji Moran's I

Nilai Z Hitung dan Moran's I untuk variabel Y sebesar :

Zhitung = 5,472455      Nilai Moran's I = 0,7133

$H_0$  ditolak jika nilai  $|Z_{hitung}| > Z_{1-\alpha/2}$ , dan nilai  $Z_{0,025} = 1,96$ . Dapat dilihat bahwa untuk variabel Y memiliki nilai  $Z_{hitung} > 1,96$ , yang berarti  $H_0$  ditolak. Hal ini menunjukkan bahwa terdapat autokorelasi spasial antar lokasi pada variabel dependen. Karena variabel dependen menunjukkan adanya autokorelasi spasial maka model SAR dapat digunakan dalam penelitian.

## 4.3. Model Spatial Autoregressive

### 4.5.1. Estimasi Parameter Model SAR dengan 6 Variabel Prediktor

Tabel 1. Estimasi Parameter SAR dengan 6 Variabel Prediktor

Parameter	Koefisien	SE Beta	$\chi^2_{hitung}$	p-value
$\rho$	0,4296	0,11112	14,9440	$1,8904 \times 10^{-8}$
$\beta_0$	33,7438	7,3713	154,4702	1,0000
$\beta_1$	0,2511	0,2105	0,2995	0,4158
$\beta_2$	$1,6571 \times 10^{-6}$	$1,0822 \times 10^{-6}$	$2,5372 \times 10^{-6}$	0,0013
$\beta_3$	-0,0022	0,0018	0,0028	0,0424
$\beta_4$	$4,3120 \times 10^{-6}$	$3,1584 \times 10^{-6}$	5,8870e-06	0,0019
$\beta_5$	-0,0017	$7,9553 \times 10^{-4}$	0,0038	0,0492
$\beta_6$	0,0523	0,0138	0,1977	0,3434

Model SAR dengan 6 Variabel Prediktor adalah sebagai berikut:

$$\hat{y}_i = 33,7438 + 0,4296w_{ij}y_i + 0,2511x_{i1} + 1,6571 \times 10^{-6}x_{i2} - 0,0022x_{i3} + 4,3120 \times 10^{-6}x_{i4} - 0,0017x_{i5} + 0,0523x_{i6}$$

### 4.5.2. Uji Hipotesis Model SAR dengan 6 Variabel Prediktor

#### ✓ Uji F

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab, diperoleh nilai  $F_{hitung} = 16,5428$ , lebih besar dari  $F_{tabel} = 2,41$ . Kemudian diperoleh nilai  $p-value = 1,8904 \times 10^{-8} < \alpha = 0,05$ . Sehingga dapat disimpulkan bahwa  $H_0$  ditolak, yang artinya paling sedikit ada satu variabel  $X_j$  yang berpengaruh terhadap variabel Y.

#### ✓ Uji Wald

Berdasarkan tabel 1, dapat dilihat bahwa  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$  dan  $\beta_5$  berpengaruh secara signifikan terhadap model SAR, karena memiliki nilai  $p-value$  kurang dari  $\alpha = 0,05$ . Sedangkan untuk parameter  $\beta_0$ ,  $\beta_1$  dan  $\beta_6$  berpengaruh secara tidak signifikan terhadap model SAR, karena memiliki nilai  $p-value$  lebih besar dari  $\alpha = 0,05$ .

### 4.5.3. Uji Asumsi Model SAR dengan 6 Variabel Independen

#### ✓ Asumsi Normalitas Residual

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab, diperoleh nilai hitung  $T_2 = 0,1204 < \text{nilai } T_{2(0,05;38)} = 0,1437$  yang diperoleh dari tabel lilliefors dan nilai probabilitas yang didapat adalah sebesar  $0,1708 > \alpha = 0,05$  yang menunjukkan bahwa  $H_0$  diterima atau residual model berdistribusi normal.

✓ **Asumsi Autokorelasi Residual**

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab di dalam Lampiran 6, diperoleh nilai hitung  $d = 1,9345$ . Kemudian didapatkan nilai  $d_U$  dan  $d_L$  yang diperoleh dari tabel *Durbin - Watson* yaitu sebesar 1,8641 dan 1,1463. Dapat dilihat bahwa nilai  $d_U \leq d \leq 4 - d_U$  dan nilai probabilitas yang didapat adalah sebesar  $0,6539 > \alpha = 0,05$  yang menunjukkan bahwa  $H_0$  diterima atau tidak ada autokorelasi positif atau negatif antar residual.

✓ **Asumsi Heteroskedastisitas Residual**

Berdasarkan hasil komputasi dengan GUI Matlab diperoleh nilai  $\text{Prob}(X_1) = 0,9289$ ,  $\text{Prob}(X_2) = 0,9272$ ,  $\text{Prob}(X_3) = 0,4832$ ,  $\text{Prob}(X_4) = 0,7595$ ,  $\text{Prob}(X_5) = 0,4961$ , dan  $\text{Prob}(X_6) = 0,0140$ . Hasil pengujian menunjukkan bahwa ada 1 variabel independen yang signifikan pada tingkat signifikansi 5% dengan probabilitas  $0,0140 < \alpha = 0,05$ . Sehingga dapat disimpulkan asumsi heteroskedastisitas pada model SAR dengan 6 variabel independen tidak terpenuhi.

✓ **Asumsi Multikolinieritas**

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab, diperoleh hasil yaitu nilai  $\text{VIF}(X_1) = 1,4236$ ,  $\text{VIF}(X_2) = 4,2876$ ,  $\text{VIF}(X_3) = 1,5975$ ,  $\text{VIF}(X_4) = 2,7264$ ,  $\text{VIF}(X_5) = 3,2557$  dan  $\text{VIF}(X_6) = 1,6771$ . Dapat dilihat bahwa nilai  $\text{VIF} < 10$  pada semua variabel independen, maka dapat disimpulkan bahwa tidak terjadi multikolinieritas pada variabel independen.

**4.5.4. Estimasi Parameter Model SAR dengan 3 Variabel Prediktor**

**Tabel 2.** Estimasi Parameter SAR dengan 3 Variabel Prediktor

Parameter	Koefisien	SE Beta	$\chi^2_{hitung}$	p-value
$\rho$	0,6203	0,12673	20,556	$2,3167 \times 10^{-6}$
$\beta_0$	27,2734	8.7458	85.0508	1,0000
$\beta_1$	$1,2395 \times 10^{-6}$	$1.1846 \times 10^{-6}$	$1.2970 \times 10^{-6}$	$9.0867 \times 10^{-4}$
$\beta_2$	-0,0025	0.0023	0.0027	0.0414
$\beta_3$	$-4,0552 \times 10^{-6}$	$3.8387 \times 10^{-6}$	$4.2839 \times 10^{-6}$	0.0017

Maka model regresinya adalah sebagai berikut:

$$\hat{y}_i = 0,6203w_{ij}y_i + 27,2734 + 1,2395 \times 10^{-6}x_{i1} - 0,0025x_{i2} - 4,0552 \times 10^{-6}x_{i3}$$

Dengan:

$$i = 1, 2, \dots, 38.$$

$$\hat{y}_i = \text{Dugaan AHH daerah ke } - i$$

$$x_{i1} = \text{Jumlah Penduduk Usia } > 15 \text{ tahun yang bekerja daerah ke } - i$$

$$x_{i2} = \text{Jumlah Gizi Buruk daerah ke } - i$$

$$x_{i3} = \text{Jumlah Penduduk Miskin daerah ke } - i$$

**4.5.5. Pengujian Model SAR dengan 3 Variabel Prediktor**

✓ **Uji F**

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab diperoleh nilai  $F_{hitung} = 14,9067$ , lebih besar dari  $F_{tabel} = 2,88$ . Kemudian diperoleh nilai  $p\text{-value} = 2,31669 \times 10^{-6} < \alpha = 0,05$ . Sehingga Dapat disimpulkan bahwa  $H_0$  ditolak, yang artinya paling sedikit ada satu variabel  $X_j$  yang berpengaruh terhadap variabel  $Y$ .

✓ **Uji Wald**

Berdasarkan tabel 2 dapat dilihat bahwa  $\beta_1$ ,  $\beta_2$  dan  $\beta_3$  berpengaruh secara signifikan terhadap model SAR, karena memiliki nilai *p-value* kurang dari  $\alpha = 0,05$ .

**4.5.6. Uji Asumsi Model SAR dengan 3 Variabel Independen**

✓ **Asumsi Normalitas Residual**

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab diperoleh nilai hitung  $T_2 = 0,2094 > \text{nilai } T_{2(0,05;38)} = 0,1437$  yang diperoleh dari tabel lilliefors dan nilai probabilitas yang didapat adalah sebesar  $0,0010 < \alpha = 0,05$  yang menunjukkan bahwa  $H_0$  ditolak atau residual model tidak berdistribusi normal.

✓ **Asumsi Autokorelasi Residual**

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab diperoleh nilai hitung  $d = 1,8400$ . Kemudian didapatkan nilai  $d_U$  dan  $d_L$  yang diperoleh dari tabel *Durbin - Watson* yaitu sebesar 1,7223 dan 1,2614. Dapat dilihat bahwa nilai  $d_U \leq d \leq 4 - d_U$  dan nilai probabilitas yang didapat adalah sebesar  $0,5099 > \alpha = 0,05$  yang menunjukkan bahwa  $H_0$  diterima atau tidak ada autokorelasi positif atau negatif antar residual.

✓ **Asumsi Homoskedastisitas Residual**

Berdasarkan hasil komputasi dengan GUI Matlab diperoleh nilai  $\text{Prob}(X_1) = 0,8172$ ,  $\text{Prob}(X_2) = 0,5441$  dan  $\text{Prob}(X_3) = 0,9508$ . Hasil pengujian menunjukkan bahwa pada tingkat signifikansi 5% dapat disimpulkan tidak ada gejala heteroskedastisitas pada model SAR dengan 3 variabel independen .

✓ **Asumsi Multikolinieritas**

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab, diperoleh hasil yaitu nilai  $\text{VIF}(X_1) = 2,8009$ ,  $\text{VIF}(X_2) = 1,5234$  dan  $\text{VIF}(X_3) = 2,0721$ . Dapat dilihat bahwa nilai  $\text{VIF} < 10$  pada semua variabel independen, maka dapat disimpulkan bahwa tidak terjadi multikolinieritas pada variabel independen.

**4.5.7. Model Akhir SAR**

Model akhir untuk *Spatial Autoregressive* (SAR) yang dipilih adalah model dengan 3 variabel prediktor yaitu model yang variabel prediktornya berpengaruh signifikan terhadap variabel Y. Model yang terbentuk adalah sebagai berikut :

$$\hat{y}_i = 0,6203w_{ij}y_j + 27,2734 + 1,2395 \times 10^{-6}x_{i1} - 0,0025x_{i2} - 4,0552 \times 10^{-6}x_{i3}$$

**4.5.8. Ukuran Kebaikan Model**

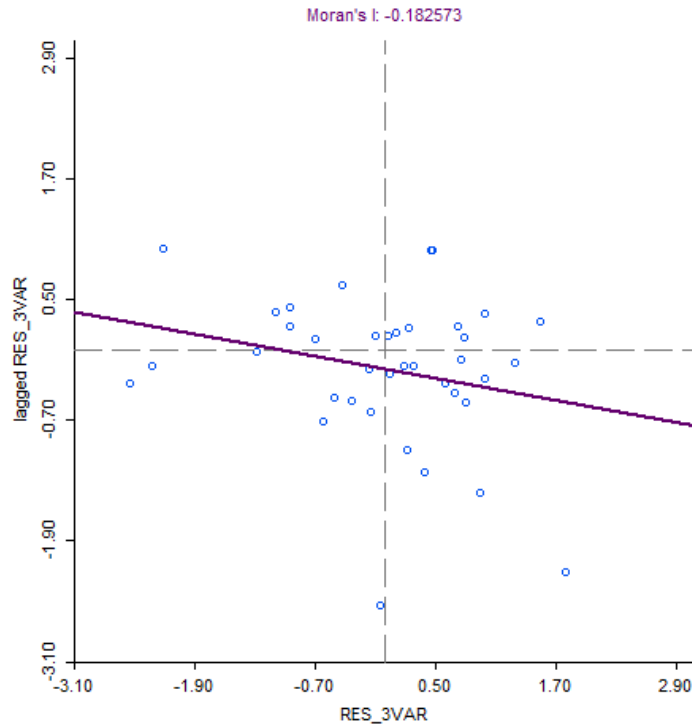
✓ **Adjusted R-Square**

Berdasarkan output di GUI Matlab di dalam Lampiran 6 didapatkan nilai  $\text{Adjusted } R^2 = 0,5554$  atau 55,54% yang artinya bahwa 55,54% Angka Harapan Hidup (Y) dipengaruhi oleh Jumlah Penduduk Usia > 15 Tahun yang Bekerja ( $X_1$ ), Jumlah Penderita Gizi Buruk ( $X_2$ ) dan Jumlah Penduduk Miskin ( $X_3$ ). Sedangkan sisanya 44,46% Angka Harapan Hidup (Y) dipengaruhi oleh faktor yang lain.

✓ **Mean Square Error (MSE)**

Nilai *Mean Square Error* (MSE) model SAR dapat dilihat pada output GUI Matlab di dalam Lampiran 6 yaitu sebesar 1,7521.

#### 4.5.9. Deteksi Pencilan Spasial



**Gambar 1.** Grafik Moran Scatterplot

Dari Grafik Moran Scatterplot dapat disimpulkan bahwa terdapat pencilan spasial, yaitu plot yang tersebar pada kuadran II (Kiri atas) dan Kuadran IV (Kanan Bawah). Selain menggunakan Grafik Moran Scatterplot pencilan spasial juga dapat diidentifikasi menggunakan perhitungan yang sudah disusun dalam GUI Matlab. Output deteksi pencilan spasial dapat dilihat pada Lampiran 6. Dari hasil *output*, didapatkan bahwa ada sebanyak 18 pencilan spasial yang terdeteksi, yaitu data ke 1,2,8, 10, 12, 15, 16, 17, 18, 21, 22, 24, 26, 28, 30, 33, 36 dan 38.

#### 4.6. Robust Spatial Autoregressive S-Estimator

##### 4.6.1. Estimasi Parameter Robust Spatial Autoregressive S-Estimator

**Tabel 13.** Estimasi Parameter Robust SAR

Parameter	Koefisien	SE Beta	$\chi^2_{hitung}$	$p - value$
$\rho$	0,4661	0,12673	20,556	$3,92293 \times 10^{-7}$
$\beta_0$	38,6404	39,8239	236,5986	1,0000
$\beta_1$	$1,0204 \times 10^{-6}$	$8,2336 \times 10^{-13}$	$1,1475 \times 10^{-6}$	$8,5472 \times 10^{-4}$
$\beta_2$	-0,0019	$2,9568 \times 10^{-6}$	0,0022	0,0374
$\beta_3$	$-4,1074 \times 10^{-6}$	$9,9789 \times 10^{-12}$	$5,3406 \times 10^{-6}$	0,0018

Maka model regresinya adalah sebagai berikut:

$$\hat{y}_i = 0,4661w_{ij}y_i + 38,6404 + 1,0204 \times 10^{-6}x_{i1} - 0,0019x_{i2} - 4,1074 \times 10^{-6}x_{i3}$$

#### 4.6.2. Uji Hipotesis Model *Robust SAR S-Estimator*

##### ✓ Uji F

Berdasarkan hasil komputasi dengan menggunakan GUI Matlab diperoleh nilai  $F_{hitung} = 18,4551$  dimana nilai tersebut lebih besar dari  $F_{tabel} = 2,88$ . Kemudian diperoleh nilai  $p\text{-value} = 3,9229 \times 10^{-7} < \alpha = 0,05$ . Sehingga Dapat disimpulkan bahwa  $H_0$  ditolak, yang artinya paling sedikit ada satu variabel  $X_j$  berpengaruh terhadap variabel  $Y$ .

##### ✓ Uji Wald

Berdasarkan Tabel 3, dapat dilihat bahwa  $\beta_1$ ,  $\beta_2$  dan  $\beta_3$  berpengaruh secara signifikan terhadap model *Robust SAR*, karena memiliki nilai  $p\text{-value}$  kurang dari  $\alpha = 0,05$ . Sedangkan untuk parameter  $\beta_0$  tidak berpengaruh secara signifikan terhadap model SAR, karena memiliki nilai  $p\text{-value}$  lebih besar dari  $\alpha = 0,05$ .

#### 4.6.3. Ukuran Kebaikan Model

##### ✓ *Adjusted R-Square*

Berdasarkan output di GUI Matlab didapatkan nilai  $adjusted R^2 = 0,6230$  atau 62,30 % yang artinya bahwa 62,30% Angka Harapan Hidup (Y) dipengaruhi oleh Jumlah Penduduk Usia > 15 Tahun yang Bekerja ( $X_1$ ), Jumlah Penderita Gizi Buruk ( $X_2$ ) dan Jumlah Penduduk Miskin ( $X_3$ ). Sedangkan sisanya 37,70% Angka Harapan Hidup (Y) dipengaruhi oleh faktor yang lain.

##### ✓ *Mean Square Error (MSE)*

Nilai *Mean Square Error (MSE)* model *Robust SAR* dapat dilihat pada output GUI Matlab di dalam Lampiran 7 yaitu sebesar 0,7456.

#### 4.7. Pemilihan Model Terbaik

Tabel 4. Perbandingan MSE Model SAR dan *Robust SAR*

No.	Model	MSE	<i>Adjusted R<sup>2</sup></i>
1.	<i>Spatial Autoregressive (SAR)</i>	1,7521	55,54%
2.	<i>Robust SAR S-Estimator</i>	0,7456	62,30 %

Dari Tabel 4 diperoleh bahwa model *Robust SAR S-Estimator* menghasilkan nilai MSE yang lebih kecil dari model *Spatial Autoregressive (SAR)*, yaitu sebesar 0,7456. Maka dari itu dapat disimpulkan bahwa model *Robust SAR S-Estimator* lebih baik daripada model *Spatial Autoregressive (SAR)* dalam memodelkan variabel AHH beserta faktor-faktor yang mempengaruhinya di Provinsi Jawa Timur ada Tahun 2017.

## 5. PENUTUP

### 5.1. KESIMPULAN

1. GUI pemodelan SAR dan *Robust SAR* dibuat melalui empat tahap yaitu mendesain *figure* atau tampilan, mengatur *layout* komponen untuk masing-masing *property inspector*, menyusun perintah (sintaks) sekaligus meletakkan sintaks dan menyesuaikannya pada lembar komputasi GUI, serta *finishing*.
2. GUI ini dapat digunakan untuk membentuk model SAR dan *Robust SAR* dari berbagai kasus yang variabel dependen dan variabel independennya memenuhi uji *Moran's I*.
3. Menggunakan estimasi parameter *Spatial Autoregressive (SAR)* diperoleh model akhir regresi sebagai berikut :

$$\hat{y}_i = 0,6203w_{ij}y_i + 1,2395 \times 10^{-6}x_{i1} - 0,0025x_{i2} - 4,0552 \times 10^{-6}x_{i3}$$

Asumsi normalitas dan homoskedastisitas tidak terpenuhi, dan terdapat 18 pencilan spasial. Oleh karena itu digunakan metode regresi *Robust Spatial Autoregressive S-Estimator* yang kekar terhadap pencilan.



4. Dengan menggunakan regresi *Robust Spatial Autoregressive S-Estimator* dengan pembobot *Tukey Bisquare*. Kemudian dibandingkan berdasarkan nilai *Adjusted R<sup>2</sup>* dan MSE, diperoleh model terbaik adalah :

$$\hat{y}_i = 0,4661w_{ij}y_i + 1,0204 \times 10^{-6}x_{i1} - 0,0019x_{i2} - 4,1074 \times 10^{-6}x_{i3}$$

5. Angka Harapan Hidup Provinsi Jawa Timur pada Tahun 2017 dipengaruhi oleh Jumlah Penduduk Usia > 15 Tahun yang Bekerja ( $X_1$ ), Jumlah Penderita Gizi Buruk ( $X_2$ ) dan Jumlah Penduduk Miskin ( $X_3$ ) sebesar 62,30 % dan nilai MSE nya 0,7456.

### 5.1. Saran

Efek spasial yang diperhatikan dalam kasus ini hanya dependensi spasial, sedangkan efek spasial yang lain belum ditangani yaitu heteroskedastisitas spasial, sehingga disarankan untuk penelitian selanjutnya, dapat menggunakan model lain yang digunakan untuk menangani efek heteroskedastisitas spasial.

### DAFTAR PUSTAKA

- [1] Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- [2] Away, G. A. 2014. *The Shortcut of Matlab Programming*. Bandung: Informatika Bandung.
- [3] Badan Pusat Statistik Provinsi Jawa Timur. 2017. *Konsep Indeks Pembangunan Manusia*.<https://jatim.bps.go.id/subject/26/indeks-pembangunan-manusia.html#subjekViewTab1>. Diakses : 15 September 2019.
- [4] Chen, C. 2002. *Robust Regression and Outlier Detection with the Robustreg Procedur*. Conference paper. Hal. 1-13.
- [5] Conover, W. L. 1980. *Practical Nonparametric Statistics*. Second Edition. New York: John Wiley and Sons, Inc.
- [6] Fox, J. 2002. *Robust Regression : Appendix to An R and S-Plus Companion to Applied Regression*.
- [7] Goodchild, M. F. 1986. *Spatial Autocorrelation*. Norwich: Geo Books.
- [8] Gujarati, D. N. 1978. *Ekonometrika Dasar*. Diterjemahkan Oleh : Zain, S. Bandung : Universitas Padjajaran. Terjemahan Dari : Basic Econometrics New York: The McGraw-Hill Companies.
- [9] Haliciouglu, F. 2010. *Modelling life expectancy in Turkey*. Yeditepe University
- [10] LeSage, J. P. 1999. *The Theory and Practice of Spatial Econometrics*. Ohio: Department of Economics, University of Toledo.
- [11] LeSage, J., dan Pace, R. K. 2009. *Introduction to Spatial Econometrics*. New York: Taylor & Francis Group..
- [12] Maronna, R. A., Martin, R. D., dan Yohai, V. J. 2006. *Robust Statistics Theory and Methods*. New York : John Wiley and Sons, Ltd.
- [13] Montgomery, D. C., dan Runger, G. C. 2011. *Applied Statistics and Probability for Engineers*. New York: John Wiley & Sons, Inc.
- [14] Rousseeuw, P. J. dan Yohai, V. 1984. "Robust Regression by Means of S Estimators".256-272. New York : Springer Verlag.
- [15] Shekhar, S., Lu, C. T., dan Zhang,P. 2003. "A Unified Approach to Detecting Spatial Outliers". Journal Geoinformatica : Vol 7. No.2 : Hal. 139-166.