

---

## Identifikasi Faktor-Faktor yang Memengaruhi Angka Harapan Hidup di Sumatera Tahun 2018 Menggunakan Analisis Regresi Spasial Pendekatan Area

Evi Ramadhani<sup>1</sup>, Nany Salwa<sup>2\*</sup>, dan Medina Suha Mazaya<sup>3</sup>

<sup>1,2,3</sup>Jurusan Statistika, FMIPA, Universitas Syiah Kuala, Banda Aceh, Indonesia

E-mail: evi.ramadhani@unsyiah.ac.id<sup>1</sup>, nany.salwa@unsyiah.ac.id<sup>2</sup>, medina@mhs.unsyiah.ac.id<sup>3</sup>

\* = Corresponding author

---

### Abstrak

Angka Harapan Hidup (AHH) merupakan perkiraan usia hidup yang dapat dicapai oleh penduduk pada suatu wilayah. AHH merupakan salah satu indikator derajat kesehatan masyarakat suatu negara yang digunakan sebagai tolok ukur dalam mengevaluasi kinerja pemerintah di bidang kesehatan, lingkungan, dan sosial ekonomi. Salah satu faktor yang memengaruhi pencapaian AHH adalah lokasi antar wilayah, sehingga dalam melakukan analisis perlu mempertimbangkan unsur lokasi di dalamnya. Penelitian ini bertujuan mengidentifikasi faktor-faktor yang berpengaruh signifikan terhadap AHH di 154 kabupaten/kota Pulau Sumatera dengan analisis regresi spasial pendekatan area dan mendapatkan model regresi spasial terbaik pada pemodelan AHH Pulau Sumatera. Regresi spasial merupakan analisis statistika untuk memodelkan dan mengevaluasi hubungan antara variabel dependen dan independen dengan memperhatikan keterkaitan unsur lokasi. Model regresi spasial pendekatan area SAR, SEM, dan SARMA dikaji dengan melibatkan 16 variabel independen terpilih dari 17 variabel independen yang teridentifikasi. Data bersumber dari BPS dan IPKM tahun 2018. Hasil penelitian menunjukkan, bahwa model SEM merupakan model regresi spasial pendekatan area terbaik dengan nilai  $R^2$  sebesar 58,23% dan nilai AIC sebesar 600,27. Variabel yang berpengaruh signifikan memengaruhi AHH Pulau Sumatera secara spasial, diantaranya yaitu proporsi balita gizi buruk dan kurang ( $X_1$ ), proporsi desa dengan kecukupan jumlah bidan per 1.000 penduduk ( $X_7$ ), proporsi rumah tangga dengan akses sanitasi ( $X_9$ ), persentase penduduk miskin ( $X_{13}$ ), angka buta huruf penduduk usia 15 tahun ke atas ( $X_{14}$ ), dan rata-rata lama sekolah ( $X_{15}$ ).

### Abstract

Life expectancy is an estimate of the life span that can be achieved by residents in a region. Life expectancy is one of the indicators of a country's public health degree that is used as a benchmark in evaluating government performance in the health, environmental, and socioeconomic fields. One of the factors that influence the achievement of life expectancy is the location between regions, so in conducting the analysis necessary to consider the element of location. This study aims to identify factors that have a significant effect on life expectancy in 154 districts/cities of Sumatra Island with spatial regression analysis of the area approach and to obtain the best model of spatial regression in the life expectancy modeling in Sumatra Island. Spatial regression is a statistical analysis to model and evaluate relationships between dependent variables and independent

---

### Informasi Artikel

#### *Sejarah Artikel:*

Diajukan 23 Agt 2020

Diterima 5 Oktober 2020

---

#### *Kata Kunci:*

Angka Harapan Hidup  
Regresi Spasial  
Pulau Sumatera  
Spatial Error Model  
(SEM)

---

#### *Keyword:*

Life Expectancy  
Spatial Regression  
Sumatera Island  
Spatial Error Model  
(SEM)

---

---

variables by paying attention to interrelations of location elements. The spatial regression model approaches the area of SAR, SEM, and SARMA reviewed with 16 independent variables selected from 17 identified independent variables. Data sourced from BPS and IPKM in 2018. The results show that the SEM model is the best spatial regression model for the area approach with a  $R^2$  value of 58.23% and an AIC value of 600.27. In term of spatial, variables that have a significant effect affect life expectancy in Sumatra Island is the proportion of malnourished and undernourished toddlers ( $X_1$ ), the proportion of villages with the number of adequate of midwives per 1,000 inhabitants ( $X_7$ ), the proportion of households with access to sanitation ( $X_9$ ), the percentage of population live in poverty ( $X_{13}$ ), the illiteracy rate of the population aged 15 years and over ( $X_{14}$ ), and the average length of schooling ( $X_{15}$ ).

---

## 1. Pendahuluan

Angka Harapan Hidup (AHH) adalah perkiraan lama hidup rata-rata penduduk dari sejak dilahirkan, dengan asumsi tidak ada perubahan pola mortalitas menurut umur. AHH digunakan untuk mengukur keberhasilan pembangunan kesehatan penduduk sampai dengan tingkat kabupaten/kota dalam meningkatkan kesejahteraan penduduk pada umumnya, dan meningkatkan derajat kesehatan pada khususnya [1]. Beberapa indikator seperti, indikator kesehatan, indikator kondisi lingkungan, serta indikator sosial ekonomi merupakan faktor utama penyebab AHH bernilai tinggi atau rendah [2].

Berdasarkan data publikasi *Population Reference Bureau*, menunjukkan bahwa AHH di Indonesia setiap tahunnya berada di bawah rata-rata AHH global. AHH di Indonesia sebagai negara berkembang, berkisar antara 69-71 tahun selama tahun 2018-2020. Sedangkan AHH global pada tahun 2018-2020, berkisar antara 72-73 tahun. Indonesia menduduki posisi AHH ke-7 di ASEAN pada tahun 2018-2020 [3].

Pulau Sumatera merupakan pulau dengan kepadatan penduduk terbanyak kedua setelah Pulau Jawa, dengan persentase persebaran penduduk tahun 2018 adalah sebesar 21,80% [4]. Jumlah penduduk yang besar jika diikuti dengan kualitas penduduk yang memadai, dapat menjadi pendorong bagi pertumbuhan ekonomi [5]. Laju pertumbuhan ekonomi secara agregat dapat diasumsikan berdasarkan tingginya nilai Produk Domestik Bruto (PDB) di suatu wilayah/regional dalam satu periode tertentu. Menurut data yang dirilis dari Badan Pusat Statistik (BPS) tahun 2018, PDB Pulau Sumatera berkontribusi sebesar 21,58 % [6].

Capaian rata-rata AHH Pulau Sumatera tahun 2018, yaitu sebesar 68,67 tahun, berada di bawah rata-rata AHH Pulau Jawa sebesar 72,40 tahun dan Pulau Kalimantan sebesar 70,30 tahun. Rata-rata AHH pada seluruh provinsi (10 provinsi) di Pulau Sumatera juga menunjukkan posisi berada di bawah standar nasional [7]. Hal ini menggambarkan, bahwa masih perlunya perhatian dan penanganan lebih lanjut baik dalam segi pelayanan kesehatan, lingkungan, serta sosial ekonomi masyarakat di wilayah Sumatera yang belum terealisasi secara optimal.

Analisis statistik yang dapat menggambarkan hubungan sebab akibat antar variabel dependen dengan variabel independen adalah analisis regresi linier. Jika dalam pemodelan regresi linier data yang dikumpulkan berdasarkan wilayah atau lokasi, analisis yang sering digunakan umumnya menggunakan metode regresi spasial. Regresi spasial merupakan metode pengembangan dari regresi linier yang melibatkan unsur lokasi ke dalam model. Hal ini menunjukkan kemungkinan adanya ketergantungan spasial dalam data. Ketergantungan spasial dan keragaman spasial merupakan pemicu adanya efek spasial.

Jenis pemodelan spasial terdiri atas pendekatan titik dan pendekatan area. Adapun beberapa model utama spasial pada pendekatan area, yaitu *Spatial Autoregressive* (SAR) yang didasari pada efek lag spasial, *Spatial Error Model* (SEM) yang didasari pada efek *error* spasial, dan *Spatial Autoregressive Moving Average* (SARMA) yang didasari pada gabungan efek lag dan *error* spasial.

Berdasarkan uraian di atas, tujuan dalam penelitian ini adalah mengidentifikasi faktor-faktor yang berpengaruh signifikan terhadap AHH Pulau Sumatera dengan analisis regresi spasial pendekatan area dan mendapatkan model regresi spasial pendekatan area terbaik pada pemodelan AHH Pulau Sumatera.

## 2. Tinjauan Kepustakaan

### 2.1 Koefisien Korelasi Pearson

Koefisien korelasi *pearson* adalah suatu hubungan linier antara dua variabel acak independen ( $X$ ) dan dependen ( $Y$ ) yang disimbolkan dengan  $r$ . Ukuran hubungan linier antara variabel  $X$  dan  $Y$  didefinisikan dalam persamaan:

$$r = \frac{n \sum_{i=1}^n X_i Y_i - \left( \sum_{i=1}^n X_i \right) \left( \sum_{i=1}^n Y_i \right)}{\sqrt{\left[ n \sum_{i=1}^n X_i^2 - \left( \sum_{i=1}^n X_i \right)^2 \right] \left[ n \sum_{i=1}^n Y_i^2 - \left( \sum_{i=1}^n Y_i \right)^2 \right]}} \quad (2.1)$$

Nilai koefisien korelasi berkisar antara  $-1 \leq r \leq 1$ , artinya jika nilai  $r$  mendekati 1 atau -1, maka hubungan antara variabel  $X$  dan  $Y$  kuat. Sebaliknya apabila nilai  $r$  mendekati nilai 0 maka hubungan variabel  $X$  dan  $Y$  sangat lemah atau diindikasikan tidak terdapat korelasi [8].

### 2.2 Regresi Linier

Regresi linier merupakan suatu model yang digunakan untuk mengetahui hubungan antara variabel dependen ( $Y$ ) dengan satu atau lebih variabel independen ( $X$ ). Model regresi linier untuk  $p$  variabel independen secara umum ditulis sebagai berikut [9].

$$Y_i = \beta_0 + \sum_{k=1}^p \beta_k X_{ki} + \varepsilon_i \quad (2.2)$$

Keterangan :

$Y_i$  = Nilai amatan variabel dependen pada pengamatan ke- $i$

$X_{ki}$  = Nilai amatan variabel independen ke- $k$  pada pengamatan ke- $i$ , dengan  
 $k = 1, 2, \dots, p$

$\beta_0$  = Nilai intersep model regresi

$\beta_k$  = Koefisien regresi variabel independen ke- $k$

$\varepsilon_i$  = Nilai residual ke- $i$ ,  $\varepsilon_i \sim N(0, \sigma^2 \mathbf{I})$

### 2.3 Matriks Pembobot Spasial

Matriks pembobot spasial ( $\mathbf{W}$ ) adalah matriks yang menunjukkan hubungan ketetanggaan antara wilayah dalam pengamatan [10]. Dalam pembentukan suatu matriks pembobot spasial terdiri dari baris ke- $i$  ( $i = 1, 2, \dots, n$ ) dan kolom ke- $j$  ( $j = 1, 2, \dots, n$ ). Wilayah ke- $i$  yang bertetangga akan diberi bobot/nilai 1 (satu), sementara wilayah ke- $i$  yang tidak bertetangga akan diberi bobot 0 (nol). Berikut adalah bentuk umum dari matriks  $\mathbf{W}$ .

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

Salah satu metode matriks pembobot spasial adalah metode pendekatan jarak tetangga terdekat atau *K-Nearest Neighbor* (K-NN). Penentuan wilayah terdekat (tetangga) dipilih menggunakan pendekatan jarak (*distance*) [10]. Salah satu ukuran jarak yang digunakan untuk menentukan wilayah tetangga adalah *Great Circle Distance* yang didefinisikan dengan persamaan berikut.

$$d_{ij} = r \times \arccos^{-1}[\cos |x_i - x_j| \cos y_i \cos y_j + \sin y_i \sin y_j] \quad (2.3)$$

dimana  $d_{ij}$  menyatakan jarak antara wilayah ke- $i$  dan wilayah ke- $j$ ,  $x_i$  dan  $x_j$  adalah koordinat jarak (*longitude* dan *latitude*) wilayah ke- $i$ ,  $y_i$  dan  $y_j$  adalah koordinat jarak (*longitude* dan *latitude*) wilayah ke- $j$ ,  $r$  adalah radius bumi, jarak *arc* ditentukan jika mil dengan  $r=3959$  dan dalam kilometer dengan  $r=6371$  [11].

Standarisasi terhadap matriks pembobot spasial dilakukan dengan normalisasi baris pembobot (*row normalized weights*). Jika baris ke- $i$  pada matriks pembobot spasial dinormalisasikan, maka setiap baris harus memiliki nilai yang jika dijumlahkan bernilai 1.

#### 2.4 Model Regresi Spasial

Regresi spasial merupakan suatu metode untuk menilai hubungan antara satu variabel dengan beberapa variabel lain dengan memperhatikan efek spasial [12]. Adapun bentuk umum model regresi spasial dinyatakan pada persamaan sebagai berikut.

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \mathbf{u} \quad (2.4)$$

dengan,

$$\mathbf{u} = \lambda \mathbf{W}_2 \mathbf{u} + \boldsymbol{\varepsilon} \quad (2.5)$$

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

Keterangan :

$\mathbf{y}$  = Vektor variabel dependen, berukuran  $n \times 1$

$\mathbf{X}$  = Matriks variabel independen, berukuran  $n \times (k + 1)$

$\boldsymbol{\beta}$  = Vektor koefisien parameter regresi, berukuran  $(k + 1) \times 1$

$\mathbf{W}_1$  = Matriks pembobot spasial, berukuran  $n \times n$

$\mathbf{W}_2$  = Matriks bobot *error* spasial, berukuran  $n \times n$

$\rho$  = Koefisien parameter autoregresif lag spasial

$\lambda$  = Koefisien parameter autoregresif *error* spasial

$\mathbf{u}$  = Vektor residual yang diasumsikan mengandung autokorelasi, berukuran  $n \times 1$

$\boldsymbol{\varepsilon}$  = Vektor residual yang diasumsikan tidak mengandung autokorelasi, berukuran  $n \times 1$

$n$  = Banyaknya amatan atau lokasi ( $i = 1, 2, \dots, n$ )

$k$  = Banyaknya variabel independen ( $k = 1, 2, \dots, p$ )

$\mathbf{I}$  = Matriks identitas, berukuran  $n \times n$

Berdasarkan model umum yang tertera dapat diperoleh model regresi spasial diantaranya sebagai berikut [13].

a. *Spatial Autoregressive (SAR)*

SAR terbentuk apabila nilai  $\rho \neq 0$  dan  $\lambda = 0$ , sehingga model ini mengasumsikan bahwa proses autoregresif hanya pada variabel dependen. Bentuk umum model SAR ditunjukkan dengan persamaan sebagai berikut.

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.6)$$

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

b. *Spatial Error Model (SEM)*

SEM terbentuk apabila nilai  $\rho = 0$  dan  $\lambda \neq 0$ , sehingga model ini mengasumsikan bahwa proses autoregresif hanya pada *error* model. Bentuk umum model SEM ditunjukkan dengan persamaan sebagai berikut.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (2.7)$$

dengan,

$$\mathbf{u} = \lambda \mathbf{W}_2 \mathbf{u} + \boldsymbol{\varepsilon} \quad (2.8)$$

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

c. *Spatial Autoregressive Moving Average (SARMA)*

SARMA terbentuk apabila nilai  $\rho \neq 0$  dan  $\lambda \neq 0$ . Model SARMA merupakan model regresi linier yang variabel residualnya dan variabel dependennya mengandung unsur korelasi spasial. Bentuk umum model SARMA ditunjukkan dengan persamaan sebagai berikut.

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \lambda \mathbf{W}_2 \mathbf{u} + \boldsymbol{\varepsilon} \quad (2.9)$$

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

### 2.5 Uji Moran's I

Uji *Moran's I* merupakan uji ketergantungan spasial pada residual terhadap model regresi linier OLS [14]. Adapun formulasi statistik *Moran's I* dapat didefinisikan sebagai berikut.

$$I = \frac{\boldsymbol{\varepsilon}' \mathbf{W} \boldsymbol{\varepsilon} / S_0}{\boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} / N} \quad (2.10)$$

dimana  $\boldsymbol{\varepsilon}$  merupakan vektor residual dengan ukuran  $n \times 1$  dari model regresi metode pendugaan OLS,  $S_0 = \sum_i \sum_j w_{ij}$ ,  $N$  adalah banyaknya amatan, dan  $\mathbf{W}$  adalah matriks pembobot spasial terstandarisasi berukuran  $n \times n$ . Pengujian dilakukan dengan distribusi normal, seperti dalam persamaan berikut [15].

$$Z(I) = \frac{I - E(I)}{\sigma(I)} \approx N(0,1) \quad (2.11)$$

dengan,

$$E(I) = \frac{tr(\mathbf{M}\mathbf{W})}{N - K} \quad (2.12)$$

$$\sigma(I) = \sqrt{\frac{tr(\mathbf{M}\mathbf{W}\mathbf{M}\mathbf{W}') + tr(\mathbf{M}\mathbf{W})^2 + [tr(\mathbf{M}\mathbf{W})]^2}{(N - K)(N - K + 2)} - [E(I)]^2} \quad (2.13)$$

dimana  $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ ,  $tr$  menyatakan operasi matriks *trace* yakni penjumlahan elemen diagonal pada suatu matriks, dan  $\mathbf{W}$  adalah matriks pembobot spasial terstandarisasi. Adapun hipotesis uji *Moran's I* adalah:

$H_0 : I = 0$  (Tidak terdapat autokorelasi spasial residual model regresi linier)

$H_1 : I \neq 0$  (Terdapat autokorelasi spasial residual model regresi linier)

Pengambilan keputusan  $H_0$  ditolak, apabila  $Z(I) > Z_{\alpha/2}$  atau  $p\text{-value} < \alpha$ . Sehingga, terdapat ketergantungan spasial dan dapat dilanjutkan dalam pemodelan regresi spasial.

### 2.6 Uji Keragaman Spasial

Pengujian efek spasial dilakukan dengan uji keragaman (heterogenitas) yaitu menggunakan uji *Breusch-Pagan test (BP test)* [12].

Pengujian hipotesis heterogenitas spasial adalah sebagai berikut:

$H_0 = \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$  (tidak terdapat keragaman spasial antar lokasi)

$H_1 =$  minimal ada satu  $\sigma_i^2 \neq \sigma^2$  (terdapat keragaman spasial antar lokasi)

Statistik uji BP dapat dilihat pada persamaan berikut [16].

$$BP = \frac{1}{2} \left( \sum_{i=1}^n x_i f_i \right) \left( \sum_{i=1}^n x_i x_i' \right) \left( \sum_{i=1}^n x_i f_i \right) \quad (2.14)$$

dengan  $f_i = \left( \frac{\hat{\varepsilon}_i}{\hat{\sigma}} - 1 \right)$ ,  $\hat{\varepsilon}_i = (y_i - \hat{\beta}'x_i)$ , dan  $\hat{\sigma}^2 = \sum_{i=1}^n \hat{\varepsilon}_i^2$ . Statistik uji BP mengikuti distribusi  $\chi_{(k)}^2$ , dengan  $k$  merupakan banyaknya parameter regresi. Pengambilan keputusan  $H_0$  ditolak, apabila  $BP > \chi_{(k)}^2$  atau  $p\text{-value} < \alpha$ .

### 2.7 Uji Lagrange Multiplier (LM)

Uji *Lagrange Multiplier (LM)* merupakan uji untuk mengidentifikasi adanya pengaruh spasial terhadap model dan menentukan model regresi yang sesuai. Uji LM mencakup *LM lag* dan *LM error*.

Pengujian hipotesis uji LM lag beserta statistik ujinya dalam mengidentifikasi model SAR adalah sebagai berikut [12].

$H_0 : \rho = 0$  (tidak terdapat ketergantungan lag spasial terhadap variabel dependen)

$H_1 : \rho \neq 0$  (terdapat ketergantungan lag spasial terhadap variabel dependen)

Statistik uji yang digunakan yaitu:

$$LM_{\rho} = \frac{\left[ \frac{\boldsymbol{\varepsilon}'\mathbf{W}\mathbf{y}}{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/N} \right]^2}{D} \quad (2.15)$$

dengan,

$$D = \left[ \frac{(\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}})}{\hat{\sigma}^2} \right] + tr(\mathbf{W}'\mathbf{W} + \mathbf{W}\mathbf{W}) \quad (2.16)$$

dimana  $\boldsymbol{\varepsilon}$  adalah vektor residual dari model regresi metode OLS berukuran  $n \times 1$ ,  $\hat{\boldsymbol{\beta}}$  dan  $\hat{\sigma}^2$  dihasilkan dari model regresi metode OLS, dan  $tr$  menyatakan operasi matriks *trace*.

Sedangkan pada pengujian *LM error*, hipotesis beserta statistik ujinya dalam mengidentifikasi model SEM adalah sebagai berikut.

$H_0 : \lambda = 0$  (tidak terdapat ketergantungan *error* spasial terhadap variabel dependen)

$H_1 : \lambda \neq 0$  (terdapat ketergantungan *error* spasial terhadap variabel dependen)  
 Statistik uji yang digunakan yaitu:

$$LM_{\lambda} = \frac{\left[ \frac{\boldsymbol{\varepsilon}' \mathbf{W} \boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} / N} \right]^2}{tr(\mathbf{W}' \mathbf{W} + \mathbf{W} \mathbf{W})} \quad (2.17)$$

Dalam mengidentifikasi model SARMA dapat dideteksi dengan memastikan pengujian pada uji LM lag dan LM *error*, apabila kedua LM tersebut memiliki kriteria keputusan yang tolak  $H_0$  yaitu terdapat ketergantungan lag spasial terhadap variabel dependen serta terdapat ketergantungan *error* terhadap variabel dependen, maka model SARMA menjadi model yang sesuai untuk digunakan. Berikut uji hipotesis yang digunakan:

$H_0 : \rho, \lambda = 0$  (tidak ada pengaruh lag dan *error* pada model)

$H_1 : \rho, \lambda \neq 0$  (ada pengaruh lag dan *error* pada model)

Statistik uji yang digunakan:

$$LM_{SARMA} = E^{-1} \{ (R_y)^2 T - 2R_y R_{\varepsilon} T + (R_{\varepsilon})^2 (D + T) \} \sim \chi^2_{(k)} \quad (2.18)$$

dengan,

$$E = (D+T)T - (T)^2 \qquad R_{\varepsilon} = \frac{\boldsymbol{\varepsilon}' \mathbf{W}_{\varepsilon}}{\sigma^2}$$

$$R_y = \frac{\boldsymbol{\varepsilon}' \mathbf{W}_y}{\sigma^2} \qquad T = tr\{(\mathbf{W}' + \mathbf{W})\mathbf{W}\}$$

$$D = \sigma^2 (\mathbf{W} \mathbf{X} \boldsymbol{\beta})' \mathbf{M} (\mathbf{W} \mathbf{X} \boldsymbol{\beta}), \text{ dimana } \mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}'$$

Kriteria pengambilan keputusan tolak  $H_0$  pada pengujian LM, apabila nilai  $LM > \chi^2_{(k+1)}$  atau *p-value*  $< \alpha$ , sehingga dapat disimpulkan bahwa model yang sesuai adalah SARMA.

### 3. Metode Penelitian

#### 3.1 Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari situs resmi Badan Pusat Statistik (BPS) dan publikasi Kementerian Kesehatan RI tahun 2018. Penelitian ini menggunakan data pada seluruh provinsi (10 provinsi) di Pulau Sumatera yang terdiri atas 154 kabupaten/kota. Rincian variabel yang digunakan dalam penelitian ini, ditampilkan pada Tabel 1.

**Tabel 1** Variabel penelitian

Indikator	Notasi	Definisi Variabel
Pembangunan Manusia	$Y$	Angka Harapan Hidup (AHH)
Lokasi	$i$	Kabupaten/ Kota di Pulau Sumatera
Kesehatan	$X_1$	Proporsi balita gizi buruk dan kurang
	$X_2$	Proporsi anak usia 12-59 bulan yang mendapat imunisasi dasar lengkap
	$X_3$	Proporsi pasangan usia subur usia 15-49 tahun pengguna KB

Indikator	Notasi	Definisi Variabel
	$X_4$	Persalinan oleh tenaga kesehatan di fasilitas kesehatan
	$X_5$	Proporsi kecamatan dengan kecukupan jumlah dokter per 2.500 penduduk
	$X_6$	Proporsi desa dengan kecukupan jumlah posyandu per desa
	$X_7$	Proporsi desa dengan kecukupan jumlah bidan per 1.000 penduduk
	$X_8$	Proporsi penduduk yang memiliki jaminan pelayanan kesehatan
Lingkungan	$X_9$	Proporsi rumah tangga dengan akses sanitasi
	$X_{10}$	Proporsi rumah tangga dengan akses air bersih
Sosial Ekonomi	$X_{11}$	Tingkat Pengangguran Terbuka (TPT)
	$X_{12}$	Besaran pengeluaran per kapita riil yang disesuaikan
	$X_{13}$	Persentase penduduk miskin
	$X_{14}$	Angka buta huruf penduduk usia 15 tahun ke atas
	$X_{15}$	Rata-rata Lama Sekolah (RLS)
	$X_{16}$	Harapan Lama Sekolah (HLS)
	$X_{17}$	Tingkat Partisipasi Angkatan Kerja (TPAK)

### 3.2 *Prosedur Analisis Data*

Adapun tahapan analisis yang dilakukan adalah sebagai berikut.

1. Mempersiapkan data berdasarkan variabel penelitian yang digunakan.
2. Melakukan eksplorasi data secara deskriptif pada data penelitian.
3. Memilih variabel independen dengan analisis korelasi *pearson*
4. Melakukan analisis regresi linier berganda AHH Pulau Sumatera tahun 2018 dengan melibatkan 154 kabupaten/kota dan 17 variabel independen.
5. Membentuk matriks pembobot spasial (**W**) dengan metode *K-Nearest Neighbors (K-NN)*.
6. Menguji efek ketergantungan dan keragaman spasial.
  - a. Uji *Moran's I*  
Pengujian *Moran's I* dilakukan untuk memeriksa autokorelasi spasial terhadap residual model regresi linier.
  - b. Uji keragaman spasial  
Pengujian keragaman spasial dilakukan dengan statistik uji *Breusch-Pagan*.
  - c. Menentukan model regresi spasial.  
Uji *Lagrange Multiplier (LM)* digunakan untuk menentukan model regresi spasial yang akan digunakan dalam penelitian ini.
7. Melakukan pendugaan dan pengujian parameter model regresi spasial.
8. Menginterpretasi model regresi spasial terpilih dan menarik kesimpulan.

## 4. Hasil dan Pembahasan

### 4.1 *Analisis Deskriptif*

Ringkasan statistik dapat memberikan informasi mengenai pemusatan, penyebaran, dan pola kecenderungan data. Nilai rata-rata AHH pada kabupaten/kota di Pulau Sumatera tahun 2018 adalah sebesar 68,68 tahun dengan nilai median sebesar 68,72 tahun. Nilai rata-rata yang lebih kecil dari nilai median menunjukkan bahwa distribusi data AHH Pulau Sumatera tahun 2018 cenderung menjulur ke arah kiri (*negative skewness*). Capaian AHH tertinggi dengan nilai sebesar



73,91 tahun berada pada Kota Bukittinggi dan capaian AHH terendah dengan nilai sebesar 61,44 tahun berada pada Kabupaten Lingga.

#### 4.2 Model Regresi Linier Berganda AHH Pulau Sumatera

Tahap awal menyusun suatu model regresi adalah memilih variabel independen untuk melihat variabel independen apa saja yang memiliki keterkaitan terhadap AHH. Berikut adalah tabel hasil analisis korelasi *pearson* yang disajikan pada Tabel 2.

**Tabel 2** Analisis korelasi *pearson*

Variabel	Nilai	<i>P-value</i>
$X_1$	-0,285	0,000
$X_2$	0,258	0,001
$X_3$	0,197	0,014
$X_4$	0,456	0,000
$X_5$	0,372	0,000
$X_6$	0,401	0,000
$X_7$	-0,475	0,000
$X_8$	-0,013	0,875
$X_9$	0,543	0,000
$X_{10}$	0,426	0,000
$X_{11}$	0,308	0,000
$X_{12}$	0,469	0,000
$X_{13}$	-0,410	0,000
$X_{14}$	-0,173	0,032
$X_{15}$	0,513	0,000
$X_{16}$	0,326	0,000
$X_{17}$	-0,212	0,008

Berdasarkan Tabel 4.1 menunjukkan bahwa, terdapat 16 variabel independen terpilih dalam penelitian ini berdasarkan signifikansi taraf nyata ( $\alpha$ ) sebesar 0,1 (10%) pada pengujian analisis korelasi *pearson*. Terdapat 16 variabel yang terpilih diantaranya, yaitu  $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}, X_{17}$ .

Regresi diterapkan untuk mengidentifikasi hubungan antara peubah respon dan peubah bebas. Jika peubah respon mempunyai nilai biner maka regresi logistik biner dapat diterapkan [17]. Pada kasus ini, regresi linear berganda diterapkan karena peubah respon memiliki skala rasio. Pemodelan regresi linier berganda menghasilkan nilai F sebesar 9,816 ( $p\text{-value} = 0,000$ ). Hal ini menunjukkan bahwa variabel independen berpengaruh secara simultan terhadap variabel dependen pada taraf nyata ( $\alpha$ ) sebesar 0,1 (10%). Dalam pengujian koefisien regresi secara individu, dilakukan untuk menguji ada tidaknya pengaruh dari masing-masing variabel independen terhadap model regresi linier. Nilai penaksir parameter model regresi linier diperoleh dengan menggunakan estimasi *Ordinary Least Square* (OLS). Berikut merupakan uji signifikansi koefisien regresi secara parsial berdasarkan hasil estimasi parameter regresi linier antara variabel dependen dan variabel independen.

**Tabel 3** Uji signifikansi koefisien regresi linier secara parsial

Variabel	Estimasi $\beta$	Standar Error	t-hitung	P-value
<i>Intercept</i>	62,7571	4,2353	14,818	0,0000
$X_1$	-0,0353	0,0347	-1,017	0,3109
$X_2$	0,0087	0,0106	0,825	0,4109
$X_3$	0,0095	0,0147	0,642	0,5219
$X_4$	-0,0021	0,0116	-0,179	0,8583
$X_5$	0,0081	0,0097	0,831	0,4076
$X_6$	0,0056	0,0064	0,874	0,3836
$X_7$	-0,0233	0,0096	-2,416	0,0170**
$X_9$	0,0014	0,0171	1,863	0,0646*
$X_{10}$	0,0014	0,0130	0,105	0,9164
$X_{11}$	0,0906	0,1061	0,854	0,3947
$X_{12}$	-0,0001	0,0001	-0,719	0,4733
$X_{13}$	-0,1321	0,0394	-3,355	0,0010**
$X_{14}$	0,4509	0,1284	3,512	0,0006**
$X_{15}$	0,6156	0,2846	2,163	0,0323**
$X_{16}$	0,0298	0,2679	0,111	0,9115
$X_{17}$	0,0015	0,0370	0,040	0,9680

Keterangan: \*) nyata pada  $\alpha = 0,1$

\*\*) nyata pada  $\alpha = 0,05$

Peubah bebas yang berpengaruh terhadap AHH perlu diidentifikasi. Identifikasi peubah bebas dapat dilakukan dengan memeriksa selang kepercayaan dari parameter atau dengan memeriksa nilai-p dari t hitung. Pada pemeriksaan selang kepercayaan, jika penduga parameter terletak pada selang yang berisi angka nol maka peubah bebas tidak berpengaruh secara signifikan [18]. Pada persoalan ini, nilai p digunakan untuk memeriksa pengaruh dari peubah bebas. Tabel 3 menunjukkan bahwa terdapat 5 variabel independen yang signifikan berpengaruh nyata terhadap AHH kabupaten/kota di Pulau Sumatera dengan taraf nyata ( $\alpha$ ) sebesar 0,1 (10%). Variabel yang signifikan, terdiri atas proporsi desa dengan kecukupan jumlah bidan per 1.000 penduduk ( $X_7$ ), proporsi rumah tangga dengan akses sanitasi ( $X_9$ ), persentase penduduk miskin ( $X_{13}$ ), angka buta huruf penduduk usia 15 tahun ke atas ( $X_{14}$ ), dan rata-rata lama sekolah ( $X_{15}$ ). Hal ini diperoleh dengan melihat nilai *p-value* dari masing masing variabel yang lebih kecil dari  $\alpha$  (0,1). Adapun hasil pendugaan dan pengujian parameter regresi linier untuk 5 variabel yang signifikan dapat dilihat pada Tabel 4 sebagai berikut.

**Tabel 4** Hasil pendugaan parameter regresi variabel yang signifikan pada model regresi linier secara parsial

Variabel	Estimasi $\beta$	Standar Error	t-hitung	P-value
<i>Intercept</i>	61,8721	1,6006	38,655	0,0000
$X_7$	-0,0260	0,0075	-3,471	0,0007
$X_9$	0,0366	0,0139	2,635	0,0093
$X_{13}$	-0,1478	0,0327	-4,519	0,0000
$X_{14}$	0,4627	0,1098	4,213	0,0000
$X_{15}$	0,7396	0,1457	5,076	0,0000

Berdasarkan Tabel 4, maka persamaan model regresi linier yang dibentuk pada pemodelan AHH Pulau Sumatera tahun 2018, dapat ditulis sebagai berikut:

$$\hat{Y} = 61,8721 - 0,0260X_7 + 0,0366X_9 - 0,1478X_{13} + 0,4627X_{14} + 0,7396X_{15}$$

#### 4.3 Pembentukan Matriks Pembobot Spasial

Tahap pertama dalam pembentukan matriks pembobot spasial adalah menentukan nilai  $k$  (jumlah tetangga terdekat) yang paling optimum berdasarkan perhitungan statistik *Moran's I*. Berdasarkan hasil analisis K-NN dengan *Moran's I*, dipilih nilai  $k$  yang paling optimum sebesar  $k = 6$  yang berarti setiap kabupaten/kota akan memiliki 6 tetangga terdekat. Pemilihan nilai  $k$  tersebut ditentukan berdasarkan kriteria nilai *Moran's I* terbesar yang dihasilkan dari proses iterasi numerik pada perhitungan jarak tetangga terdekat. Matriks pembobot spasial yang dibentuk bersifat tidak simetris dengan ukuran  $154 \times 154$ .

#### 4.4 Pengujian Aspek Data Spasial

Uji ketergantungan spasial pada residual dari regresi linier menggunakan *Moran's I*. Nilai statistik uji *Moran's I* yang diperoleh adalah sebesar 4,9689 ( $p\text{-value} = 0,000$ ). Dengan taraf nyata ( $\alpha$ ) yang digunakan sebesar 0,1 (10%), maka dapat disimpulkan bahwa terdapat autokorelasi spasial positif pada residual regresi linier.

Pengujian heterogenitas spasial dilakukan dengan statistik uji *Breusch-Pagan* (BP). Nilai statistik uji BP yang diperoleh adalah sebesar 9,2613 ( $p\text{-value} = 0,9023$ ). Dengan menggunakan taraf nyata ( $\alpha$ ) sebesar 0,1 (10%), maka dapat disimpulkan bahwa tidak terdapat heterogenitas spasial atau ragam spasial homogen. Setelah kedua pengujian efek spasial yang dilakukan terpenuhi, maka dilakukan penentuan model regresi spasial.

Selanjutnya adalah pengujian *Lagrange Multiplier* (LM) dalam menentukan model regresi spasial yang sesuai dalam pemodelan. Hasil dari uji LM dapat dilihat pada Tabel 5 sebagai berikut.

**Tabel 5** Uji *Lagrange Multiplier* (LM)

Statistik Uji	Nilai	$P\text{-value}$	AIC
<i>Lagrange Multiplier</i> lag (SAR)	9,779	0,0018	620,85
<i>Lagrange Multiplier</i> error (SEM)	13,354	0,0003	<b>615,48</b>
<i>Lagrange Multiplier</i> SARMA	13,624	0,0011	617,20

Hasil dari pengujian LM menunjukkan, bahwa ketiga model regresi spasial tersebut signifikan pada taraf nyata ( $\alpha$ ) sebesar 0,1 (10%). Berdasarkan perolehan nilai AIC terkecil sebesar 615,48 dari tiga model tersebut, menunjukkan bahwa model SEM adalah model terpilih yang sesuai digunakan dalam menganalisis ketergantungan spasial pada AHH Pulau Sumatera. Sehingga, model regresi spasial yang digunakan dalam penelitian ini adalah *Spatial Error Model* (SEM).

#### 4.5 *Spatial Error Model* (SEM) Sebagai Model Regresi Spasial Terbaik

Pendugaan parameter regresi spasial menggunakan metode *maximum likelihood*. Hasil pendugaan dan pengujian parameter regresi spasial pada model SEM disajikan pada Tabel 6 sebagai berikut.

**Tabel 6** Hasil pendugaan parameter regresi spasial model SEM

Variabel	Estimasi $\beta$	Standar Error	Z-hitung	P-value
<i>Intercept</i>	62,152	3,9436	15,7601	0,0000
Lambda ( $\lambda$ )	0,5169	0,1099	4,6998	0,0000
$X_1$	-0,0551	0,0304	-1,8145	0,0610*
$X_2$	0,0118	0,0098	1,2062	0,2278
$X_3$	0,0175	0,0149	0,7913	0,4288
$X_4$	-0,0031	0,0106	-0,2961	0,7671
$X_5$	0,0057	0,0085	0,6672	0,5046
$X_6$	0,0047	0,0063	0,7436	0,4571
$X_7$	-0,0227	0,0094	-2,4222	0,0154**
$X_9$	0,0268	0,0161	1,6654	0,0958*
$X_{10}$	0,0109	0,0118	0,9248	0,3551
$X_{11}$	0,0813	0,0945	0,8605	0,3895
$X_{12}$	-0,0001	0,0001	-0,7151	0,4745
$X_{13}$	-0,1010	0,0367	-2,7525	0,0059**
$X_{14}$	0,3657	0,1221	2,9938	0,0028**
$X_{15}$	0,5974	0,2571	2,3235	0,0202**
$X_{16}$	-0,0265	0,2466	-0,1075	0,9144
$X_{17}$	0,0183	0,0337	0,5424	0,5876

Keterangan: \*) nyata pada  $\alpha = 0,1$   
\*\*) nyata pada  $\alpha = 0,05$

Tabel 6 menunjukkan bahwa terdapat 6 variabel independen yang signifikan berpengaruh nyata terhadap AHH Pulau Sumatera dengan taraf nyata yang digunakan sebesar 0,1 (10%). Variabel yang signifikan, terdiri atas proporsi balita gizi buruk dan kurang ( $X_1$ ), proporsi desa dengan kecukupan jumlah bidan per 1.000 penduduk ( $X_7$ ), proporsi rumah tangga dengan akses sanitasi ( $X_9$ ), persentase penduduk miskin ( $X_{13}$ ), angka buta huruf penduduk usia 15 tahun ke atas ( $X_{14}$ ), dan rata-rata lama sekolah ( $X_{15}$ ).

Hasil pendugaan dan pengujian parameter regresi spasial untuk variabel yang signifikan dapat dilihat pada Tabel 7 berikut.

**Tabel 7** Hasil pendugaan parameter regresi spasial variabel yang signifikan pada model SEM

Variabel	Estimasi $\beta$	Standar Error	Z-hitung	P-value
<i>Intercept</i>	63,0491	1,8622	33,8568	0,0000
Lambda ( $\lambda$ )	0,5019	0,0982	5,1124	0,0000
$X_1$	-0,0667	0,0279	-2,3902	0,0168
$X_7$	-0,0232	0,0081	-2,8565	0,0043
$X_9$	0,0357	0,0143	2,5035	0,0123
$X_{13}$	-0,1048	0,0347	-3,0191	0,0025
$X_{14}$	0,3869	0,1127	3,4344	0,0006
$X_{15}$	0,6805	0,1567	4,3421	0,0000

Berdasarkan Tabel 7, maka persamaan model SEM untuk AHH di kabupaten/kota pada Pulau Sumatera dapat ditulis dalam persamaan sebagai berikut:

$$\hat{y}_i = 63,0491 - 0,0667X_1 - 0,0232X_7 + 0,0357X_9 - 0,1048X_{13} + 0,3869X_{14} \quad (4.1)$$

$$+ 0,6805X_{15} + u_i,$$

$$\text{dengan } u_i = 0,5019 \sum_{j=1, i \neq j}^n w_{ij} u_j + \varepsilon_i. \quad (4.2)$$

Hasil dari koefisien residual spasial ( $\lambda$ ) yang nyata secara spasial menunjukkan bahwa, terdapat adanya keterkaitan pengaruh residual spasial antar satu kabupaten/kota dengan kabupaten/kota lainnya yang menjadi tetangganya sebesar 0,5019 dikalikan dengan rata-rata residual pada wilayah tetangganya.

## 5. Kesimpulan

Berdasarkan hasil penelitian, maka dapat disimpulkan:

1. Faktor-faktor yang memengaruhi AHH Pulau Sumatera berdasarkan hasil analisis regresi spasial pendekatan area terbaik yaitu *Spatial Error Model* (SEM) dari pengujian *Lagrange Multiplier* (LM) adalah proporsi balita gizi buruk dan kurang ( $X_1$ ), proporsi desa dengan kecukupan jumlah bidan per 1.000 penduduk ( $X_7$ ), proporsi rumah tangga dengan akses sanitasi ( $X_9$ ), persentase penduduk miskin ( $X_{13}$ ), angka buta huruf penduduk usia 15 tahun ke atas ( $X_{14}$ ), serta rata-rata lama sekolah ( $X_{15}$ ). Faktor-faktor tersebut berpengaruh signifikan pada taraf nyata sebesar 0,1 (10%).
2. Proporsi balita gizi buruk dan kurang ( $X_1$ ), proporsi desa dengan kecukupan jumlah bidan per 1.000 penduduk ( $X_7$ ), dan persentase penduduk miskin ( $X_{13}$ ) berpengaruh negatif terhadap AHH Pulau Sumatera. Sedangkan pada proporsi rumah tangga dengan akses sanitasi ( $X_9$ ), angka buta huruf penduduk usia 15 tahun ke atas ( $X_{14}$ ), dan rata-rata lama sekolah ( $X_{15}$ ) berpengaruh positif terhadap AHH Pulau Sumatera.

## Daftar Pustaka

- [1] Badan Pusat Statistik. (2020). *Angka Harapan Hidup*. Diakses pada 20 Desember 2020, dari <https://sirusa.bps.go.id/sirusa/index.php/indikator/48>.
- [2] Halicioglu, F. (2011). Modeling Life Expectancy in Turkey, *Economic Modeling. Jurnal Publikasi Universitas Yeditepe*, 28(5), 2075–2082.
- [3] Population Reference Bureau. (2020). *2020 World Population Data Sheet*. Washington DC: Population Reference Bureau Inc.
- [4] Kementerian Kesehatan RI. (2019). *Profil Kesehatan Indonesia Tahun 2018*. Jakarta: Balitbangkes.
- [5] Ekwarso, E. & Sari, L. (2010). Penyerasian Kebijakan Kependudukan di Provinsi Riau. *Jurnal Ekonomi*, 18(02), 36–49. <https://doi.org/http://dx.doi.org/10.31258/je.18.02.p.%25p>.
- [6] Badan Pusat Statistik. (2019). *Ekonomi Indonesia Triwulan IV-2018 Tumbuh 5,17 Persen*. Diakses pada 15 Desember 2020, dari <https://www.bps.go.id/pressrelease/2019/02/06/1619/ekonomi-indonesia-2018-tumbuh-5-17-persen.html>.
- [7] Badan Pusat Statistik. (2018). *Statistik Indonesia Tahun 2018*. Jakarta: Badan Pusat Statistik RI.
- [8] Walpole, R. E. (1992). *Pengantar Statistika: Edisi ke-3*. Terjemahan dari *Introduction to Statistics 3<sup>rd</sup> ed*, oleh Ir. Bambang Sumantri. Jakarta: PT Gramedia Pustaka Utama.

- [9] Draper, N. R., & Smith, H. (1992). *Applied Regression Analysis* (2nd ed.). New York: John Wiley and Sons Inc.
- [10] LeSage, J. P., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Boca Raton: CRC Press.
- [11] Sarrias, M. (2020). *Lecture 1: Introduction to Spatial Econometrics*. Chile: Universidad de Talca.
- [12] Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- [13] Ward, M. D., & Gleditsch, K. S. (2008). *Spatial Regression Model*. United States: Sage Punlicaton Inc.
- [14] Anselin, L., & Bera, A. (1998). *Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics*. New York: Marcel Dekker.
- [15] Lee, J., & Wong, D. W. (2001). *Statistical analysis with ArcView GIS*. Canada: John Wiley & Sons Inc.
- [16] Arbia, G. (2006). *Spatial Econometrics: Statistical Foundation Application to Regional Convergence*. Berlin: Springer.
- [17] Nurhasanah, Rusyana, A. and Fitriana, AR. 2021. Binary logistic regression for identification of high school student interest in Banda Aceh city in continuing study at Universitas Syiah Kuala. *J. Phys.* **1882** 012034.
- [18] Marzuki, Sofyan, H. dan Rusyana, A., 2010. Pendugaan Selang Kepercayaan Persentil Bootstrap Nonparametrik untuk Parameter Regresi. *Statistika*, **10**(1), pp.13-23.