

MODEL REGRESI SPASIAL UNTUK ANAK TIDAK BERSEKOLAH USIA KURANG 15 TAHUN DI KOTA MEDAN

MUSFIKA RATI, ESTHER NABABAN, SUTARMAN

Abstrak. Penelitian ini dilakukan untuk menentukan model anak tidak bersekolah usia kurang 15 tahun di kota Medan menggunakan regresi spasial, menganalisis faktor-faktor yang mempengaruhinya serta mengkaji efektifitas metode regresi spasial dalam menganalisis kasus tersebut. Analisis yang digunakan yaitu Spatial Autoregressive Model (SAR). Hasil analisis menunjukkan bahwa variabel prediktor yang mempengaruhi variabel respon adalah jumlah penduduk prasejahtera, jumlah sekolah SD, dan rasio antara anak yang bersekolah dengan anak tidak bersekolah (ATB) kurang 15 tahun. Nilai koefisien determinasi (R^2) adalah 95.70 %.

1. PENDAHULUAN

Regresi spasial merupakan pengembangan dari metode regresi linier klasik. Pengembangan itu berdasarkan hukum Tobler geografi pertama yang menyatakan bahwa segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang dekat mempunyai pengaruh yang lebih daripada sesuatu yang jauh. Ini berarti adanya pengaruh tempat atau spasial pada data yang dianalisis [1]. Beberapa peneliti yang mengembangkan regresi

Received 20-01-2013, Accepted 23-02-2013.

2010 Mathematics Subject Classification: 91B72

Key words and Phrases: Regresi spasial, Spatial Autoregressive Model (SAR), anak tidak bersekolah.

spasial diantaranya ialah Anselin, *et al.* [2] dan LeSage [3]. Selain itu, metode ini juga telah banyak digunakan sebagai alat analisis data pada beberapa bidang, diantaranya ialah untuk memodelkan harga jual apartemen di Surabaya [4] dan penggunaan regresi spasial pendekatan titik dalam memodelkan kejadian diare dengan studi kasus di Kabupaten Tuban Jawa Timur [5].

Banyaknya anak tidak bersekolah di suatu daerah sangat mungkin dipengaruhi oleh lingkungan atau kondisi geografis daerahnya, termasuk posisinya terhadap daerah lain. Regresi spasial dapat digunakan untuk menentukan dan menganalisis model anak tidak bersekolah di suatu lokasi, yang dikaji dari sisi pengaruh antara lokasi lain disekitarnya.

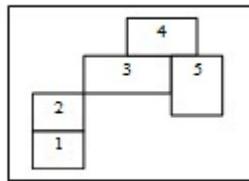
2. LANDASAN TEORI

2.1. Matrik Ketetanggaan Spasial

Bentuk umum matrik spasial \mathbf{W} ialah:

$$\mathbf{W} = \begin{pmatrix} W_{11} & \cdots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{n1} & \cdots & W_{nn} \end{pmatrix}.$$

Pembentukan matrik \mathbf{W} diperoleh dari berbagai teknik pembobotan. Salah satu teknik ialah berdasarkan persentuhan batas wilayah (*contiguity*) yang menyatakan interaksi spasial yang terjadi antar wilayah yang bertetangga. Secara umum terdapat berbagai tipe interaksi, yaitu *Rook contiguity*, *Bishop contiguity* dan *Queen contiguity* [3]. Ilustrasi ketetanggaan ini disajikan pada Gambar 1 berikut.



Gambar 1: Ilustrasi *Contiguity*

Matriks normalitas \mathbf{W} yang merefleksikan *queen contiguity* pada Gambar 1 adalah:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 1/3 & 1/3 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix}.$$

2.2. Model Regresi Spasial

Model umum regresi spasial atau juga biasa disebut *Spatial Autoregressive Moving Average* (SARMA) dalam bentuk matriks [2] dinyatakan sebagai berikut:

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

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \varepsilon \quad (2)$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

dengan

β : vektor koefisien parameter regresi dengan ukuran $(k + 1) \times 1$

ρ : parameter koefisien spasial lag variabel dependen

λ : parameter koefisien spasial lag pada *error*

\mathbf{u}, ε : vektor *error* dengan ukuran $n \times 1$

\mathbf{W} : matriks pembobot dengan ukuran $n \times n$

n : jumlah amatan atau lokasi

k : jumlah variabel independen ($k = 1, 2, \dots, l$)

\mathbf{I} : matriks identitas dengan ukuran $n \times n$

Pada persamaan (1), jika nilai $\rho \neq 0$ dan $\lambda = 0$ maka model regresi spasial akan menjadi model regresi spasial *Mixed Regressive-Autoregressive* atau *Spatial Autoregressive Model (SAR)* atau disebut juga *Spatial Lag Model (SLM)* [1]. Bentuk persamaannya ialah:

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

$$\varepsilon \sim N(0, \sigma^2 I)$$

Pada persamaan (1) jika nilai $\lambda \neq 0$ atau $\rho = 0$ maka model regresi spasial akan menjadi model *Spatial Error Model (SEM)*. Bentuk persamaannya ialah:

$$\mathbf{y} = \mathbf{X}\beta + \lambda \mathbf{W}\mathbf{u} + \varepsilon \quad (4)$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

Penduga masing-masing model regresi spasial dalam bentuk matriks untuk SARMA dan SAR [1] ialah:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \rho \mathbf{W}) \mathbf{y} \quad (5)$$

dan pada SEM ialah:

$$\hat{\beta} = [(\mathbf{X} - \lambda \mathbf{W} \mathbf{X})^T (\mathbf{X} - \lambda \mathbf{W} \mathbf{X})]^{-1} (\mathbf{X} - \lambda \mathbf{W} \mathbf{X})^T (\mathbf{y} - \lambda \mathbf{W} \mathbf{y}) \quad (6)$$

Statistik uji signifikansi parameter yang dipergunakan ialah:

$$Z_{hitung} = \frac{\hat{\theta}}{s.b_{(\hat{\theta})}}$$

Pada persamaan tersebut, $s.b_{(\hat{\theta})}$ merupakan *asymptotic standard error* dan $\hat{\theta}$ adalah parameter regresi spasial (β , λ , dan ρ) dengan hipotesis $H_0 : \hat{\theta} = 0$ dan $H_1 : \hat{\theta} \neq 0$. Apabila $Z_{hitung} \geq Z_{(\alpha/2)}$ atau $p\text{-value} < \alpha/2$, maka keputusan tolak H_0 , artinya koefisien regresi layak digunakan pada model.

2.3. Efek Spasial

Dependensi spasial terjadi akibat adanya dependensi dalam data wilayah. *Spatial dependence* muncul berdasarkan hukum Tobler I (1979). Uji untuk mengetahui *spatial dependence* di dalam suatu model adalah statistik *Moran's I* dan *Langrange Multiplier* (LM) [1].

Moran's I adalah sebuah tes statistik lokal untuk melihat nilai autokorelasi spasial, yang digunakan untuk mengidentifikasi suatu lokasi dari pengelompokan spasial atau autokorelasi spasial. Rumus *Moran's I* dalam bentuk normalitas, yaitu:

$$I = \frac{e' W e}{e' e} \quad (7)$$

dengan $e_i = y_i - \frac{1}{n} \sum_{i=1}^n y_i$ adalah suatu vektor deviasi untuk rata-rata sampel dan $W = [w_{ij}]$ adalah matrik bobot spasial. Nilai ekspektasi dari *Moran's I* [6] adalah:

$$E(I) = I_o = -\frac{1}{n-1} \quad (8)$$

Jika $I > I_0$, maka nilai autokorelasi bernilai positif, hal ini berarti bahwa pola data membentuk kelompok (*cluster*), $I = I_0$ artinya tidak terdapat autokorelasi spasial, dan $I < I_0$ artinya nilai autokorelasi bernilai negatif, hal ini berarti pola data menyebar. Selanjutnya uji LM digunakan untuk menentukan kehadiran efek spasial atau tidak didalam model. Bentuk tes LM [1], yaitu:

a). Pada SEM

$$LM = (1/T)(e'Wy/\sigma^2)^2 \sim \chi^2(1) \quad (9)$$

$$T = \text{trace}[(W + W) \cdot * W]$$

$$\sigma^2 = \frac{e'e}{n}$$

b). Pada SAR

$$LM = (e'Wy/\sigma^2)[T_{22} - (T_{21})^2 \text{var}(\rho)]^{-1} \sim \chi^2(1) \quad (10)$$

$$T_{22} = \text{trace}(W \cdot * W + W'W)$$

$$T_{21} = \text{trace}(W \cdot * CA^{-1} + W'CA^{-1})$$

$$A = (I_n - \rho C)$$

dimana $\cdot *$ melambangkan operasi perkalian titik pada elemen matriks. Hipotesis yang digunakan pada Uji *Lagrange Multiplier* (LM) ialah:

1. Untuk SAR, $H_0 : \lambda = 0$ dan $H_1 : \lambda \neq 0$
2. Untuk SEM, $H_0 : \rho = 0$ dan $H_1 : \rho \neq 0$
3. Untuk *mixture* Model, $H_0 : \rho, \lambda = 0$ dan $H_1 : \rho, \lambda \neq 0$

Dalam mengambil keputusan, tolak H_0 jika $LM > \chi^2$ atau $p\text{-value} < \alpha$.

3. METODE PENELITIAN

Pengolahan data dilakukan dengan urutan [7] sebagai berikut:

- a. Melakukan eksplorasi peta tematik untuk mengetahui pola penyebaran dan dependensi pada masing-masing variabel serta *scatterplot* untuk mengetahui pola hubungan variabel X dan Y .

- b. Melakukan pemodelan regresi dengan metode *Ordinary Least Square* (OLS) yang meliputi estimasi parameter dan estimasi signifikansi model.
- c. Uji dependensi atau korelasi.
- d. Identifikasi keberadaan efek spasial dengan uji *Lagrange Multiplier* (LM) dan *Moran's I Statistics* [1].
- e. Proses pemodelan, yaitu data dimodelkan dengan *Spatial Autoregressive Model* (SAR), *Spatial Error Model* (SEM), atau *Spatial Autoregressive Moving Average* (SARMA).

4. HASIL DAN PEMBAHASAN

Berikut adalah data variabel prediktor dan respon di dalam Tabel 1. Variabel yang diteliti dipilih berdasarkan pertimbangan ketersediaan data yang ada pada BAPEDDA Kota Medan.

Tabel 1: Data Variabel Prediktor dan Respon

No	Nama Kecamatan	Y	X_1	X_2	X_3	X_4
1	M. Tuntungan	150	2547	36	9	14,99
2	M. Johor	234	5017	47	30	21,34
3	M. Amplas	96	3711	38	1	41,43
4	M. Denai	293	5634	69	10	21,18
5	M. Area	96	2267	41	10	24,92
6	M. Kota	68	2142	40	1	31,66
7	M. Maimun	92	1926	22	8	21,05
8	M. Polonia	128	2048	16	9	16,78
9	M. Baru	18	566	24	0	30,00
10	M. Selayang	143	2784	28	13	17,17
11	M. Sunggal	376	3650	40	19	9,31
12	M. Helvetia	227	4015	52	27	17,41
13	M. Petisah	57	1473	22	12	24,30
14	M. Barat	202	2377	27	4	11,39
15	M. Timur	135	3571	43	8	23,17
16	M. Perjuangan	140	3649	32	6	25,79
17	M. Tembung	249	4529	41	20	19,04
18	M. Deli	464	6821	51	77	15,47

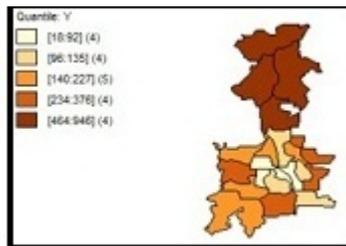
(sambungan)

No	Nama Kecamatan	Y	X_1	X_2	X_3	X_4
19	M. Labuhan	643	6512	46	16	10,28
20	M. Marelan	685	7707	49	38	11,53
21	M. Belawan	946	9201	41	53	10,48

Keterangan:

 Y = Jumlah Anak Tidak Bersekolah (ATB) usia kurang 15 tahun X_1 = Jumlah penduduk prasejahtera X_2 = Jumlah sekolah SD X_3 = Jumlah anak bekerja usia kurang 15 tahun X_4 = Rasio anak bersekolah dengan ATB usia kurang 15 tahun

Berdasarkan jumlahnya, Anak Tidak Bersekolah (ATB) dapat dikelompokan atau dipetakan menjadi 5 wilayah. Pemetaan tersebut dapat dilihat pada Gambar 2 yaitu:



Gambar 2: Peta Tematik ATB di kota Medan

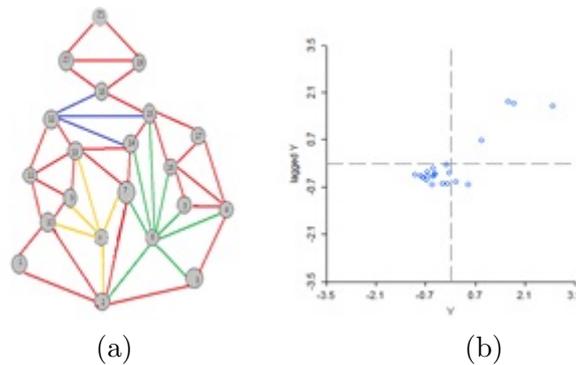
Sumber: OpenGeoda

Berdasarkan Gambar 2, lima wilayah kota Medan tersebut disajikan dalam Tabel 2 ialah:

Tabel 2: Jumlah ATB Berdasarkan Wilayah

No	Wilayah 1 (18 - 92)	Wilayah 2 (96 - 135)	Wilayah 3 (140 - 227)	Wilayah 4 (234 - 376)	Wilayah 5 (464 - 946)
1	Petisah	Timur	Helvetia	Sunggal	Belawan
2	Baru	Area	Perjuangan	Johor	Labuhan
3	Maimun	Polonia	Tuntungan	Tembung	Marelan
4	Kota	Amplas	Barat	Denai	Deli
5	-	-	Selayang	-	-

Selanjutnya, disajikan *graph contiguity* yaitu gambar ketetanggaan kecamatan satu dengan yang lain di Kota Medan sebagai dasar untuk pembuatan matrik spasial W , dan diagram *Moran's* yaitu diagram untuk melihat pola pengelompokan data beserta nilai-nilai *Moran's I*.

Gambar 3: (a) *Graph Contiguity*, (b) *Moran's Scatterplot*

Tabel 3: *Moran's I*

	<i>Moran's I</i>
Y	0,699580
X ₁	0,640032
X ₂	0,298701
X ₃	0,249088
X ₄	0,285518

Pada Tabel 3, terlihat bahwa semua nilai *Moran's I* bernilai lebih besar dari $I_0 = -0,05$ yang artinya semua data pada variabel baik bebas maupun terikat berautokorelasi positif.

Berikutnya pada Tabel 4 diperlihatkan pengaruh jumlah tetangga dengan anak tidak bersekolah.

Tabel 4: Pengaruh Jumlah Tetangga dengan ATB

No	Nama kecamatan	Jlh ATB	Jlh Tetangga	Nama kecamatan	Jlh ATB	Jlh Tetangga
1	Tuntungan	150	2	Johor	234	6
2	Amplas	96	3	Denai	293	5
3	M. Area	96	3	M. Kota	68	8
4	Maimun	92	5	Polonia	128	5
5	M. Baru	18	4	Selayang	143	5
6	Sunggal	376	4	Helvetia	227	5
7	Petisah	57	6	M. Barat	202	5
8	M. Timur	135	6	Perjuangan	140	5
9	Tembung	249	3	M. Deli	464	4
10	Labuhan	643	3	Marelan	685	3
11	Belawan	946	2	-	-	-

Tabel 4 memperlihatkan bahwa semakin banyak jumlah tetangga pada suatu kecamatan relatif mengakibatkan semakin sedikit jumlah anak tidak bersekolah di kecamatan tersebut. Sebagai contoh pada M. Belawan yang memiliki 2 tetangga merupakan kecamatan yang paling banyak jumlah anak tidak bersekolah yaitu 946 orang. Begitu pula pada kecamatan M. Baru yang memiliki 4 tetangga yang merupakan kecamatan yang paling rendah jumlah anak tidak bersekolah yaitu 18 orang. Hal ini dapat disebabkan faktor lain yang tidak turut dikaji dalam penelitian ini, seperti ada atau

tidaknya sekolah SMP di kecamatan tersebut.

Berikutnya pada Tabel 5 diperlihatkan hasil uji dependensi spasial sebagai berikut:

Tabel 5: Hasil Analisis Dependensi Spasial

Uji Dependensi Spasial	Nilai	Prob
<i>Moran's I</i> (error)	2,1297	0,0332
<i>Lagrange Multiplier</i> (lag)	5,9335	0,0149
<i>Lagrange Multiplier</i> (error)	0,6934	0,4050
<i>Lagrange Multiplier</i> (SARMA)	5,9444	0,0512
Taraf signifikan $\alpha = 0,05$		

Berdasarkan Tabel 5, diketahui bahwa nilai probabilitas dari *Moran's I* sebesar 0,0332 lebih kecil dari α . Artinya ada dependensi spasial dalam eror regresi. Begitu pula LM (lag) sebesar 0,0149 lebih kecil dari α artinya terdapat dependensi lag sehingga perlu dilanjutkan ke pembuatan *Spatial Autoregressive Model* (SAR).

Berdasarkan uji LM, telah diketahui bahwa pada kasus ATB di kota Medan terdapat pengaruh spasial dalam data. Hal ini mengidentifikasi bahwa pemodelan kurang akurat dengan menggunakan metode OLS karena pada OLS mengabaikan unsur spasial dalam data. Maka pemodelan akan diselesaikan dengan menggunakan regresi spasial.

Selanjutnya disajikan hasil estimasi regresi pada OLS dan SAR pada Tabel 6 yaitu:

Tabel 6: Hasil Estimasi Koefisien Regresi pada OLS dan SAR

Metode	OLS	SAR
Konstanta	165,8063	128,3035
X_1	0,1141	0,0842
X_2	-4,7741	-3,0008
X_3	-1,3169	-0,9627
X_4	-7,6519	-7,3893
R^2	0,9372	0,9570
Rho (ρ)		0,3048
Jmlh kuadrat eror	74364,7300	53566,7820
Taraf signifikansi (α) = 0,05		

Dari Tabel 6, diperoleh model umum *Spatial Autoregressive* (SAR) sebagai berikut:

$$y_i = 128,3035 + 0,3048 \sum_{j=1, i \neq j}^n w_{ij} y_j + 0,0842X_1 - 3,0008X_2 - 7,3893X_4 + \varepsilon \quad (14)$$

Berikut salah satu contoh model SAR pada kecamatan Medan Belawan yang merupakan daerah yang paling banyak memiliki anak tidak bersekolah. Pada kecamatan tersebut memiliki 2 tetangga yaitu M. Labuhan dan M. Marelan. Persamaan tersebut ialah:

$$\hat{y}_{21} = 128,3035 + 0,1524y_{19} + 0,1524y_{20} + 0,0842X_1 - 3,0008X_2 - 7,3893X_4 \quad (15)$$

Pada persamaan (15) terlihat bahwa apabila faktor lain dianggap konstan, maka ketika jumlah penduduk kesejahteraan di suatu kecamatan (X_1) naik sebesar 100 orang maka bisa menambah ATB usia kurang 15 tahun di M. Belawan sebesar 8 orang. Jumlah sekolah SD di suatu kecamatan (X_2) naik 100 buah maka mengurangi ATB sebesar 300 orang. Rasio anak bersekolah dengan ATB (X_4) naik 100 satuan maka juga akan mengurangi jumlah ATB sebesar 739 orang. Selanjutnya banyak ATB di kecamatan M. Belawan juga dipengaruhi kecamatan tetangganya yaitu M. Labuhan dan M. Marelan sehingga jika banyak ATB usia kurang 15 tahun pada M. Labuhan naik sebesar 100 orang maka akan menambah ATB usia kurang 15 tahun pada kecamatan M. Belawan sebesar 15 anak dan apabila banyak ATB usia kurang 15 tahun pada M. Marelan naik sebesar 100 orang maka akan menambah ATB usia kurang 15 tahun pada kecamatan M. Belawan sebesar 15 orang.

Pada Tabel 6 dapat dilihat bahwa model SAR memiliki nilai R^2 sebesar 95,70% dan jumlah kuadrat eror yang lebih kecil dari OLS yaitu sebesar 53566,782. Jumlah variabel yang berpengaruh pada SAR adalah jumlah penduduk prasejahtera (X_1), jumlah sekolah (X_2) dan rasio anak bersekolah dengan ATB kurang 15 tahun (X_4) sedangkan pada jumlah anak yang bekerja usia kurang 15 tahun (X_3) tidak menjadi variabel yang mempengaruhi di dalam pemodelan kasus anak tidak bersekolah usia kurang 15 tahun di kota Medan.

5. KESIMPULAN

Dari hasil dan pembahasan, dapat diambil kesimpulan sebagai berikut:

1. Model SAR adalah model regresi spasial yang digunakan karena pada kasus anak tidak bersekolah usia kurang 15 tahun di kota Medan hanya bergantung pada lag saja.
2. Model SAR di setiap kecamatan adalah berbeda satu dengan yang lain karena memiliki ketergantungan spasial yang berbeda-beda.
3. Koefisien determinasi pada model SAR sebesar 95,70% dan jumlah kuadrat eror sebesar 53566,782.
4. Semakin banyak jumlah tetangga pada suatu kecamatan relatif mengakibatkan semakin sedikitnya jumlah anak tidak bersekolah di kecamatan tersebut.

Penelitian ini dapat dilanjutkan dengan mengkaji pengaruh variabel lain seperti luas wilayah dan lain-lain.

Daftar Pustaka

- [1] Anselin, L. *Spatial Econometrics : Methods and Models*. Netherlands: Kluwer Academic Publishers, (1988).
- [2] Anselin, L., Syahbri, I., dan Youngihn, K. 2004. *GeoDa: An Introduction to Spatial Data Analysis*. Urbana: University of Illionis, (2004).
- [3] LeSage, J. P. *Spatial Econometrics. Departement of Economics*. University of Toledo, (1998).
- [4] Halim., S., Anastasya, S., Evalina, A., Tobing, A. F.. Penentuan harga jual hunian pada apartemen di Surabaya dengan menggunakan metode regresi spasial. *Jurnal Teknik Industri* 10(2), hal. 151-157, (2008).
- [5] Sutikno dan Arumsari, N. *Permodelan Kejadian Diare Dengan Pendekatan Regresi Spasial*, Studi Kasus : Kabupaten Tuban Jawa Timur. Seminar Nasional Pascasarjana X ITS, Surabaya, (2010).
- [6] Lee, J dan Wong, D. W. S. *Statistical Analysis with Arcview GIS*. John Willey and Sons Inc. New York, (2001).

- [7] Septiana, L. dan Wulandari, S.P. *Pemodelan Remaja Putus Sekolah Usia SMA di Provinsi Jawa Timur dengan Menggunakan Metode Regresi Spasial*, (2009). digilib.its.ac.id/public/ITS-Undergraduate-16199-Cover_id-pdf.pdf.

MUSFIKA RATI: Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Sumatera Utara, Medan 20155, Indonesia
E-mail: musfikatorati654@gmail.com

ESTHER NABABAN: Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Sumatera Utara, Medan 20155, Indonesia
E-mail: esther@usu.ac.id

SUTARMAN: Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Sumatera Utara, Medan 20155, Indonesia
E-mail: sutarman@usu.ac.id