Statistika, Vol. 12 No. 1, 1 – 8
Mei 2012

Regresi Spasial untuk Menentuan Faktor-faktor Kemiskinan di Provinsi Jawa Timur

ANIK DJURAIDAH DAN AJI HAMIM WIGENA

Departemen Statistika FMIPA-IPB, Kampus IPB Darmaga, Bogor
E-mail: anikdjuraidah@gmail.com; ajihamim@yahoo.com

ABSTRAK

Dalam menentukan suatu wilayah kabupaten tergolong miskin umumnya masih digunakan analisis regresi. Padahal kemiskinan sangat mungkin terpengaruh oleh ruang dan daerah sekitarnya. Kondisi ini menyebabkan data antar pengamatan sulit memenuhi asumsi saling bebas sebagai salah satu asumsi pada analisis regresi. Analisis yang dapat mengakomodir masalah spasial ini adalah model otoregresif spasial (spatial autoregressive models, SAR), model galat spasial galat (spatial error models, SEM), dan model spasial umum (spatial general models, SGM). Tujuan penelitian ini adalah menentukan faktor-faktor yang mempengaruhi kemiskinan dengan model regresi spasial. Hasil penelitian menunjukkan bahawa model terbaik adalah SAR dan faktor-faktor yang mempengaruhi kemiskinan adalah persentase penduduk yang tidak tamat Sekolah Dasar (SD) atau tidak bersekolah, persentase penduduk yang menggunakan air minum yang tidak berasal dari air mineral, air PAM, pompa air, sumur atau mata air terlindung, dan persentase penduduk yang menempati rumah dengan kategori sehat yaitu dengan luas lantai lebih dari 8 m2.
Kata kunci : Regresi spasial, model otoregresif spasial, model galat spasial, kemiskinan

ABSTRACT

The determination of whether areas are considered poor are usually based on the average cost per capita with a global analysis that needs independent observations and the results are applied to all villages. But it is very likely that poverty would be influenced by space and neighboring areas, so the data between observations are rarely independent. The statistical analysis that encounters this spatial problem are spatial autoregressive models (SAR), spatial error models (SEM), and spatial general models (SGM). The objective of this study is determining the factors that affect poverty using spatial regression model. The results showed that the best model is SAR and the factors that affect poverty are the percentage of people who did not complete primary school (SD), the percentage of people who drink another kind of water instead of drinking water, and the percentage of people who live in healthy houses with floor area at least 8 m2 per capita.
Keywords: spatial regression, spatial autoregressive models, spatial error models, poverty.

1. PENDAHULUAN

Persoalan kemiskinan masih menjadi salah satu masalah besar di Indonesia. Hasil survey Badan Pusat Statistik (BPS) Maret 2008 menyatakan jumlah orang miskin di Indonesia sebanyak 34.96 juta jiwa atau 15.42 persen dari total jumlah penduduk (BPS 2008). Strategi penanggulangan kemiskinan lebih efektif dengan pendekatan geografis yang akan berhubungan dengan sumber daya alam dan manusia. Hakim & Zuber (2008) menyatakan bahwa lokasi tempat tinggal, akses ke teknologi, dan ketersediaan sumber alam berpengaruh terhadap kemiskinan.

Kemiskinan suatu wilayah dipengaruhi oleh kemiskinan di wilayah sekitarnya. Hal ini berdasarkan hukum geografi yang dikemukakan Tobler (Tobler's *first law of geography*) dalam Schabenberger dan Gotway (2005), yang menyatakan "*everything is related to everything else, but near things are more related than distant things*". Segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang lebih dekat akan lebih berpengaruh daripada sesuatu yang jauh. Hukum Tobler digunakan sebagai pilar kajian analisis data spasial. Pada data spasial, seringkali pengamatan di suatu lokasi (*space*) bergantung pada pengamatan di lokasi lain yang berdekatan (*neighboring*).

Dalam menentukan suatu wilayah tergolong miskin, analisis yang digunakan biasanya masih bersifat global, misalnya analisis regresi. Model analisis ini akan memberikan informasi yang reliabel untuk wilayah yang lebih kecil (wilayah lokal) jika tidak ada atau hanya ada sedikit keragaman antar wilayah lokal tersebut (Fotheringham *et al.* 2002). Dalam analisis regresi sendiri, salah satu asumsi yang diperlukan adalah antar pengamatan harus saling bebas. Sementara kondisi kemiskinan suatu desa sangat mungkin dipengaruhi oleh lokasi pengamatan atau kondisi geografis desa, termasuk posisinya terhadap desa lain di sekitarnya. Hal ini akan menyebabkan asumsi kebebasan antar pengamatan dalam analisis regresi sulit terpenuhi. Anselin (1988) menjelaskan apabila model regresi klasik digunakan sebagai alat analisis pada data spasial, maka bisa menyebabkan kesimpulan yang kurang tepat karena unsur spasial tidak tercakup didalamnya.

Komponen yang mendasar dari model spasial adalah matriks pembobot spasial, matriks inilah yang mencerminkan adanya hubungan antara satu wilayah dengan wilayah lainnya (Grasa 1989). Banyak metode dalam membuat matriks pembobot. Salah satunya dengan pendekatan area berupa ketetanggaan antar wilayah (LeSage 1999). Diharapkan penggunaan model regresi spasial ini mampu menentukan faktor-faktor yang berpengaruh terhadap kemiskinan di setiap wilayah sehingga dapat dijadikan salat satu rujukan dalam program pengentasan kemiskinan yang tepat sasaran.

Penelitian ini bertujuan untuk menentukan faktor-faktor yang berpengaruh terhadap kemiskinan dengan menggunakan model regresi spasial yaitu model otoregresif spasial (*spatial autoregressive model,* SAR), model galat spasial (*spatial error model,* SEM), dan model umum spasial (*general spatial model,* GSM).

2. REGRESI DAN REGRESI SPASIAL

Pada bab ini akan dijelaskan tentang prinsip dasar regresi spasial. Penjelasan ini diawali dari regeresi linear yang telah banyak dikenal.

2.1. Regresi Linear

Persamaan regresi linear yang biasa didefinisikan dengan menggunakan metode pendugaan parameter *Ordinary Least Square* (OLS), secara umum dapat dituliskan:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_p X_{ij} + \varepsilon_i\ , \qquad i = 1,2,\dots,n\,; j = 1,\dots p$$

dengan y_i merupakan peubah respon, β_0 merupakan konstanta, X_{ij} merupakan nilai peubah bebas ke-j pada amatan ke-i, β_p merupakan nilai koefisien peubah penjelas X_j dan ε_i merupakan galat acak pengamatan ke-i. Vektor galat $\boldsymbol{\varepsilon}$ diasumsikan menyebar $N(\mathbf{0}, \sigma^2\mathbf{I})$.

Dalam notasi matriks, persamaan regresi di atas dapat dituliskan sebagai:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Pendugaan $\boldsymbol{\beta}$ dilakukan dengan menggunakan metode OLS yaitu dengan meminimumkan jumlah kuadrat galat $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$. Nilai $\boldsymbol{\beta}$ diduga dengan rumus:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

dimana $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)'$ adalah vektor $p + 1$ sebagai koefisien regresi, $\mathbf{X}$ adalah matriks peubah penjelas berukuran $n \times (p + 1)$ dengan kolom pertama bernilai 1 untuk konstanta, dan $\mathbf{Y}$ adalah vektor peubah respon.

2.2 Model Umum Regresi Spasial (GSM)

Bentuk persamaan model umum regresi spasial sebagai berikut :

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

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

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

dengan $\mathbf{y}$ adalah vektor peubah respon berukuran $n \times 1$, $\mathbf{X}$ adalah matriks peubah penjelas berukuran $n \times (p + 1)$, $\boldsymbol{\beta}$ adalah vektor koefisien parameter regresi yang berukuran $p \times 1$, ρ adalah koefisien otokorelasi spasial pada peubah respon yang bernilai $|\rho| < 1$, λ adalah koefisien otokorelasi spasial pada galat yang bernilai $|\lambda| < 1$, $\mathbf{u}$ adalah vektor galat berukuran

$n \times 1$, $\mathbf{W}$ adalah matriks pembobot spasial yang berukuran $n \times n$, $\boldsymbol{\varepsilon}$ adalah galat acak yang diasumsikan menyebar normal dengan nilaitengah $\mathbf{0}$ dan ragam $\sigma^2\mathbf{I}$, dan n adalah banyak pengamatan.

Pendugaan parameter pada model GSM diperoleh dengan metode penduga kemungkinan maksimum (**Anselin 1988**). Dari persamaan (1) dapat dinyatakan dalam bentuk:

$$\mathbf{y} - \rho\mathbf{W}\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad \text{atau}$$

$$(\mathrm{I} - \rho\mathbf{W})\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{3}$$

Dan dari persamaan (2) dapat dinyatakan dalam bentuk:

$$(\mathbf{I} - \lambda\mathbf{W})\mathbf{u} = \boldsymbol{\varepsilon} \quad \text{atau}$$

$$\mathbf{u} = (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon} \tag{4}$$

persamaan (4) disubstitusi ke persamaan (3) diperoleh:

$$(\mathbf{I} - \rho\mathbf{W})\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon}$$

$$(\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon} = (\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}$$

jika semua ruas dikalikan dengan $(\mathbf{I} - \lambda\mathbf{W})$, maka:

$$\boldsymbol{\varepsilon} = (\mathbf{I} - \lambda\mathbf{W})(\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta} \tag{5}$$

Nilai fungsi kemungkinan (*likelihood*) dari galat $\boldsymbol{\varepsilon}$ adalah:

$$\mathrm{L}(\sigma^2; \boldsymbol{\varepsilon}) = \mathrm{c}(\boldsymbol{\varepsilon})|\mathbf{V}|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{V}^{-1}\boldsymbol{\varepsilon}\right] \tag{6}$$

dengan $\mathbf{V}$ adalah matriks ragam-koragam dari $\boldsymbol{\varepsilon}$ yang bernilai $\mathbf{V} = \sigma^2\mathbf{I}$. Determinan matriks $\mathbf{V}$ adalah σ^{2n} dan kebalikan dari matriks ragam koragam dari $\mathbf{V}^{-1} = 1/(\sigma^2\mathbf{I})$. Dengan mensubstitusikan nilai $|\mathbf{V}|$ dan $\mathbf{V}^{-1}$ pada persamaan (6) maka diperoleh:

$$\mathrm{L}(\sigma^2; \boldsymbol{\varepsilon}) = \mathrm{c}(\boldsymbol{\varepsilon})\sigma^{2n}\exp\left[-\frac{1}{2\sigma^2}\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon}\right] \tag{7}$$

Dari hubungan $\boldsymbol{\varepsilon}$ dan $\mathbf{y}$ pada persamaan (5), didapatkan nilai Jacobian:

$$\mathrm{J} = \left|\frac{\partial\boldsymbol{\varepsilon}}{\partial\mathbf{y}}\right| = |\mathbf{I} - \lambda\mathbf{W}||\mathbf{I} - \rho\mathbf{W}|$$

Dengan mensubstitusikan persamaan (5) ke dalam persamaan (7) diperoleh fungsi kemungkinan untuk $\mathbf{y}$ yaitu:

$$\mathrm{L}(\rho, \lambda, \sigma^2, \boldsymbol{\beta}; \mathbf{y}) = \mathrm{c}(\mathbf{y})(\sigma^{2n})^{-\frac{1}{2}}|\mathbf{I} - \lambda\mathbf{W}||\mathbf{I} - \rho\mathbf{W}|$$
$$\exp\left[-\frac{1}{2\sigma^2}\{(\mathbf{I} - \rho\mathbf{W})[(\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]\}^{\mathbf{T}}\{(\mathbf{I} - \rho\mathbf{W})[(\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]\}\right] \tag{8}$$

dan fungsi log kemungkinan (*log-likelihood*) diperoleh persamaan (8) berikut:

$$\mathrm{l}(\rho, \lambda, \sigma^2, \boldsymbol{\beta}; \mathbf{y}) = \mathrm{c}(\mathbf{y}) - \frac{n}{2}\ln(\sigma^2) + \ln|\mathrm{I} - \lambda\mathrm{W}| + \ln|\mathrm{I} - \rho\mathrm{W}|$$
$$- \frac{1}{2\sigma^2}\{(\mathbf{I} - \rho\mathbf{W})[(\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]\}^{\mathbf{T}}\{(\mathbf{I} - \rho\mathbf{W})[(\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]\} \tag{9}$$

Misalkan kuadrat matriks pembobot $(\mathbf{I} - \rho\mathbf{W})^{\mathbf{T}}(\mathbf{I} - \rho\mathbf{W})$ dinotasikan sebagai $\boldsymbol{\Omega}$ dan penduga $\boldsymbol{\beta}$ diperoleh dengan memaksimalkan fungsi log kemungkinan pada persamaan (9), akan diperoleh penduga $\boldsymbol{\beta}$ yaitu:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Omega}(\mathbf{I} - \lambda\mathbf{W})\mathbf{y}$$

2.3 Model Otoregresif Spasial (SAR)

Jika $\rho \neq 0$ dan $\lambda = 0$, maka persamaan (1) menjadi

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{10}$$

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

Peubah respon pada model SAR berkorelasi spasial. Fungsi log kemungkinan (*log-likelihood*) model SAR diperoleh dari persamaan (9) dengan menggantikan nilai $\lambda = 0$ dan akan diperoleh

$$\mathrm{l} = \mathrm{L}(\boldsymbol{\beta}, \rho, \sigma^2; \mathbf{y})$$
$$= \ln\left(\frac{|\mathrm{I} - \rho\mathbf{W}|}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{(\mathrm{y} - \rho\mathbf{W}\mathrm{y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}(\mathrm{y} - \rho\mathbf{W}\mathrm{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right]\right)$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 + \ln|I - \rho\mathbf{W}| - \frac{(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{2\sigma^2} \quad (11)$$

Pendugaan untuk $\sigma^2, \boldsymbol{\beta}$ dan ρ diperoleh dengan memaksimumkan fungsi log kemungkinan pada persamaan (11) yaitu :

$$\hat{\sigma}^2 = \frac{(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{y}-\rho\mathbf{W}\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})}{n} \quad (12)$$

Persamaan (12) dapat ditulis sebagai:

$$\hat{\sigma}^2 = \frac{\sum(y_i - \hat{y}_i)^2}{n} = \frac{SS_E}{n}$$

dengan y_i adalah peubah respon pada lokasi i, $\hat{y}_i$ adalah nilai penduga peubah respon pada lokasi i, n adalah banyak pengamatan, dan SS_E adalah jumlah kuadrat galat.

Penduga untuk $\boldsymbol{\beta}$ adalah:

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

dan penduga untuk ρ adalah:

$$\hat{\rho} = (\mathbf{y}^T\mathbf{W}^T\mathbf{W}\mathbf{y})^{-1}\,\mathbf{y}^T\mathbf{W}^T\mathbf{y}$$

2.4 Model Galat Spasial (SEM)

Jika $\rho = 0$ dan $\lambda \neq 0$, maka persamaan (1) menjadi

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}\,, \quad \mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \quad (13)$$

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

Model galat spasial adalah model regresi linier yang pada peubah galatnya terdapat korelasi spasial. Fungsi log kemungkinan (*log-likelihood*) model SEM diperoleh dari persamaan (9) dengan menggantikan nilai $\rho = 0$ dan akan diperoleh

$$l = L(\boldsymbol{\beta}, \lambda, \sigma^2; \mathbf{y}_1, \ldots, \mathbf{y}_n)$$

$$= \ln\left(\frac{|I-\rho\mathbf{W}|}{(2\pi)^{n/2}\sigma^n}\, \mathrm{e}\mathbf{X}\mathrm{p}\left[-\frac{(\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{I}-\lambda\mathbf{W})^T(\mathbf{I}-\lambda\mathbf{W})(\mathbf{y}-\mathbf{X}\beta)}{2\sigma^2}\right]\right)$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 + \ln|\mathbf{I} - \lambda\mathbf{W}| - \frac{(\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{I}-\lambda\mathbf{W})^T(\mathbf{I}-\lambda\mathbf{W})(\mathbf{y}-\mathbf{X}\beta)}{2\sigma^2} \quad (14)$$

Pendugaan untuk $\sigma^2, \boldsymbol{\beta}$ dan ρ diperoleh dengan memaksimumkan fungsi log kemungkinan (*log-likelihood*) pada persamaan (14) dan diperoleh :

$$\hat{\sigma}^2 = \frac{\left[(\mathbf{I} - \lambda\mathbf{W})(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{I} - \lambda\mathbf{W})(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})\right]}{n}$$

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

Untuk menduga parameter λ diperlukan suatu iterasi numerik untuk mendapatkan penduga untuk λ yang memaksimalkan fungsi log kemungkinan tersebut.

2.5 Matriks Pembobot Spasial

Matriks pembobot spasial pada dasarnya merupakan matriks yang menggambarkan hubungan antar wilayah. Pada penelitian ini matriks pembobot spasial yang digunakan adalah matriks pembobot spasial Queen. Matriks pembobot spasial Queen mendefinisikan w_{ij} =1 untuk wilayah yang bersebelahan atau titik sudutnya bertemu dengan wilayah yang menjadi perhatian sedangkan w_{ij} = 0 untuk wilayah lainnya. Matriks pembobot spasial merupakan matriks simetris dan diagonal utama selalu bernilai nol (Lee & Wong 2001).

3. METODOLOGI PENELITIAN

Data yang digunakan dalam penelitian ini adalah Data dan Informasi Kemiskinan tahun 2008 yang telah dipublikasikan oleh BPS. Data ini adalah data sekunder yang berasal dari data Potensi Desa tahun 2008 yang dilakukan oleh. Provinsi Jawa Timur terdiri atas 29 kabupaten dan 9 kota.

Peubah tak bebas pada penelitian ini adalah *headcount index* kemiskinan di tingkat kabupaten. *Head Count Index* adalah persentase penduduk yang berada di bawah Garis Kemiskinan (GK).

GK merupakan penjumlahan dari GKM dan GKNM. Penduduk yang yang memiliki rata-rata pengeluaran per kapita per bulan di bawah GK dikategorikan penduduk miskin (BPS 2008).GKM adalah jumlah nilai pengeluaran dari 52 komoditi dasar makanan yang riil dikonsumsi penduduk referensi yang kemudian disetarakan dengan 2100 kilo kalori perkapita per hari. Penyetaraan nilai pengeluaran kebutuhan minimum makanan dilakukan dengan menghitung harga rata-rata kalori dari ke-52 komoditi tersebut. GKNM adalah penjumlahan nilai kebutuhan minimum dari komoditi-komoditi non makanan terpilih yang meliputi perumahan, sandang, pendidikan, dan kesehatan.

Peubah-peubah penjelas yang diduga mempengaruhi Y diperoleh dari kriteria kemiskinan menurut informasi kemiskinan BPS, yaitu :

1. Pendidikan, yang terdiri dari
   (a) X_1 adalah persentase penduduk yang tidak dapat membaca pada usia 15-55 tahun.
   (b) X_2 adalah persentase penduduk yang tidak tamat Sekolah Dasar.
2. Fasilitas Perumahan, yang terdiri dari
   (a) X_3 adalah persentase rumah tangga yang menggunakan air minum yang tidak berasal dari air mineral, air PAM, pompa air, sumur atau mata air terlindung.
   (b) X_4 persentase penduduk yang menempati rumah sehat. Menurut Departemen Kesehatan, sebuah rumah dikategorikan sebagai rumah sehat apabila luas lantai per kapita yang ditempati minimal 8 m^2.
3. Ketenagakerjaan
   (a) X_5 adalah persentase penduduk yang bekerja di sektor pertanian.
   (b) X_6 adalah persentase penduduk yang bekerja pada sektor non pertanian.
   (c) X_7 adalah persentase penduduk yang bekerja di sektor formal.
   (d) X_8 adalah persentase penduduk yang bekerja di sektor informal

Langkah-langkah analisis yang dilakukan adalah:

1. Pemilihan peubah-peubah penjelas dari data Podes 2008.
2. Menguji efek spasial yaitu uji dependensi spasial dan uji heterogenitas spasial. Uji dependensi spasial dilakukan dengan metode uji Pengganda Lagrange (*Lagrange Multiplier*, LM) dan uji keragaman spasial dilakukan dengan uji Breusch-Pagan (Anselin 1988).
3. Menentukan matriks pembobot spasial **W**.
4. Menduga parameter untuk persamaan model regresi spasial dengan metode penduga kemungkinan maksimum.
5. Menguji asumsi model regresi spasial.
6. Menentukan model yang paling sesuai dengan kriteria nilai koefisien determinasi (R^2), dan nilai akar rata-rata kuadrat tengah galat (*Root Mean Square Error*, RMSE) terkecil. Kemudian menentukan faktor-faktor yang mempengaruhi kemiskinan.

4. HASIL DAN PEMBAHASAN

4.1 Pemilihan Peubah Penjelas

Peubah-peubah penjelas dipilih yang secara nyata tidak saling berkorelasi, untuk memenuhi asumsi regresi bahwa tidak ada autokorelasi antar peubah penjelas. Peubah-peubah penjelas yang akhirnya digunakan adalah X_2 (persentase penduduk yang tidak tamat Sekolah Dasar), X_3 (persentase rumah tangga yang menggunakan air minum yang tidak layak), X_4 (persentase penduduk yang menempati rumah sehat).

Nilai korelasi Pearson dan nilai-p untuk setiap korelasi antara peubah respon dengan peubah-peubah penjelas yang terpilih serta antar peubah penjelas terdapat pada Tabel 1. Nilai-nilai korelasi antar peubah penjelas yang kecil dan tidak nyata pada taraf 5%, dapat menjadi indikator tidak adanya autokorelasi dalam model regresi yang terbentuk.

Tabel 1
Korelasi Pearson antara peubah respon dengan peubah-peubah penjelas dan antar peubah penjelas

Peubah	korelasi Pearson	nilai-p
Y dengan X_2	0,730	0,000 **
Y dengan X_3	0,800	0,000 **
Y dengan X_4	0,642	0,000 **
X_2 dengan X_3	-0,210	0,205
X_2 dengan X_4	0,415	0,067
X_3 dengan X_4	-0,194	0,235

Keterangan : **: nyata pada α=1%,

4.2 Pengujian Efek Spasial

Pengujian ini bertujuan untuk mengetahui adanya heterogenitas spasial dan ketergantungan spasial. Kedua hal di atas dilakukan untuk menentukan model spasial yang akan digunakan untuk memodelkan persentase kemiskinan. Lagrange Multiplier (LM) digunakan untuk mendeteksi otokorelasi spasial secara pada peubah respon, peubah galat, atau keduanya. Sedangkan untuk menguji heterogenitas spasial digunakan dengan uji Breusch-Pagan. Hasil uji ketergantungan dapat dilihat pada Tabel 2.

Model SAR mempunyai nilai LM tertinggi dan nyata dengan nilai-p = 0.0002. Sedangkan untuk model GSM dan SEM, hasil uji LM menunjukkan tidak nyata. Dengan demikian model spatial yang dapat menggambarkan data kemiskinan adalah model SAR.

Tabel 2.
Hasil Uji Otokorelasi Spasial dengan LM

Model	Nilai LM	Khi-kuadrat	nilai-p	Kesimpulan
GSM	0.4773	5.99	0.4897	Terima Ho
SAR	13.278	3.84	0.0002	Tolak Ho
SEM	1.4002	3.84	0.2367	Terima Ho

Pengujian efek spasial selanjutnya adalah uji heterogenitas spasial yaitu dengan uji Breusch-Pagan. Nilai statistik Breusch-Pagan sebesar 9.677025 dengan nilai khi kuadrat pada derajat bebas 3 sebesar 7.81 dan nilai-p sebesar 0.0079, maka tolak H_0 yang berarti terdapat keragaman antar wilayah. Hasil kedua uji di atas (ketergantungan dan heterogenitas spasial) mengindikasikan terdapat efek spasial dalam data sehingga model regresi yang digunakan sebaiknya memasukkan pengaruh lokasi ke dalam model.

4.3 Model SAR

Koefisien model OLS dan SAR tertera pada Tabel 2. Koefisien determinasi (R^2) pada model SAR lebih tinggi dibandingkan model OLS, demikian juga nilai RMSE SAR lebih kecil dari model OLS. Dari Tabel 2 terlihat bahwa faktor-faktor yang berpengaruh pada peningkatan persentase kemiskinan adalah persentase penduduk yang tidak tamat Sekolah Dasar (SD) atau tidak bersekolah dan persentase penduduk yang menggunakan air minum yang tidak berasal dari air mineral, air PAM, pompa air, sumur atau mata air terlindung serta persentase rumah tangga yang tidak menempati rumah dengan kategori sehat yaitu dengan luas lantai lebih dari 8 m^2.

Model SAR yang terbentuk adalah

$$\hat{y} = 0.11 + 0.3Wy + 1.38\ X_2 + 0.15\ X_3 + 0.72\ X_4$$

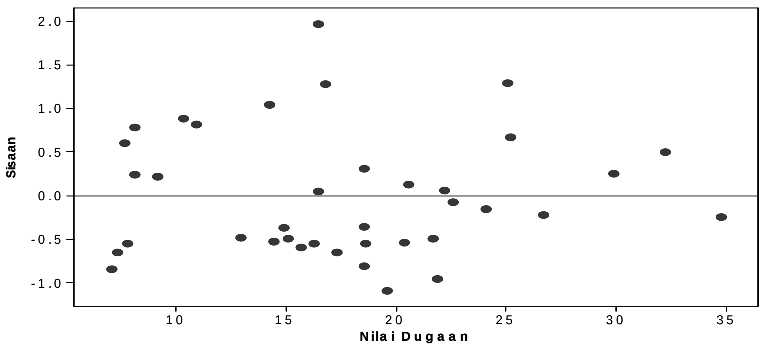
Nilai dugaan parameter ($\hat{\boldsymbol{\beta}}$) ketiga peubah penjelas yaitu X_2, X_3, dan X_4 bernilai positif berarti semakin meningkatnya persentase penduduk yang tidak tamat Sekolah Dasar (SD) dan persentase rumah tangga yang menggunakan air minum yang tidak berasal dari air mineral, air PAM, pompa air, sumur atau mata air terlindung serta meningkatnya rumah tangga yang tidak menempati rumah dengan kategori sehat di suatu kabupaten/kota mengindikasikan meningkatnya persentase kemiskinan di kabupaten/kota tersebut.

Tabel 3.
Koefisien Pada Model OLS dan SAR

	OLS	SAR
Intercept	1.13*	0.11*
X_2	0.47*	1.38*
X_3	0.18*	0.15*
X_4	0.7*	0.72*
R^2	0.8230	0.9989
Rho		0.30*
R^2	0.8230	0.9989
RMSE	0.7814	0.5867

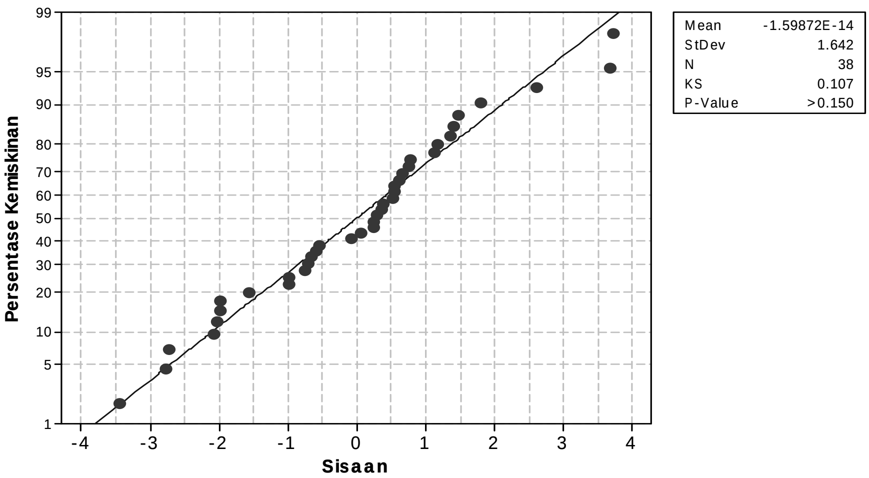
*) nyata pada α = 5%

Pengujian asumsi pada model SAR adalah uji homoskedastisitas, kenormalan, kebebasan dari galat model. Hasil pengujian menunjukkan bahwa ketiga asumsi tersebut terpenuhi. Uji asumsi kehomogenandapat dilihat dari plot sisaan pada Gambar 1. Dari plot pada Gambar 1 terlihat sebaran sisaan menyebar tidak membentuk pola sehingga asumsi kehomogenan terpenuhi.



Gambar 1
Plot Kehomogenan Sisaan pada Model SAR

Uji kenormalan dari sisaan digunakan metode Kolmogorov-Smirnov (KS). Hasil pengolahan diperoleh nilai KS adalah 0.097 dengan nilai p-value lebih dari 0.15 (< 0.15), ini menunjukkan sisaan berdistribusi normal. Uji kenormalan dapat dilihat pada Gambar 2.



Gambar 2.
Uji Kenormalan pada Model SAR

Uji asumsi kebebasan galat dilakukan dengan uji Durbin Watson. Hasil pengolahan diperoleh nilai Durbin Watson sebesar 1.68. Pada k=3, α = 5%, n=38, d_L = 1.32, d_U = 1.66, karena d > d_U yaitu 1.68 > 1.66 maka d tidak nyata yang berarti tidak tolak H_0 sehingga dapat disimpulkan asumsi tidak ada otokorelasi pada sisaan terpenuhi.

5. SIMPULAN

Model regresi SAR lebih baik dibandingkan model klasik OLS dalam penentuan faktor-faktor yang berpengaruh terhadap kemiskinan di Provinsi Jawa Timur. Faktor-faktor yang berpengaruh pada persentase kemiskinan berdasarkan model SAR adalah persentase penduduk yang tidak tamat Sekolah Dasar (SD) atau tidak bersekolah, persentase penduduk yang menggunakan air minum yang tidak berasal dari air mineral, air PAM, pompa air, sumur atau mata air terlindung, dan persentase penduduk yang menempati rumah dengan kategori sehat yaitu dengan luas lantai lebih dari 8 m^2 berbanding terbalik dengan persentase kemiskinan.

DAFTAR PUSTAKA

[1]. Anselin L., (1988), Spatial Econometrics: Methods and Models, Academic Publishers , Dordrecht.
[2]. [BPS] Badan Pusat Statistik, 2008, Data dan Informasi Kemiskinan, Badan Pusat Statistik, Jakarta
[3]. Hakim L. and Zuber A, 2008, *Dimensi Geografis dan Pengentasan Kemiskinan Pedesaan. Media Ekonom*i, Fakultas Ekonomi Universitas Trisakti, Jakarta
[4]. Fotheringham A.S., Brunsdon C., Chartlon M., 2002, Geographically Weighted Regression, the analysis of spatially varying relationships, John Wiley and Sons, LTD.
[5]. Lee J.and Wong D.W.S. 2001, *Statistic for Spatial Data*, John Wiley & Sons, Inc, New York
[6]. Le Sage J.P. ,1999, *The Theory and Practice of Spatial Econometrics*, Department of Economics University of Toledo, Toledo.