

PEMODELAN *SPATIAL AUTOREGRESSIVE* (SAR) UNTUK PRESENTASE PENDUDUK MISKIN DI JAWA BARAT TAHUN 2018

Dea Handayani Juniar¹, Muhammad Ulinnuha²

^{1,2} Mahasiswa Pasca Sarjana Statistika Terapan FMIPA IPB University

¹deajuniar@apps.ipb.ac.id ²muhammadulinnuhaulinnuha@apps.ipb.ac.id

Abstrak: Badan Pusat Statistik (BPS) memandang kemiskinan sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Penanggulangan kemiskinan merupakan salah satu upaya untuk meningkatkan kualitas hidup masyarakat sehingga diperlukan analisis mendalam untuk penanggulangan kemiskinan. Analisis penanggulangan kemiskinan akan lebih efektif dengan pendekatan geografis. Penelitian ini bertujuan untuk memodelkan persentase jumlah penduduk miskin di Jawa Barat dengan memperhatikan efek ketergantungan antar wilayah. Pada penelitian ini peubah bebas yang digunakan adalah Laju PDRB, Indeks Pendidikan dan Indeks Kesehatan. Metode yang digunakan model regresi spasial *Spatial Autoregressive Model* (SAR) dan *Spatial Error Model* (SEM). Hasil penelitian menunjukkan Model SAR lebih baik dibandingkan model SEM dan OLS karena pada model SAR memiliki AIC lebih kecil dan nilai Pseudo R^2_{adj} lebih tinggi dibandingkan model SEM dan OLS. Peubah bebas yang berpengaruh signifikan terhadap presentase penduduk miskin di Jawa Barat pada model SAR adalah indeks pendidikan. Pemodelan dengan *Spatial Autoregressive Model* menghasilkan nilai AIC sebesar 116.99 dengan R^2_{adj} sebesar 56.4%.

Kata Kunci: *Ordinary Least Square* (OLS), *Regresi Spasial*, *Spatial Autoregressive* (SAR) dan *Spatial Error Model* (SEM).

1. Pendahuluan

1.1 Latar Belakang

Analisis regresi spasial merupakan salah satu pengembangan dari analisis regresi klasik. Analisis regresi klasik atau sering disebut *Ordinary Least Square* (OLS) merupakan metode yang sering digunakan untuk melihat hubungan dan memodelkan peubah-peubah bebas dengan suatu peubah respon. Pada analisis regresi OLS ada beberapa asumsi yang harus dipenuhi yaitu asumsi normalitas residual, tidak terdapat heteroskedastisitas & autokorelasi pada residual dan antar peubah bebas tidak terdapat multikolinieritas (Gujarati, 2003). Penggunaan pemodelan analisis regresi linier pada data yang mengandung informasi ruang atau tempat dapat menyebabkan pelanggaran asumsi seperti sisaan menjadi berkorelasi yang bisa disebut adanya efek autokorelasi dan dapat juga menyebabkan sisaan ragam tidak homogen dengan kata lain terjadi heteroskedastik. Data yang memiliki informasi ruang atau tempat apabila diabaikan dalam pemodelan, maka dapat

mengakibatkan koefisien regresi yang diperoleh menjadi bias dan tak konsisten sehingga penarikan kesimpulan dari hasil pemodelan yang diperoleh tidak tepat, untuk mengatasi hal tersebut dapat digunakan pendekatan statistik lainnya seperti regresi spasial. Regresi spasial merupakan suatu metode statistik yang dapat diterapkan untuk memodelkan suatu permasalahan yang dipengaruhi oleh beberapa peubah tertentu dengan memperhatikan adanya ketergantungan antar lokasi atau pengamatan (Anselin, 1988).

Kemiskinan merupakan salah satu permasalahan dalam makro ekonomi. Kemiskinan memberikan dampak negatif ke semua sektor seperti meningkatkan pengangguran dan kriminalitas, menjadi pemicu timbulnya bencana sosial, dan akan menghambat kemajuan suatu daerah. Penanggulangan kemiskinan merupakan salah satu upaya untuk meningkatkan kualitas hidup masyarakat sehingga diperlukan analisis mendalam untuk penanggulangan kemiskinan. Analisis penanggulangan kemiskinan akan lebih efektif dengan pendekatan geografis. Hal ini dikarenakan kemiskinan suatu wilayah bisa dipengaruhi oleh wilayah sekitarnya. Hal ini berkaitan dengan hukum geografi yang dikemukakan Tobler (1979), yang menyatakan segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang lebih dekat akan lebih berpengaruh daripada sesuatu yang lebih jauh. Berdasarkan hukum tersebut, dalam penelitian ini akan dikaji apakah analisis regresi spasial yang mempertimbangkan faktor lokasi lebih tepat digunakan dibandingkan dengan analisis regresi klasik OLS pada pemodelan persentase kemiskinan di Jawa Barat tahun 2018.

1.2 Tinjauan Pustaka

a. Analisis Regresi Spasial

Analisis regresi spasial adalah analisis yang mengevaluasi hubungan antara variabel respon dengan beberapa variabel penjelas dengan memberikan efek spasial pada beberapa lokasi yang menjadi pusat pengamatan. Menurut Anselin (1988), Bentuk umum model regresi spasial adalah sebagai berikut :

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

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

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

dengan \mathbf{y} adalah vektor variabel dependen, \mathbf{X} adalah matriks variabel independen dengan ukuran $n \times (k+1)$, $\boldsymbol{\beta}$ adalah vektor koefisien parameter regresi dengan ukuran $(k+1) \times 1$, \mathbf{W} adalah matriks pembobot spasial dengan ukuran $n \times n$, \mathbf{u} dan $\boldsymbol{\varepsilon}$ adalah vektor error dengan ukuran $n \times 1$, ρ adalah parameter koefisien spasial lag variabel respon, λ adalah parameter koefisien spasial lag error, \mathbf{I} adalah matriks identitas dengan ukuran $n \times n$, n adalah banyaknya amatan atau lokasi dan k adalah jumlah variabel bebas.

Jika pada persamaan model (1), nilai $\rho = 0$ dan $\lambda = 0$, maka model yang terbentuk adalah model regresi linear klasik atau OLS. Secara umum bentuk model regresi linear dalam bentuk matriks (Myers, 1990) adalah sebagai berikut :

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3)$$

Pendugaan parameter $\boldsymbol{\beta}$ menggunakan metode kuadrat terkecil. Penduga $\hat{\boldsymbol{\beta}}$ untuk model ini adalah sebagai berikut :

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{Y} \quad (4)$$

Jika pada persamaan model (1) nilai $\rho \neq 0$ dan $\lambda = 0$, maka model yang terbentuk adalah model *Spatial Autoregressive* (SAR). Model SAR merupakan model ketika nilai observasi dari peubah respon di lokasi (i) memiliki hubungan (korelasi) dengan nilai observasi peubah respon di lokasi sekelilingnya (j). Bentuk umum persamaan model SAR (LeSage, 1999) adalah sebagai berikut :

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (5)$$

Bentuk penaksir parameter untuk model SAR adalah sebagai berikut :

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t(\mathbf{I} - \rho\mathbf{W})\mathbf{y} \quad (6)$$

Jika pada persamaan model (1) nilai $\rho = 0$ dan $\lambda \neq 0$, maka model yang terbentuk adalah model *Spatial Error Model* (SEM). Model spasial error muncul ketika nilai error pada suatu lokasi (i) memiliki korelasi dengan nilai error di lokasi sekitarnya (j) atau dengan kata lain terdapat korelasi spasial antar error. Bentuk umum persamaan model SEM (LeSage, 1999) adalah sebagai berikut :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\varepsilon} \quad (7)$$

Bentuk penaksir parameter untuk model SEM adalah sebagai berikut :

$$\hat{\boldsymbol{\beta}} = [(\mathbf{X} - \lambda\mathbf{W}\mathbf{X})^t(\mathbf{X} - \lambda\mathbf{W}\mathbf{X})]^{-1}(\mathbf{X} - \lambda\mathbf{W}\mathbf{X})^t(\mathbf{I} - \lambda\mathbf{W})\mathbf{y} \quad (8)$$

Pengujian signifikasni parameter pada analisis regresi spasial menggunakan uji wald (Anselin, 1988). Hipoteisis dari uji wald, yaitu

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

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

Statistik uji wald :

$$W = \left[\frac{\hat{\beta}_j^2}{\text{var}(\hat{\beta}_j)} \right] \sim \chi_{\alpha,1}^2 \quad (9)$$

Kriteria dalam pengambilan kesimpulan adalah tolak H_0 , jika $W > \chi_{\alpha,1}^2$

b. Matriks Pembobot Spasial

Matriks pembobot spasial adalah matriks yang berdimensi $n \times n$ yang menyatakan hubungan antara observasi spasial dependency (Dubin,2009). Matriks pembobot spasial dapat dibedakan menjadi 2 yaitu matriks berdasarkan jarak dan matriks berdasarkan persinggungan (*contiguity*). Salah satu matriks berdasarkan persinggungan (*contiguity*) adalah *Queen Contiguity*. Matriks *Queen Contiguity*, dimana elemen matriks bernilai satu ($W_{ij} = 1$) jika bersisian atau bersinggungan dengan wilayah yang menjadi perhatian dan bernilai nol ($W_{ij} = 0$) untuk wilayah yang tidak bersisian atau bersinggungan.

c. Efek Spasial

Efek spasial terjadi akibat adanya dependensi dalam data wilayah. Efek Spasial dependence muncul berdasarkan hukum Tobler I (1979). Uji untuk mengetahui efek spasial pada suatu model adalah statistik Moran's I dan Langrange Multiplier (LM). Hipotesis untuk uji Moran's I adalah sebagai berikut :

$H_0 : I = 0$ (tidak ada autokorelasi antar wilayah)

$H_1 : I \neq 0$ (ada autokorelasi antar wilayah)

Statistik Uji Moran's I (Anselin, 1988) yaitu:

$$Z(I) = \frac{I - E(I)}{\sqrt{\text{var}(I)}} \sim N(0,1) \quad (10)$$

dimana :

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}} \left[\frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \right], E(I) = I_0 = -\frac{1}{n-1}$$

Nilai statistik uji dimana $Z(I)$ mengikuti sebaran normal, yang artinya tolak H_0 jika $|Z(I)| > Z_{\alpha/2}$.

Hipotesis untuk Uji LM model *spatial autoregressive* adalah sebagai berikut :

$H_0: \rho = 0$ (Tidak ada ketergantungan galat spasial)

$H_1: \rho \neq 0$ (Ada ketergantungan galat spasial)

Hipotesis untuk Uji LM model *spatial error model* adalah sebagai berikut:

$H_0: \lambda = 0$ (Tidak ada ketergantungan galat spasial)

$H_1: \lambda \neq 0$ (Ada ketergantungan galat spasial)

Statistik Uji LM (Anselin, 1988):

$$LM_{\rho} = \left[\frac{\boldsymbol{\varepsilon}^t \mathbf{W} \mathbf{y}}{\boldsymbol{\varepsilon}^t \boldsymbol{\varepsilon} / N} \right]^2 / D \quad (11)$$

$$D = \left[(\mathbf{W} \mathbf{X} \hat{\boldsymbol{\beta}})^t [\mathbf{I} - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t] (\mathbf{W} \mathbf{X} \hat{\boldsymbol{\beta}}) / \hat{\sigma}^2 \right] + \text{tr}(\mathbf{W}^t \mathbf{W} + \mathbf{W} \mathbf{W})$$

$$LM_{\lambda} = \frac{\left[\frac{\boldsymbol{\varepsilon}^t \mathbf{W} \boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}^t \boldsymbol{\varepsilon} / N} \right]^2}{\text{tr}(\mathbf{W}^t \mathbf{W} + \mathbf{W} \mathbf{W})} \quad (12)$$

Kriteria keputusan tolak H_0 dilakukan jika nilai statistik uji LM lebih besar dari χ_q^2 dengan q adalah banyaknya parameter spasial (1).

2. Metode Penelitian

2.1 Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari beberapa sumber BPS Provinsi Jawa Barat tahun 2018. Pada penelitian ini menggunakan variabel respon yaitu persentase penduduk yang berada dibawah garis kemiskinan di Jawa Barat pada tahun 2018, sedangkan variabel bebas terdiri dari 3 peubah yaitu Laju PDRB (X_1), Indeks Pendidikan (X_2) dan Indeks Kesehatan (X_3).

2.2 Tahapan Analisis Data

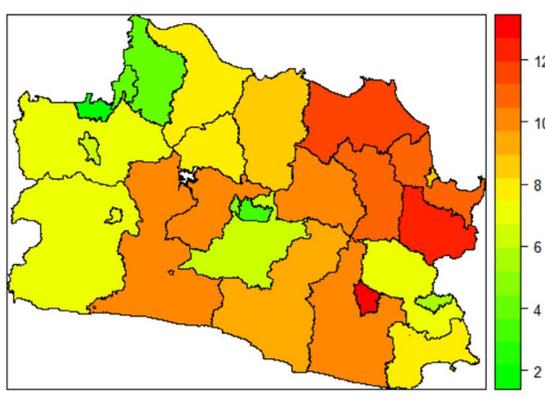
Langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut :

- 1) Melakukan eksplorasi terhadap data variabel respon dengan menggunakan peta tematik.
- 2) Melakukan pemodelan analisis regresi dengan *Ordinary Least*.

- 3) Membuat matriks pembobot spasial dengan metode *Queen Contiguity* (Persinggungan sisi-sudut).
- 4) Melakukan uji Moran's I untuk mengukur autokorelasi antar wilayah.
- 5) Melakukan Uji *Langrange Multiplier* (LM) untuk mengetahui model yang cocok dengan data.
- 6) Melakukan estimasi parameter analisis regresi spasial yang signifikan pada tahap (5)
- 7) Melakukan pemeriksaan asumsi pada model spasial (kehomogenan, kenormalan, dan saling bebas) yang terpilih.
- 8) Membandingkan hasil analisis regresi spasial dengan analisis regresi *ordinary least square* dan memilih model terbaik.

3. Hasil dan Pembahasan

Provinsi Jawa Barat terdiri dari 27 kabupaten/kota, meliputi 18 kabupaten dan 9 kota. Penyebaran persentase penduduk miskin di Jawa Barat tahun 2019 pada setiap Kabupaten/Kota disajikan pada peta tematik pada Gambar 1. Berdasarkan gambar tersebut terlihat adanya kecenderungan spasial dependensi atau ketergantungan spasial yang dilihat dari plot berwarna yang mengumpul. Hal ini mengindikasikan adanya ketergantungan spasial. Gradasi warna menunjukkan bahwa warna semakin hijau memiliki persentase penduduk miskin rendah dan jika semakin merah memiliki persentase penduduk miskin semakin tinggi.



Gambar 1. Peta Tematik Persentase Penduduk Miskin Jawa Barat Tahun 2018

Tahap selanjutnya melakukan pemodelan tanpa mempertimbangkan efek spatial dependensi dengan menggunakan analisis regresi linear berganda. Berdasarkan metode kuadrat terkecil diperoleh hasil pada Tabel 1.

Tabel 1. Hasil Regresi Linear Berganda

Parameter	Estimasi	P-Value	VIF
Konstanta	34.7729	0.0553	
X ₁	-0.2219	0.6217	1.3861
X ₂	-0.2294	0.0250	2.6551
X ₃	-0.1392	0.5989	2.1850
R ² Adj			0.4408
AIC			120.8187

Berdasarkan Tabel 1 peubah Laju PDRB (X_1), Indeks Pendidikan (X_2) dan Indeks Kesehatan (X_3) memiliki nilai VIF kurang dari 10. Hal ini berarti menyatakan bahwa peubah bebas sudah bebas dari multikolinearitas. Pada uji parsial Nilai p-value kurang dari taraf signifikansi 5% yang artinya peubah bebas signifikan terhadap peubah respon. Pada tabel 1 menunjukkan bahwa hanya peubah indeks pendidikan (X_2) yang signifikan.

Setelah melakukan pemodelan menggunakan analisis regresi linear berganda, selanjutnya melakukan pengujian efek spatial dependensi menggunakan Moran's I dan uji Lagrange Multiplier.

Tabel 2. Uji Autokorelasi dengan Moran's I

Variabel	Indeks Moran	P-Value
Y	0.4351	0.0007
X ₁	0.1302	0.2266
X ₂	0.2523	0.0370
X ₃	0.3242	0.0093
Error	0.1397	0.0782

Pada Tabel 2 menunjukkan jika nilai p-value kurang dari 5 %, keputusan yang diambil adalah tolak H₀ yang mengidentifikasi peubah Y, X₂, dan X₃ signifikan. Hal ini menyatakan bahwa peubah persentase penduduk miskin (Y), Indeks Pendidikan (X_2) dan Indeks Kesehatan (X_3) memiliki autokorelasi antar

wilayah kabupaten/kota. Selanjutnya melakukan uji LM untuk menentukan model regresi spasial dependensi yang cocok dengan data.

Tabel 3. Hasil Uji LM dengan *Queen Contiguity*

Pengujian	Nilai	P-Value
SAR	4.6290	0.0314
Robust SAR	9.0170	0.0027
SEM	0.8499	0.3566
Robust SEM	5.2379	0.0221

Pengujian efek dependensi dengan uji *Lagrange multiplier (LM)* bertujuan untuk mendeteksi ketergantungan spasial secara spesifik, yaitu ketergantungan dalam lag, sisaan atau keduanya (lag dan sisaan). Ringkasan Hasil uji penganda *Lagrange multiplier (LM)* dapat dilihat pada tabel 3. Berdasarkan p-value hasil pada output yang diperoleh dari uji penganda *Lagrange multiplier (LM)* menunjukkan bahwa terdapat ketergantungan dalam lag (SAR), Robust SAR, dan robust SEM karena mempunyai nilai p-value < 0.05.

Berdasarkan pengujian Moran's I dan uji LM menunjukkan bahwa terdapat efek spasial dependensi pada data persentase penduduk miskin di Jawa Barat. Hal ini mengidentifikasi bahwa pemodelan dengan metode kuadrat terkecil belum cukup akurat karena mengabaikan unsur spasial dalam data. Oleh karena itu, pemodelan akan diselesaikan menggunakan regresi spasial. Tabel 4 menunjukkan hasil estimasi model SAR dan SEM.

Tabel 4. Hasil Estimasi SAR dan SEM

Parameter	SAR		SEM	
	Koef	P-Value	Koef	P-Value
Rho	0.519	0.016		
Lamda			0.387	0.058
Koefisien	20.351	0.155	26.293	0.109
X ₁	-0.136	0.701	-0.184	0.632
X ₂	-0.185	0.015	-0.198	0.019
X ₃	-0.053	0.802	-0.059	0.807
AIC	116.990		121.300	
Pseudo R ²	0.564		0.469	

Berdasarkan Tabel 4 pada model SAR dan SEM dapat disimpulkan bahwa hanya parameter rho (efek spatial lag) dan peubah bebas X_2 yaitu Indeks Pendidikan yang signifikan ($p\text{-value} < 5\%$). Oleh karena itu, dapat disimpulkan bahwa model regresi spatial yang cocok adalah model *spatial autoregressive*. Setelah melakukan pemodelan, langkah selanjutnya adalah melakukan analisis uji asumsi normalitas, autokorelasi, dan heteroskedastisitas pada data error pada model SAR.

Berdasarkan uji *Anderson-Darling normality test* diperoleh nilai p-value sebesar 0.1021 yang menunjukkan bahwa residual berdistribusi normal. Selanjutnya pengujian Breusch-Pagan test diperoleh nilai sebesar 0.9405 yang menunjukkan bahwa residual bersifat homoskedastisitas dan pengujian moran test memperoleh nilai p-value sebesar 0.3731 yang menunjukkan tidak terdapat autokorelasi pada data error model SAR. Oleh karena itu, dapat disimpulkan bahwa model SAR telah memenuhi semua asumsi klasik. Jika dibandingkan dengan hasil analisis regresi linear pada Tabel 1, dapat disimpulkan bahwa model SAR memperoleh model yang lebih baik. Hal ini dapat ditunjukkan dengan nilai AIC SAR (116.99) yang lebih kecil dari AIC OLS (120.818) dan nilai R^2 adj yang lebih besar ($56.4\% > 44.08\%$). Berdasarkan Tabel 4, diperoleh pengujian signifikansi parameter (uji wald) pada model SAR hanya parameter ρ dan X_2 yang signifikan, maka model yang terbentuk adalah

$$\hat{y}_i = 20.315 + 0.519 \sum_{j=1, i \neq j}^m w_{ij} y_j - 0.1852 X_{2i} + \varepsilon_i \quad (13)$$

Koefisien ρ sebesar 0.519 pada model SAR menyatakan bahwa persentase penduduk miskin disuatu wilayah dipengaruhi oleh wilayah lain yang bersinggungan sebesar 0.519 kali. Sedangkan untuk pengaruh peubah bebas X_2 harus memperhatikan marginal efek dari peubah bebas terhadap variabel respon.

Tabel 5. Marginal Efek SAR

Variabel	Langsung	Tidak Langsung	Total
X1	-0.14841234	-0.13526822	-0.2836806
X2	-0.20149349	-0.18364824	-0.3851417
X3	-0.05717489	-0.05211121	-0.1092861

Berdasarkan pengujian signifikansi parameter hanya peubah indeks pendidikan (X_2) yang signifikan. Maka interpretasi pengaruh langsung untuk

variabel Indeks Pendidikan adalah jika variabel Indeks Pendidikan di suatu lokasi bertambah satu satuan, maka nilai Persentase Penduduk Miskin pada lokasi tersebut akan menurun sebesar 0.20149349 satuan. Sedangkan pengaruh tidak langsung yaitu, jika variabel Indeks Pendidikan di suatu lokasi bertambah satu satuan, maka nilai Persentase Penduduk Misikin pada tetangga lokasi tersebut akan menurun sebesar 0.18364824 satuan.

4. Kesimpulan

Model regresi SAR lebih baik dibandingkan model OLS dan SEM dalam pemodelan persentase jumlah penduduk miskin di Provinsi Jawa Barat tahun 2018. Hal ini dilihat dari nilai Psedo R^2 yang lebih besar dan nilai AIC yang lebih kecil. Pola ketergantungan spasial pada persentase jumlah penduduk miskin di Provinsi Jawa Barat tahun 2018 adalah model SAR dengan nilai rho sebesar 0.519 dengan peubah bebas yang berpengaruh secara signifikan adalah indeks pendidikan (X_2) sehingga model yang terbentuk adalah sebagai berikut:

$$\hat{y}_i = 20.315 + 0.519 \sum_{j=1, i \neq j}^m w_{ij} y_j - 0.1852 X_{1i} + \varepsilon_i$$

Daftar Pustaka

- Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers. The Netherlands
- Dubin R. 2009. *Spatial Weight*. Fotheringham AS, PA Rogerson, editor, *Handbook of Spatial Analysis*. London: Sage Publications.
- Gujarati, D.N.2003. *Basic Econometrics*. Fourth Edition. McGraw-Hill. New York
- Le Sage JP. 1999. *The Theory and Practice of Spatial Econometrics*. Toledo: Department of Economics University of Toledo.
- Myers, R.H., 1990, *Classical and modern regression with application, second edition*, Boston: PWS-KENT Publishing Company.