



## ***STRUCTURAL EQUATION MODELLING (SEM) DENGAN MODEL STRUKTURAL REGRESI SPASIAL***

**Tisti Ilda Prihandini<sup>1</sup>, Sony Sunaryo<sup>2</sup>**

<sup>1</sup>Mahasiswa Magister Jurusan Statistika ITS

<sup>2</sup>Dosen Jurusan Statistika ITS

### **Abstrak**

Suatu variabel yang tidak dapat diukur secara langsung (variabel laten) dan mempunyai hubungan kausalitas bisa diselesaikan dengan *Structural Equation Modelling* (SEM). Pada data spasial, seringkali pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang berdekatan (*neighbouring*). Apabila variabel laten tersebut memiliki aspek spasial maka unsur spasial perlu diikuti sertakan dalam modelnya, sehingga diselesaikan dengan SEM spasial. Jika terdapat variabel laten spasial misalnya seperti *deprivation area* (daerah tertinggal) maka dapat digunakan SEM spasial dengan model strukturalnya regresi spasial. Suatu daerah dikategorikan sebagai *deprivation area* dikarenakan salah satunya oleh sisi geografisnya, sehingga daerah tertinggal merupakan variabel yang memiliki aspek spasial, tetapi tidak dapat diukur secara langsung dan harus diukur melalui indikator-indikator yang dapat menjelaskannya.

**Kata Kunci:** SEM, Model Spasial, Regresi Spasial, *Deprivation area*.

### **1. Pendahuluan**

Dalam suatu penelitian sering ditemukan suatu variabel yang tidak dapat diukur secara langsung yang disebut dengan variabel laten. Menurut Hair (1998) variabel laten adalah suatu konstruk dalam model persamaan struktural yang tidak dapat diukur secara langsung, tetapi dapat direpresentasikan atau ditentukan oleh satu atau lebih variabel indikator. Variabel laten terdiri dari 2 tipe yaitu eksogen dan endogen. Variabel eksogen adalah variabel yang menjadi variabel bebas dalam persamaan regresi dan variabel endogen yaitu variabel yang menjadi variabel tak bebas dalam persamaan regresi. Sedangkan variabel indikator adalah variabel yang dapat diukur melalui observasi, sehingga variabel laten tersebut dapat diukur secara tidak langsung oleh variabel indikator. Analisis statistik yang dapat digunakan untuk mengukur hubungan variabel laten dan variabel indikator adalah SEM (*Structural Equation Modeling*).

Seringkali pada suatu variabel penelitian dipengaruhi oleh aspek kewilayahan (spasial) maka perlu dipertimbangkan aspek spasial pada model. Data spasial merupakan data yang memuat informasi lokasi yang didasarkan pada hukum tobler I

yaitu: *Everything is related to everything else, but near things are more related than distance things*. Segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang dekat lebih mempunyai pengaruh daripada sesuatu yang jauh. Dengan demikian untuk analisis statistik SEM yang mempunyai pengaruh kewilayahan pada variabelnya digunakan SEM spasial. Pada data spasial, seringkali pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang berdekatan (*neighboring*). Sehingga untuk mengatasinya, diperlukan unsur spasial yang dimasukkan dalam model. Menurut Congdon (2008) agar terdapat korelasi diantara variabel laten pada SEM dan juga korelasi secara spasial maka digunakan *spatial SEM*.

Penelitian dalam bidang SEM spasial mulai berkembang, karena dalam penerapannya terdapat variabel yang tidak terukur secara langsung (laten) yang memiliki hubungan kausalitas serta pengaruh secara spasial. Wang dan Wall (2003) memperkenalkan efek random struktur spasial tunggal untuk menghitung korelasi diantara variabelnya yang diterapkan pada kasus kesehatan. Kemudian Hogan dan Tchernis (2004) mengadopsi model faktor konfirmatori dengan korelasi spasial tunggal diantara beberapa indikator social. Selanjutnya Congdon (2008) meneliti mengenai SEM spasial dengan dengan model spasial CAR variabel laten *area deprivation* dan *social fragmentation* untuk data kota London.

## **2. Structural Equation Model (SEM)**

SEM merupakan metode analisis multivariat yang dapat digunakan untuk menggambarkan keterkaitan hubungan linier secara simultan antara variabel pengamatan (indikator) dan variabel yang tidak dapat diukur secara langsung (variabel laten). Variabel laten merupakan variabel tak teramati (*unobserved*) atau tak dapat diukur (*unmeasured*) secara langsung. Terdapat dua tipe variabel laten dalam SEM yaitu endogen dan eksogen. Variabel laten endogen adalah variabel laten yang minimal pernah menjadi variabel tak bebas dalam satu persamaan, meskipun dalam persamaan lain (di dalam model tersebut) menjadi variabel bebas. Variabel laten eksogen adalah variabel laten yang berperan sebagai variabel bebas dalam model. SEM merupakan gabungan dari analisis jalur, analisis faktor konfirmatori dan analisis regresi. Secara garis besar sistem persamaan struktural terdiri dari model struktural (*structural model*) dan model pengukuran (*measurement model*).

Model struktural menggambarkan hubungan antar variabel laten, yang dapat dinyatakan dalam bentuk berikut:

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\zeta} + \boldsymbol{\zeta} \quad (1)$$

dengan

$\boldsymbol{\eta}$  = vektor laten endogen

$\mathbf{B}$  = matriks koefisien variabel laten endogen

$\boldsymbol{\zeta}$  = vektor laten eksogen

$\boldsymbol{\zeta}$  = vektor error pada persamaan struktural

$\boldsymbol{\Gamma}$  = matriks koefisien variabel laten eksogen

Model pengukuran adalah bagian dari suatu model persamaan struktural yang menggambarkan hubungan variabel laten dengan indikator-indikatornya.

$$\mathbf{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (2)$$

$\begin{matrix} p \times 1 & p \times m & m \times 1 & p \times 1 \end{matrix}$

$$\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (3)$$

$\begin{matrix} q \times 1 & q \times n & n \times 1 & q \times 1 \end{matrix}$

Dengan:

$\mathbf{y}$  = vektor variabel indikator pada variabel laten  $\boldsymbol{\eta}$   
 $p \times 1$

$\mathbf{x}$  = vektor variabel indikator pada variabel laten  $\boldsymbol{\xi}$   
 $q \times 1$

$\boldsymbol{\delta}$  = error untuk  $\mathbf{x}$   
 $q \times 1$

$\boldsymbol{\varepsilon}$  = error untuk  $\mathbf{y}$   
 $p \times 1$

$\boldsymbol{\Lambda}_y$  = koefisien relasi  $\mathbf{y}$  pada  $\boldsymbol{\eta}$   
 $q \times n$

$\boldsymbol{\Lambda}_x$  = koefisien relasi  $\mathbf{x}$  pada  $\boldsymbol{\xi}$   
 $q \times n$

### 3. Model Spasial

Bentuk data spasial dalam statistik pertama kali disajikan dalam bentuk peta, kemudian berkembang dalam bentuk model spasial, misalnya Fisher yang menggunakan spasial dependence pada percobaan agrikultur dengan metode *nearest-neighbor* untuk menganalisa percobaan lahan pertanian dengan mempertimbangkan aspek ketergantungan spasial (Cressie, 1991).

Salah satu model Regresi spasial adalah model *Geographically Weighted Regression* (GWR) yang merupakan suatu pengembangan dari model regresi. Setiap

parameter pada GWR dihitung pada setiap titik lokasi, sehingga setiap titik lokasi geografis punya nilai parameter regresi yang berbeda- beda. Model GWR dapat ditulis sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad i = 1, 2, \dots, n \quad (4)$$

dengan

- $y_i$  = Nilai pengamatan variabel respon ke- $i$
- $x_{ik}$  = Nilai pengamatan variabel prediktor  $k$  pada pengamatan ke- $i$
- $\beta_k(u_i, v_i)$  = Realisasi fungsi kontinu  $\beta_k(u, v)$  pada pengamatan ke- $i$
- $(u_i, v_i)$  = Titik koordinat (longitude, latitude) lokasi ke- $i$
- $\varepsilon_i$  = Error yang diasumsikan identik, independen, dan berdistribusi normal dengan mean nol dan varian konstan  $\sigma^2$

Estimasi parameter model dengan metode *Ordinary Least Square* (OLS) dalam bentuk vektor dapat ditulis sebagai berikut:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (5)$$

Dengan

- $\beta$  = Vektor parameter yang ditaksir
- $X$  = Matriks data
- $y$  = Vektor variabel respon yang berukuran  $n \times 1$
- $k$  = Banyaknya variabel bebas ( $k = 1, 2, \dots, p$ )

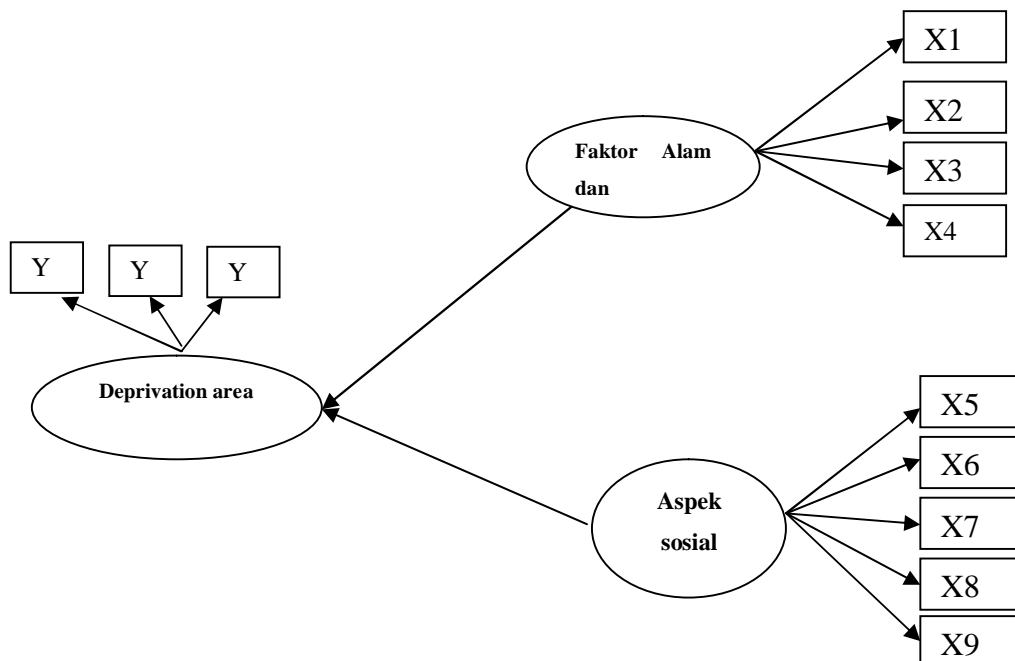
#### 4. Implementasi terhadap *Deprivation Area*

*Deprivation area* atau daerah tertinggal adalah daerah yang relatif kurang berkembang dibandingkan daerah lain dalam skala nasional, dan berpenduduk yang relative tertinggal. Suatu daerah dikategorikan sebagai daerah tertinggal, karena beberapa faktor penyebab, misalnya secara geografis atau keadaan alam, sumber daya alam, sumber daya manusia, Sarana prasarana dan pembangunannya. Variabel yang bisa diambil untuk menggambarkan deprivation area terlihat dalam Tabel 1.

**Tabel 1.** Variabel Laten dan Indikatornya

| <b>Indikator untuk<br/>Faktor Alam dan Lingkungan (<math>\xi</math>)</b>   |   |
|--|---|
| X <sub>1</sub>   | Jumlah keluarga yang tinggal dibantaran/tepi sungai           |
| X <sub>2</sub>   | Jumlah keluarga yang tinggal di permukiman kumuh              |
| X <sub>3</sub>   | Jumlah penderita wabah penyakit selama setahun terakhir       |
| X <sub>4</sub>   | Jumlah korban jiwa akibat bencana alam dalam 3 tahun terakhir |
| <b>Indikator untuk<br/>Aspek Sosial (<math>\xi</math>)</b>                 |   |
| X <sub>5</sub>   | Jumlah tempat ibadah  |
| X <sub>6</sub>   | Banyaknya lembaga atau organisasi kemasyarakatan              |
| X <sub>7</sub>   | Banyaknya penyandang cacat                                    |
| X <sub>8</sub>   | Jumlah sarana transportasi yang tersedia                      |
| X <sub>9</sub>   | Jumlah warnet yang tersedia                                   |
| <b>Indikator untuk<br/>Status Keteringgalan daerah (<math>\eta</math>)</b> |   |
| Y <sub>1</sub>   | keluarga penerima ASKESKIN                                    |
| Y <sub>2</sub>   | surat miskin yang dikeluarkan                                 |
| Y <sub>3</sub>   | penderita gizi buruk dalam 3 tahun terakhir                   |

Jika digambarkan dalam bentuk diagram jalur maka diperoleh gambar sebagai berikut:



### Gambar 1. Diagram Jalur

Dengan menggunakan data yang bersumber dari PODES tahun 2008 untuk provinsi Sumatera Barat dan diolah dengan AMOS, diperoleh hasil untuk model diatas sebagai berikut:

- Bobot faktor (regression Weight) untuk variabel laten aspek alam

**Tabel 2.** Regression Weights: (Group Number 1 – Default Model)

|                    | Estimate | S.E.  | C.R.  | P    | Label |
|--------------------|----------|-------|-------|------|-------|
| X1 <--- Aspek Alam | 1.000    |       |       |      |       |
| X2 <--- Aspek Alam | .343     | .180  | 1.900 | .057 | par_1 |
| X3 <--- Aspek Alam | 3.100    | 1.208 | 2.567 | .010 | par_2 |
| X4 <--- Aspek Alam | 5.214    | 1.737 | 3.002 | .003 | par_3 |

**Tabel 3.** Standardized Regression Weights: (Group Number 1 - Default Model)

|                    | Estimate |
|--------------------|----------|
| X1 <--- Aspek Alam | .703     |
| X2 <--- Aspek Alam | .474     |
| X3 <--- Aspek Alam | .645     |
| X4 <--- Aspek Alam | .956     |

Dari tabel hasil output di atas diketahui bahwa semua faktor merupakan dimensi dari variabel laten (aspek alam) yang dibentuk.

- Bobot faktor (regression Weight) untuk variabel laten aspek sosial

**Tabel 4.** Regression Weights: (Group Number 1 - Default Model)

|                      | Estimate | S.E. | C.R.  | P    | Label |
|----------------------|----------|------|-------|------|-------|
| X5 <--- Aspek_sosial | 1.000    |      |       |      |       |
| X6 <--- Aspek_sosial | .088     | .026 | 3.367 | ***  | par_1 |
| X7 <--- Aspek_sosial | .333     | .077 | 4.316 | ***  | par_2 |
| X8 <--- Aspek_sosial | .036     | .011 | 3.131 | .002 | par_3 |

|                      | Estimate | S.E. | C.R.  | P    | Label |
|----------------------|----------|------|-------|------|-------|
| X9 <--- Aspek_sosial | .008     | .006 | 1.427 | .154 | par_4 |

**Tabel 5.** Standardized Regression Weights: (Group Number 1 - Default Model)

|                      | Estimate |
|----------------------|----------|
| X5 <--- Aspek_sosial | .702     |
| X6 <--- Aspek_sosial | .710     |
| X7 <--- Aspek_sosial | 1.114    |
| X8 <--- Aspek_sosial | .647     |
| X9 <--- Aspek_sosial | .294     |

Dari tabel hasil output di atas diketahui bahwa semua faktor merupakan dimensi dari variabel laten (aspek alam) yang dibentuk

Langkah selanjutnya adalah dengan melakukan skor faktor diperoleh nilai masing-masing variabel laten, kemudian dapat dilakukan regresi spasial sehingga modelnya menjadi:

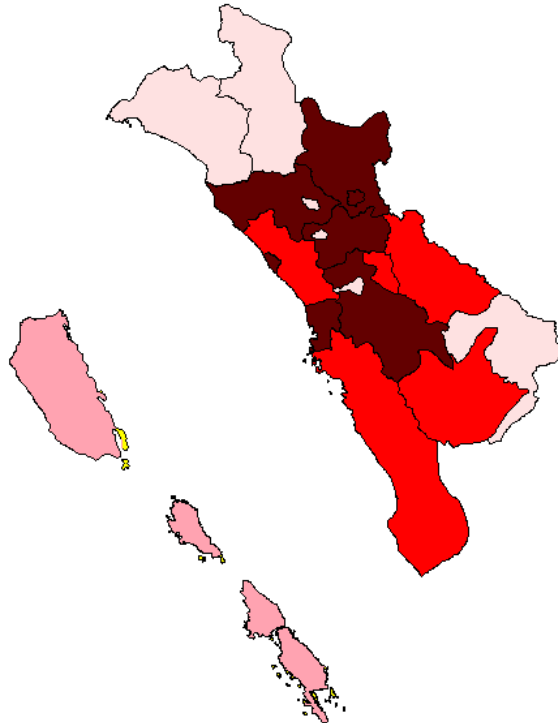
$$\eta_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)\xi_{i1} + \beta_2(u_i, v_i)\xi_{i2} + \beta_4(u_i, v_i)y_{i1} + \beta_5(u_i, v_i)y_{i2} + \beta_6(u_i, v_i)y_{i3} + \varepsilon$$

berikut hasil estimasi model spasial dengan menggunakan GWR4 yang akan dimasukkan ke dalam model umum di atas.

**Tabel 6.** Estimasi model GWR untuk Persamaan Struktural.

| Variable  | Estimate |
|-----------|----------|
| -----     |          |
| Intercept | 0.000000 |
| y1        | 0.020000 |
| y2        | 0.456000 |
| y3        | 0.102000 |
| sosial    | 0.001000 |
| alam      | 0.220000 |

Jika dipetakan dalam bentuk peta penyebaran *deprivation area* dari hasil di atas



dengan software ArcView GIS diperoleh peta seperti Gambar 2.

**Gambar 2.** Peta penyebaran *deprivation area* di Sumatera Barat

Pada Gambar 2 terlihat penyebaran *deprivation area* dengan interpretasi semakin pudar warnanya merupakan gambaran daerah yang semakin tertinggal.

## 5. Kesimpulan

Dari pembahasan di atas diperoleh kesimpulan bahwa faktor yang diduga mempengaruhi kemajuan dan ketertinggalan suatu daerah diantaranya adalah faktor alam dan faktor sosial serta penerima askeskin, surat miskin dan penderita gizi buruk dalam 3 tahun terakhir.

## Daftar Pustaka

Anselin, L., 1988, *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

- Badan Pusat Statistik, World Bank Institute, 2002, *Dasar-dasar Analisis Kemiskinan*, Jakarta.
- Bollen, K. A., 1989, *Structural Equation With Latent Variables*. New York: Wiley.
- Congdon, P., 2008, A Spatial Structural Equation Model for Health Outcomes, *Journal of Statistical Planning and Inference*, Vol 138, pages 2090-2105.
- Congdon, P., 2009, Modelling the Impact of Socioeconomic Structure on Spatial Health Outcomes, *Journal of Statistical Planning and Inference*, Vol 53, pages 3047-3056.
- Cressie, N.A.C., 1991, *Statistics for Spatial Data*, New York : John Wiley & Sons, Inc.
- Hair, J.F., Anderson, R.F ., Tatham, R.L. dan Black, W.C., 1998, *Multivariate Data Analysis*, Fifth Edition, Prentice Hall.
- Hogan, J. dan Tchernis, R., 2004, Bayesian factor analysis for spatially correlated data with application to summarizing area-level material deprivation from census data, *J. Amer. Statist Assoc.* 99,314-324.
- Johnson R.A. and Wichern D.W., 2002, *Applied Multivariate Statistical Analysis*, Fifth Edition, New Jersey: Prentice Hall, In.
- Kosfeld, R., 2010, *Spatial Econometrics*, Entry from [www.ivwl.uni-kessel.de/kosfeldlehre/spatial/spatialEconometrics1](http://www.ivwl.uni-kessel.de/kosfeldlehre/spatial/spatialEconometrics1)
- Sharma, S., 1996, *Applied Multivariate Techniques*, John Wiley & Sons, Inc.
- Wang, F dan Wall, M., 2003, Generalized Common Spatial Factor Model, *Biostatistics*, page:569-582.
- World Bank, 2006, *Era Baru dalam Pengentasan Kemiskinan di Indonesia*: World Bank, Jakarta.