ESTIMASI PARAMETER PADA MODEL REGRESI *SPATIAL LAG* YANG MENGANDUNG *OUTLIER* DENGAN METODE *WEIGHTED Z ALGORITHM*

SKRIPSI

OLEH
ASHIMATIN MUNIFAH
NIM.09610022

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERIMAULANA MALIK IBRAHIM
MALANG
2016

ESTIMASI PARAMETER PADA MODEL REGRESI *SPATIAL LAG* YANG MENGANDUNG *OUTLIER* DENGAN METODE *WEIGHTED Z ALGORITHM*

SKRIPSI

Diajukan kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Sains (S.Si)

Oleh
Ashimatin Munifah
NIM. 09610022

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI MAULANA MALIK IBRAHIM
MALANG
2016

ESTIMASI PARAMETER PADA MODEL REGRESI *SPATIAL LAG* YANG MENGANDUNG *OUTLIER* DENGAN METODE *WEIGHTED Z ALGORITHM*

SKRIPSI

Oleh
ASHIMATIN MUNIFAH
NIM. 09610022

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal 23 Juni 2016

Pembimbing I,	Pembimbing II,
Dr. Sri Harini, M.Si	Ari Kusumastuti, M.Pd, M.Si
NIP. 19731014 200112 2 002	NIP. 19770521 200501 2 004

Mengetahui,
Ketua Jurusan Matematika

Dr. Abdussakir, M.Pd
NIP. 19751006 200312 1 001

ESTIMASI PARAMETER PADA MODEL REGRESI *SPATIAL LAG* YANG MENGANDUNG *OUTLIER* DENGAN METODE *WEIGHTED Z ALGORITHM*

SKRIPSI

Oleh
ASHIMATIN MUNIFAH
NIM. 09610022

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Sains (S.Si)
Tanggal 29 Juni 2016

Penguji Utama	: Fachrur Rozi, M.Si	....
Ketua Penguji	: Dr. H. Imam Sujarwo, M.Pd	....
Sekretaris Penguji	: Dr. Sri Harini, M.Si	....
Anggota Penguji	: Ari Kusumastuti, M.Pd, M.Si	....

Mengesahkan,
Ketua Jurusan Matematika

Dr. Abdussakir, M.Pd
NIP. 19751006 200312 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Ashimatin Munifah
NIM	: 09610022
Jurusan	: Matematika
Fakultas	: Sains dan Teknologi
Judul Skripsi	: Estimasi Parameter pada Model Regresi *Spatial Lag* yang Mengandung *Outlier* dengan Metode *Weighted Z Algorithm*.

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 23 Juni 2016
Yang membuat pernyataan,

Ashimatin Munifah
NIM. 09610022

MOTO

يَٰٓأَيُّهَا ٱلَّذِينَ ءَامَنُوا۟ ٱسْتَعِينُوا۟ بِٱلصَّبْرِ وَٱلصَّلَوٰةِ ۚ إِنَّ ٱللَّهَ مَعَ ٱلصَّٰبِرِينَ ﴿١٥٣﴾

"Hai orang-orang yang beriman jadikanlah sabar dan sholat sebagai penolongmu, sesungguhnya Allah beserta orang-orang yang sabar" (QS. Al-Baqoroh/2:153)

PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:

Ibunda tercinta Ismatul Izzah dan Ayahanda Khuzaini tercinta, yang senantiasa telah membimbing, mendukung, memberi semangat, menolong, selalu mendoakan dan memberi restu kepada penulis dalam menuntut ilmu, serta memberikan teladan yang baik.

Kepada kakak kakak tercinta yang dengan sabar memberikan dorongan dan semangat dalam menuntut ilmu.

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Segala puji bagi Allah Swt. atas rahmat, taufik serta hidayah-Nya, sehingga penulis mampu menyelesaikan penyusunan skripsi ini sebagai slah satu syarat untuk memperoleh gelar sarjana dalam bidang matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Dalam proses penyusunan skripsi ini, penulis banyak mendapat bimbingan dan arahan dari berbagai pihak untuk itu ucapan terima kasih yang sebesar-besarnya dan penghargaan yang setinggi-tingginya penulis sampaikan terutama kepada:

1. Prof. Dr. H. Mudjia Rahardjo, M.Si, selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. drh. Bayyinatul Muchtaromah, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Abdussakir, M.Pd, selaku ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Sri Harini, M.Si, selaku dosen pembimbing I yang telah banyak memberikan arahan, nasihat, motivasi, dan berbagi pengalaman yang berharga kepada penulis.
5. Ari Kusumastuti,M.Pd, M.Si, selaku dosen pembimbing II yang telah banyak memberikan arahan dan berbagi ilmunya kepada penulis.
6. Segenap sivitas akademika Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang terutama seluruh dosen, terima kasih atas segala ilmu dan bimbingannya.

7. Ayah dan Ibu yang selalu memberikan doa, semangat, serta motivasi kepada penulis sampai saat ini.
8. Kakak-kakak yang selalu memberikan doa, semangat serta dukungan kepada penulis sampai saat ini.
9. Kepada teman-teman yang memberikan semangat berjuang bersama kepada penulis sampai saat ini.
10. Semua pihak yang ikut membantu dalam menyelesaikan skripsi ini baik moril maupun materil.

Akhirnya penulis berharap semoga skripsi ini bermanfaat bagi penulis dan bagi pembaca.

Wassalamu'alaikum Wr. Wb.

Malang, Juni 2016

Penulis

DAFTAR ISI

DAFTAR SIMBOL

$y_{(n\times 1)}$	:	Vektor variabel terikat
$X_{(n\times(k+1))}$	:	Matriks variabel bebas
$\beta_{((k+1)\times 1)}$	:	Vektor koefisien parameter regresi
$W_{1(n\times n)}$	:	Matriks pembobot *spatial*
$W_{2(n\times n)}$	:	Matriks bobot *spatialerror*
ρ	:	Parameter koefisien *spatial lag* variabel dependen
λ	:	Parameter koefisien *spatial lagerror*
$u_{(n\times 1)}$	:	Vektor *error* yang diasumsikan mengandung autokorelasi
$\varepsilon_{(n\times 1)}$	:	Vektor *error* yang diasumsikan tidak mengalami autokorelasi, yang berdistribusi normal dengan *mean* nol dan *varians* $I\sigma^2$
$I_{(n\times n)}$	:	Matriks identitas
φ	:	Fungsi objektif
ψ	:	Fungsi *influence* (pengaruh)

ABSTRAK

Munifah, Ashimatin. 2016. **Estimasi Parameter pada Model Regresi *Spatial Lag* yang Mengandung *Outlier* dengan Metode *Weighted Z Algorithm*.** Skripsi. Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Sri Harini, M.Si (II) Ari Kusumastuti, M.Pd, M.Si.

Kata kunci: *Spatial lag*, *outlier, Weighted Z Algorithm*

Pada suatu data spasial, sering memiliki kondisi yang tidak wajar, yaitu adanya *outlier*. *Outlier* adalah suatu keganjilan dan menandakan suatu titik yang sama sekali tidak tipikal dibanding data lainnya. Sedangkan spasial *outlier* adalah suatu titik dimana nilai-nilai atribut non-spasialnya berbeda nyata dari titik-titik yang lain. Adanya *outlier* dapat berdampak terhadap hasil estimasi parameter model yang menyebabkan estimasi parameter menjadi bias. Salah satu pendeteksian *outlier* adalah dengan metode *Weighted Z Algorithm.*

Penelitian ini bertujuan untuk mendapatkan estimasi parameter model *spasial lag* yang mengandung outlier dengan metode Weighted Z Algorithm. Dengan hasil penelitian ini adalah bentuk estimasi parameter regresi *spatial lag* yang mengandung outlier dengan metode *Weighted z algorithm* adalah

$$\hat{\beta}_{OLS} = \left(X^T\left(1 - w_j\right)\omega_i X\right)^{-1} X^T y\,(I - \rho W_1)\left(1 - w_j\right)\omega_i$$

Dengan ω_i adalah matriks pembobot yang berukuran $n \times n$ dengan elemen-elemen diagonal yang berisi pembobot $\omega_1, \omega_2, \omega_3, \dots, \omega_n$. Persamaan tersebut dikenal dengan persamaan *Weighted Least Square* (WLS). Dan terbukti estimator parameter regresi *spatial lag* yang mengandung *outlier* dengan metode *Weighted Z Algorithm* adalah estimator yang bersifat unbias.

ABSTRACT

Munifah, Ashimatin. 2016. **Parameter Estimation on Regression Spatial Lag Model that Containing Outlier Method Weighted Z Algorithm**. Thesis. Departmen of Mathematics, Faculty of Science and Technology, the State Islamic University of Maulana Malik Ibrahim Malang. Supervisor: (I) Dr. Sri Harini, M.Si (II) Ari Kusumastuti, M.Pd, M.Si.

Keywords: *Spatial lag, outlier, Weighted Z Algorithm*

In a spatial data, it has a morbid condition frequently, it is called outlier. Outlier is the anomaly and signifies a point which is not typical at all compared to other data. In other condition, spatial outlier is a point where the value of nonspatial attributes significantly different from the other dots. Outlier can affect the result of parameter estimation model that causes estimate of parameter become biased. One of outlier detection is the method Weighted Z Algorithm.

This study aims to obtain parameter estimation of spatial lag model containing outlier method Z Weighted Algorithm. The results of this research is a form of regression of parameter estimation spatial lag model that containing outlier method weighted z algorithm is:

$$\hat{\beta}_{OLS} = \left(X^T\left(1 - w_j\right)\omega_i X\right)^{-1} X^T y \, (I - \rho W_1)\left(1 - w_j\right)\omega_i$$

With ω_i is a weighting matrix size $n \times n$ with the diagonal elements that contains weighted $\omega_1, \omega_2, \omega_3, \ldots, \omega_n$. This equation is known as the equation Weighted Least Square (WLS). And proven spatial lag regression parameter estimator containing outlier method Weighted Z Algorithm is an estimator that is unbias.

ملخص

منيفة، عصيمة. ٢٠١٦. **.تقدير المعلمة على نموذج الانحدار التأخير المكانية (*Spatial Lag)* يحتوي علشذوذ (*Outlier)* مع طريقة خوارزمية موزونZ (*Weighted Z Algorithm).*** بحث جامعى.شعبة الرياضيات، كلية العلوم والتكنولوجيا، وجامعة الإسلامية الحكومية مولانا مالك إبراهيم مالانج .المشرف(١)الدكتورة سري حريني الماجستيرة، (٢) اري كوسومستوتي الماجستير.

كلمات الرئيسية: تأخر المكانية،شذوذ،خوارزمية موزونZ

في البيانات المكانية، وغالبا ما يكون حالة مرضية، وهذا شذوذ.الشذوذ هو نقطة ليست في جميع نموذجي مقارنة مع غيرها من البيانات. في حين النموذجية المكانية هو نقطة حيث قيم غير المكانية وسمات مختلفة إلى حد كبير من النقاط الأخرى. وجود الشذوذ يمكن أن تؤثر على نتيجة من نموذج تقدير المعلمة التي تسبب تقديرات المعلمة لتكون متحيزة. واحد من كشف الشذوذ هو الأسلوب خوارزمية موزونZ

هدفت هذه الدراسة لتحصل المكاني تأخر تقدير المعلمة نموذج تأخير المكانية يحتوي على شذوذ مع طريقة خوارزمية موزونZ. نتائج هذا البحث هو شكل من أشكال المكاني تقدير المعلمةنموذج الانحدار تأخير المكانية التي تحتوي على شذوذ مع طريقة خوارزمية موزونZ يعنى:

$$\hat{\beta}_{OLS} = \left(X^T(1-w_j)\omega_i X\right)^{-1} X^T y(I-\rho W_1)(1-w_j)\omega_i$$

مع ω_i هو حجم الترجيح مصفوفة $n \times n$ مع العناصر القطري التي تحتوي على ترجيح $\omega_1, \omega_2, \omega_3, \dots, \omega_n$ وتعرف هذه المعادلة كما المعادلة مربعأقلالوزني *Weighted Least Square*(WLS). وثبت مقدر المعلمة الانحدار تأخر المكاني التي تحتوي على شذوذ مع الطريقة خوارزمية موزون Z هو مقدر الذى يصف غير متحيزة

BAB I

PENDAHULUAN

1.1 Latar Belakang

Statistika adalah ilmu yang mempelajari bagaimana merencanakan, mengumpulkan, menganalisis, menginterpretasikan, dan mempresentasikan data (Harini & Turmudi, 2008). Pada statistika juga membahas tentang konsep peluang, atau konsep pendugaan.

Pada statistika, data merupakan hal yang penting. Data adalah inti dari statistika. Salah satu jenis data adalah data *spatial*. Data *spatial* yaitu data yang berorientasi geografis, memiliki sistem koordinat tertentu sebagai dasar referensinya (Hartoyo, dkk, 2010).

Pada suatu data *spatial*, sering memiliki kondisi yang tidak wajar, yaitu adanya *outlier* pada data tersebut. *Outlier* adalah suatu keganjilan dan menandakan suatu titik yang sama sekali tidak tipikal dibandingkan data lainnya (Draper dan Smith, 1992). Sedangkan *spatialoutlier* menurut Shekhar, dkk (2003) adalah suatu titik *spatial* tertentu dimana nilai-nilai atribut *non-spatial*nya berbeda nyata dari titik-titik lain yang masih menjadi tetangganya.

Dalam al-Quran dengan tersirat telah terkait permasalahan tentang *outlier* yaitu terdapat pada suratal-Jin/72:14, yaitu:

وَأَنَّا مِنَّا ٱلۡمُسۡلِمُونَ وَمِنَّا ٱلۡقَٰسِطُونَۖ فَمَنۡ أَسۡلَمَ فَأُوْلَٰٓئِكَ تَحَرَّوۡاْ رَشَدٗا ١٤

"Dan Sesungguhnya di antara kami ada orang-orang *yang taat dan ada (pula) orang-orang yang menyimpang dari kebenaran. barangsiapa yang yang taat, Maka mereka itu benar-benar telah memilih jalan yang lurus"(QS.al-Jin/72:14).*

Asal turunnya Surat al-Jin/72:14 yaitu untuk menampik dugaan bahwa semua jin baik yang mendengar langsung ayat-ayat al-Quran maupun yang belum atau tidak mendengarnya kesemuanya telah patuh kepada Allah. Kemudian pada ayat tersebut diterangkan bahwa *dan sesungguhnya diantara kami* masyarakat jin *ada orang-orang muslim* yakni yang benar-benar taat dan penuh kepatuhan kepada Allah dan ada pula *para penyimpang* yakni mereka yang telah sangat jauh dari kebenaran lagi sangat mantap kekufurannya. *Barang siapa yang patuh, maka mereka itu telah* bersungguh-sungguh *memilih arah* yang mengantar ke jalan *kebenaran* (Shihab, 2003).

Pendeteksian *outlier* sangat penting bagi semua bidang di Sistem Informasi Geografis (SIG) seperti ekologi, transportasi, kesehatan masyarakat, klimatologi, pelayanan umum, dan lain-lain. Metode pendeteksian *outlier* telah banyak dikembangkan oleh ilmuan. Algoritma pendeteksian nilai z (*z algorithm*) dikenalkan oleh Shekhar, dkk (2002). Pada jurnal *Spatial Weighted Outlier Detection* (Kou, dkk, 2006), penelitian membandingkan metode *z value approach, weighted z approach*,dan *average different algorithm* untuk mendeteksi *spatialoutlier*, dan didapatkan kesimpulan bahwa *weighted z approach* dan *average different algorithm* lebih baik dalam mendeteksi *spatialoutlier*.

Berdasarkan latar belakang tersebut, dalam penelitian ini akan membahas tentang "Estimasi Parameter pada Model Regresi *SpatialLag* yang Mengandung *Outlier* denganMetode *Weighted Z Algorithm"*.

1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas pada penelitian ini adalah bagaimana bentuk estimasi parameter pada model regresi *spatial lag* yang mengandung *outlier* dengan metode *Weighted Z Algorithm*?

1.3 Tujuan Penelitian

Tujuan yang akan dicapai pada penelitian ini adalah untuk mengetahui bentuk estimasi parameter model regresi *spatial lag* yang mengandung *outlier* dengan metode *Weighted Z Algorithm.*

1.4 Manfaat Penelitian

Bagi Penulis

1. Mampu mengaplikasikan mata kuliah statistik yang pernah dipelajari di bangku kuliah.
2. Menambah pengetahuan dan wawasan, khususnya keterkaitan dalam estimasi parameter.

Bagi Pembaca

1. Memperkaya dan memperkuat wawasan ilmu Statistika.
2. Membantu pembaca yang ingin memperdalam dan memperluas ilmu pengetahuan khususnya dalam estimasi parameter.
3. Sebagai literatur penunjang khususnya bagi mahasiswa Matematika.

1.5 Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini adalah menggunakan model *spatial lag* untuk model *spatial* dan menggunakan fungsi pembobot *tukey bisquare*.

1.6 Metode Penelitian

1.6.1 Pendekatan Penelitian

Pada penelitian ini menggunakan pendekatan kualitatif untuk mengestimasi permasalahan dengan menggunakan teori yang mendukung dalam masalah yang diangkat. Pendekatan ini menggambarkan objek penelitian yang dihubungkan dan ditelaah dengan teori-teori yang ada.

1.6.2 Sumber Penelitian

Penelitian ini adalah penelitian teoritis yaitu penelitian yang dilakukan dengan cara mengumpulkan data dan informasi yang bersumber dari majalah, artikel, buku-buku, jurnal, dan lain lain.

1.6.3 Tahap-tahap Penelitian

Pada Penelitian ini langkah-langkah untuk estimasi parameter model regresi *spatiallag* yang mengandung *outlier*dengan menggunakan metode *Weighted Z Algorithm* adalah sebagai berikut:

1. Menentukan model regresi *spatial lag* yang mengandung *outlier*.
2. Menghitung *neighborhood functiong*(x_i)
3. Menghitung fungsi pembanding $h(x_i)$.
4. Menghitung nilai jumlah kuadrat *error*.
5. Menghitung nilai $\hat{\beta}$ dan pembobot ω_i.

6. Algoritma dari estimasi parameter dengan Metode *Weighted Z Algorithm.*
7. Kajian *outlier* dalam al-Quran

1.7 Sistematika Penulisan

Untuk mempermudah dan memahami skripsi ini secara keseluruhan maka penulis menggambarkan sistematika pembahasannya yang terdiri dari empat BAB dan masing-masing akan dijelaskan sebagai berikut:

Bab I Pendahuluan

Merupakan Bab Pendahuluan yang menjelaskan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, metode penelitian, dan sistematika penulisan.

Bab II Kajian Pustaka

Dalam Bab ini akan dijelaskan beberapa pengertian dan teori-teori tentang data *spatial*, model *spatial*, pendugaan parameter, matriks pembobot, *outlier, spatialoutlier*, fungsi objektif, regresi *robust*, algoritma pendeteksian *outlier, Z Algorithm, Weighted Z Algorithm,* kajian *outlier*, dan estimasi dalam al-Quran.

Bab III Pembahasan

Merupakan Bab inti dari penulisan yang menjabarkan tentang estimasi parameter pada model regresi *spatial lag* yang mengandung *outlier* dengan metode *Weighted Z Algorithm* dan kajian *outlier* dalam al-Quran.

BabIV Penutup

Merupakan kesimpulan dari pembahasan hasil penelitian yang telah diterangkan dan dilengkapi dengan saran.

BAB II

KAJIAN PUSTAKA

2.1 Data *Spatial*

Data *spatial* yaitu sebuah data yang berorientasi geografis, memiliki sistem koordinat tertentu sebagai dasar referensinya dan mempunyai dua bagian penting yang membuatnya berbeda dari data lain, yaitu informasi lokasi (*spatial*) dan informasi deskriptif (atribut) yang dijelaskan berikut ini:

a. Informasi lokasi (*spatial*), berkaitan dengan suatu koordinat baik koordinat geografis (lintang dan bujur) dan koordinat XYZ, termasuk diantaranya informasi datum dan proyeksi.
b. Informasi deskriptif (atribut) atau informasi *non-spatial*, suatu lokasi memiliki beberapa keterangan yang berkaitan denganya, contohnya: jenis vegetasi, populasi, luasan, kode pos, dan sebagainya (Hartoyo, dkk, 2010).

2.2 Model *Spatial*

Berdasarkan tipe data, pemodelan *spatial* dapat dibedakan menjadi pemodelan dengan pendekatan titik dan area. Jenis pendekatan titik diantaranya *Geographically Weighted Regression* (GWR), *Geographically Weighted Poisson Regression* (GWPR), *Geographically Weighted Logistic Regression* (GWLR), *Space-Time Autoregressive* (STAR), dan *Generalized Space Time Autoregressive* (GSTAR). Menurut LeSage (1998), jenis pendekatan area diantaranya *Mixed Regressive-Autoregressive* atau *Spatial Autoregressive Models* (SAR), *Spatial Error Models* (SEM), *Spatial Durbin Model* (SDM), *Conditional Autoregressive*

Models (CAR), *Spatial Autoregressive Moving Average* (SARMA), dan panel data.

Menurut Anselin (1988) model umum regresi *spatial* dinyatakan pada persamaan berikut:

$$y = \rho W_1 y + X\beta + u \tag{2.1}$$

dengan :

$$u = \lambda W_2 u + \varepsilon \quad ;\ \varepsilon \sim N(0, \sigma^2 \mathrm{I}) \tag{2.2}$$

dimana :

$y_{(n\times 1)}$	:	vektor variabel dependen
$X_{(n\times(k+1))}$	:	matriks variabel independen
$\beta_{((k+1)\times 1)}$	:	vektor koefisien parameter regresi
$W_{1(n\times n)}$	:	matriks pembobot *spatial*
$W_{2(n\times n)}$	:	matriks bobot *spatial error*
ρ	:	parameter koefisien *spatial lag* variabel dependen
λ	:	parameter koefisien *spatial error*
$u_{(n\times 1)}$	:	vektor *error* yang diasumsikan mengandung autokorelasi
$\varepsilon_{(n\times 1)}$	:	vektor *error* yang diasumsikan tidak mengalami autokorelasi, yang berdistribusi normal dengan mean nol dan *varians*$I\sigma^2$
$I_{(n\times n)}$	:	matriks identitas
n	:	banyaknya amatan/lokasi $(i = 1,\ 2, \dots, n)$
k	:	banyaknya variabel independen $(k = 1,\ 2,\ 3, \dots, l)$

Model $u$mempunyai *error* yang berdistribusi normal dengan *mean* nol dan *varians*$\sigma^2 I$. Parameter yang diestimasi adalah β, ρ dan λ. ρadalah parameter koefisien *spatial lag* variabel dependen dan λ adalah parameter *spatial lag* pada *error*. n adalah banyaknya amatan atau lokasi$(i = 1,2,3,\dots,n)$ dan k adalah banyaknya variabel prediktor $(k = 1,2,3,\dots,l)$. Pengaruh *spatial* antar lokasi dalam model dibentuk dalam matriks pembobot W_1, W_2 yang berukuran $n \times n$. Dalam bentuk matriks sebagai berikut:

$$y = [y_1 y_2 \quad \dots \quad y_n]^T \; ; \; u = [u_1 u_2 \quad \dots \quad u_n]^T \; ; \; \varepsilon = [\varepsilon_1 \varepsilon_2 \quad \dots \quad \varepsilon_n]^T$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & x_{ik} & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nl} \end{bmatrix} \; ; \; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \vdots \\ \beta_l \end{bmatrix}$$

$$W_1 \text{atau} W_2 = \begin{bmatrix} W_{11} & W_{12} & W_{13} & \cdots & W_{1n} \\ W_{21} & W_{22} & W_{23} & \cdots & W_{2n} \\ \vdots & \vdots & \vdots & W_{ij} & \vdots \\ W_{n1} & W_{n2} & W_{n3} & \cdots & W_{nn} \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.3}$$

Pada persamaan (2.1) ketika $X = 0$ dan $W_2 = 0$akan menjadi *spatial autoregressive* order pertama seperti pada persamaan berikut:

$$\begin{aligned} y &= \rho W_1 y + \varepsilon \\ \varepsilon &= N(0, \sigma^2 I) \end{aligned} \tag{2.4}$$

Persamaan tersebut menunjukkan *varians* pada y sebagai kombinasi linier *varians* antar lokasi yang berdekatan dengan tanpa variabel independen.

1. Pada persamaan (2.1) jika nilai $W_2 = 0$ atau $\lambda = 0$ maka akan menjadi model regresi *spatialMixed Regressive-Autoregressive* atau *Spatial Autoregressive Model* (SAR) seperti pada persamaan berikut:

$$y = \rho W_1 y + X\beta + \varepsilon$$
$$\varepsilon = N(0, \sigma^2 I) \tag{2.5}$$

Model persamaan (2.5) mengasumsikan bahwa proses *autoregressive* hanya pada variabel dependen.

2. Jika persamaan (2.1) nilai $W_1 = 0$ atau $\rho = 0$ maka akan menjadi model *Spatial Error Model* (SEM) seperti pada persamaan berikut:

$$y = X\beta + \lambda W_2 u + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I) \tag{2.6}$$

$\lambda W_2 u$ menunjukkan *spatial* struktur λW_2 pada *spatially dependent error* (ε).

3. Jika persamaan (2.1) nilai $W_1 W_2 \neq 0$, $\lambda \neq 0$ atau $\rho \neq 0$ akan menjadi model *Spatial Autoregressive Moving Average* (SARMA) Apabila $\rho = 0$ dan $\lambda = 0$, maka persamaan menjadi model regresi linier sederhana yang estimasi parameternya dapat dilakukan melalui *Ordinary Least Square* (OLS) seperti pada persamaan berikut:

$$y = X\beta + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I) \tag{2.7}$$

Hal ini berarti dalam model persamaan tersebut tidak terdapat efek *spatial*.

2.3 Pendugaan Parameter

Pendugaan (*estimasi*) adalah proses yang menggunakan sample statistik untuk menduga atau menaksir hubungan parameter populasi yang tidak diketahui. Pendugaan merupakan suatu suatu pernyataan mengenai parameter populasi yang diketahui berdasarkan populasi dari sampel, dalam hal ini sampel random, yang diambil dari populasi yang bersangkutan. Jadi dengan pendugaan ini, keadaan parameter populasi dapat diketahui (Hasan, 2002). Menurut Yitnosumarto (1990), pendugaan adalah anggota peubah acak dari statistik yang mungkin untuk sebuah parameter (anggota peubah diturunkan). Besaran sebagai hasil penerapan penduga terhadap data dari semua contoh disebut nilai duga (*estimate*). Adapun sifat-sifat dari penduga parameter tersebut adalah:

1. Tak Bias (*Unbias*)

Menurut Yitnosumarto (1990), satu hal yang menjadi tujuan dalam pendugaan adalah pendugaan harus mendekati nilai sebenarnya dari parameter yang diduga tersebut. Misalkan terdapat parameter θ. Jika $\hat{\theta}$ merupakan penduga tak bias (*unbiased estimator*) dari parameter θ, maka:

$$E(\hat{\theta}) = \theta \tag{2.8}$$

2. Efisien

Suatu penduga (dimisalkan: $\hat{\theta}$) dikatakan efisien bagi parameter (θ) apabila penduga tersebut mempunyai *varians* yang kecil. Apabila terdapat lebih dari satu penduga, penduga yang efisien adalah penduga yang mempunyai varian terkecil. Dua buah penduga dibandingkan efisiennya dengan menggunakan efisien relatif (*relative efficiency*). Efisien relatif $\hat{\theta}_2$ terhadap $\hat{\theta}_1$ dirumuskan:

$$R(\hat{\theta}_2, \hat{\theta}_1) = \frac{E(\hat{\theta}_1 - \hat{\theta})^2}{E(\hat{\theta}_2 - \hat{\theta})^2}$$

$$= \frac{E(\hat{\theta}_1 - E(\hat{\theta}_1))^2}{E(\hat{\theta}_2 - E(\hat{\theta}_2))^2}$$

$$= \frac{var\ \hat{\theta}_1}{var \hat{\theta}_2} \quad (2.9)$$

$R = \frac{\hat{\theta}_1}{\hat{\theta}_2}$, jika $R > 1$ maka $\hat{\theta}_1 > \hat{\theta}_2$ artinya secara relatif $\hat{\theta}_2$ lebih efisien daripada $\hat{\theta}_1$, dan jika $R < 1$ maka $\hat{\theta}_1 < \hat{\theta}_2$ artinya secara relatif $\hat{\theta}_1$ lebih efisien daripada $\hat{\theta}_2$.

3. Konsisten

Suatu penduga dikatakan konsisten apabila memenuhi syarat sebagai berikut:

a. Jika ukuran sampel semakin bertambah maka penduga akan mendekati parameternya. Jika besar sampel tak terhingga maka penduga konsisten harus dapat memberi suatu penduga titik yang sempurna terhadap parameternya. Jadi, $(\hat{\theta})$ merupakan penduga konsisten, jika dan hanya jika:

$$E\left(\hat{\theta} - E(\hat{\theta})\right)^2 \to 0 \text{ jika } n \to \infty \quad (2.10)$$

b. Jika ukuran sample bertambah besar maka distribusi sampling penduga akan mengecil menjadi satu garis tegak lurus diatas parameter yang sama dengan probabilitas sama dengan 1 (Hasan, 2002).

2.4 Matriks Pembobot

Salah satu hal yang sangat penting dalam analisis adalah penentuan bobot atau penimbang. Cara untuk memperoleh matriks pembobot atau penimbang *spatial*(W) yaitu dengan menggunakan informasi jarak dari ketetanggan (*neighborhood*), atau kedekatan antara satu *region* dengan *region* yang lain. Lokasi yang dekat dengan lokasi yang diamati diberi pembobot besar, sedangkan yang jauh diberi pembobot kecil. Pemberian koding pembobot menurut Bivad dalam Kissling dan Carl (2007), diantaranya pada persamaan berikut ini:

1. Kode Biner

$$W_{ij} = \begin{cases} 1, & untuk\ i\ dan\ j\ yang\ berdekatan \\ 0, & untuk\ lainnya \end{cases}$$

2. *Row Standardization*

Didasarkan pada jumlah tetangga pada suatu baris yang sama pada matriks pembobot

$$\widehat{W}_{ij} = \frac{W_{ij}}{\sum_{j=1}^{n} W_{ij}} \tag{2.11}$$

3. *Varians Stabilization*

Menstabilkan *varians* dengan menjumlahkan semua baris dan kolom

$$\widehat{W}_{ij} = \frac{W_{ij}}{\sum_{i,j=1}^{i,j=n} W_{ij}} \tag{2.12}$$

Tobler dalam Anselin (1988), merumuskan hukum *first law of geography* yang berbunyi "*everything is related to everything else, but near things are more related than distant things*" artinya segala sesuatu saling berkaitan satu sama lainnya, wilayah yang lebih dekat cenderung akan memberikan efek yang lebih besar dari pada wilayah yang lebih jauh jaraknya. Ada beberapa metode untuk mendefinisikan hubungan persinggungan (*contiguity*) antar wilayah tersebut.

Menurut LeSage (1999), metode *contiguity* terdiri dari:

1. *Linier Contiguity* (persinggungan tepi) adalah lokasi yang berada ditepi kiri maupun kanan dari lokasi yang menjadi perhatian dari pembobotan $W_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $W_{ij} = 0$.

2. *Rook contiguity* (persinggungan sudut) adalah lokasi yang bersisian dengan lokasi yang menjadi perhatian diberi pembobotan $W_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $W_{ij} = 0$.
3. *Bishop contiguity* (persinggugan sudut) adalah lokasi yang titik sudutnya bertemu dengan sudut lokasi yang menjadi perhatian diberi pembobotan $W_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $W_{ij} = 0$.
4. *Double linier contiguity* (persinggungan dua tepi) adalah lokasi yang berada di sisi kiri kanan lokasi yang menjadi perhatian diberi pembobot $W_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $W_{ij} = 0$.
5. *Double rook contiguity* (persinggungan dua sisi) adalah lokasi yang berada di kiri, kanan, utara, dan selatan lokasi yang menjadi perhatian diberi pembobotan $W_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $W_{ij} = 0$.
6. *Queen contiguity* (persinggungan sisi-sudut) adalah lokasi yang bersisian atau titik sudutnya bertemu dengan lokasi yang menjadi perhatian diberi pembobotan $W_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $W_{ij} = 1$.

2.5 *Outlier*

Belum ada patokan yang disepakati para statistikawan kapan suatu pengamatan dapat dikategorikan sebagai *outlier*. Secara umum *outlier* dapat diartikan data yang tidak mengikuti pola umum model dan secara kasar, dapat diambil patokan yaitu yang sisanya berjarak 3 kali simpangan baku atau lebih dari rata-ratanya (yaitu nol) (Sembiring, 1995). Menurut Draper (1992), sisaan yang merupakan *outlier* adalah yang nilai mutlaknya jauh lebih besar dari pada sisaan-

sisaan lainnya dan bisa jadi terletak tiga atau empat kali simpangan baku atau lebih jauh lagi darirata-rata sisaannya.

Keberadaan *outlier*akan mengganggu dalam proses analisis data dan harus dihindari. Dalam kaitannya analisis regresi, *outlier* dapat menyebabkan hal-hal berikut:

1. *Error* yang besar dari model yang terbentuk,
2. *Varians* pada data tersebut menjadi lebih besar,dan
3. Taksiran interval memiliki rentang yang lebar (Soemartini, 2007).

Penghapusan atau penolakan terhadap *outlier* yang muncul seringkali dilakukan, hal ini kurang tepat dilakukan karena pada amatan mungkin terdapat informasi yang tidak terdapat pada titik data yang lain, misalnya timbul karena adanya kombinasi keadaan yang tidak biasa yang mungkin saja sangat penting yang perlu diselidiki lebih jauh. Oleh karena itu, sangat disarankan untuk menguji dahulu apakah *outlier* yang ada benar-benar memiliki pengaruh atau tidak (Drapper dan Smith, 1992).

Menurut Montgomery (2006) pada analisis regresi, terdapat beberapa tipe *outlier* yang mempengaruhi hasil estimasi kuadrat terkecil. Klasifikasi secara umum sebagai berikut:

1. *Vertical outlier,* merupakan suatu titik yang menjadi *outlier* karena memiliki koordinat y yang ekstrim. Dalam penduga kuadrat terkecil, *vertical outlier* sangat berpengaruh khususnya pada penduga intersep.
2. *Good leverage point*, merupakan suatu titik yang menjadi *outlier* pada variabel independen tetapi terletak dekat dengan garis linier, yang berarti bahwa observasi (x_i, y_i) apabila x_i menjauh tetapi y_i cocok dengan garis linier. *Good*

leverage ini tidak berpengaruh terhadap estimasi kuadrat terkecil, tetapi berpengaruh terhadap inferensi statistik karena dapat meningkatkan estimasi standart *error*.

3. *Bad leverage point*, merupakan suatu titik yang menjadi *outlier* pada variabel independen tetapi terletak jauh dengan garis linier. *Bad leverage*ini berpengaruh signifikan terhadap estimasi kuadrat terkecil, baik terhadap intersep maupun *slope* dari persamaan regresi.
4. Pencilan*regresi*adalah sebuah titik yang menyimpang dari hubungan linier yang ditentukan dari sisaan $(n-1)$ pengamatan.
5. Pencilan sisaanadalah sebuah titik mempunyai standar sisaan yang besar. Catatan bahwa sebuah titik dapat menjadi pencilan regresi tanpa menjadi pencilan sisaan (ini akan selalu terjadi jika titik sangat berpengaruh).

Metode yang digunakan untuk mengidentifikasi adanya *outlier* yang berpengaruh dalam koefisien regresi adalah sebagai berikut:

1. Diagram Pencar (*Scatter Plot*)

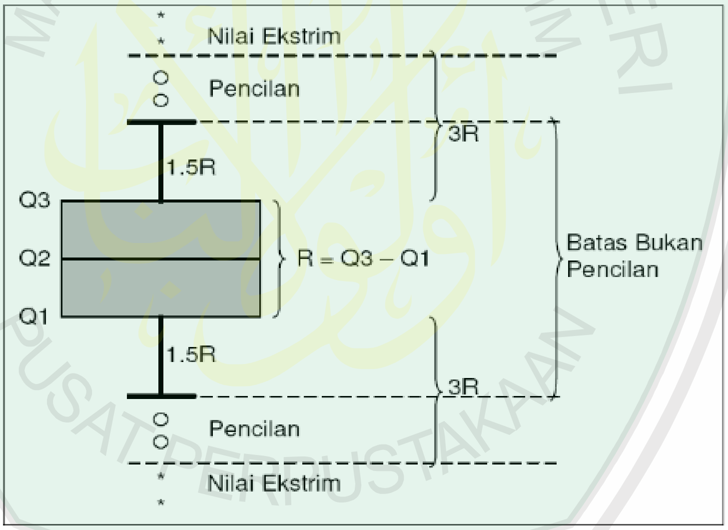
Metode diagram pencar dilakukan dengan cara memplot data dengan observasi ke-$i(i = 1,2, \dots, n)$. Selain itu, jika sudah didapatkan model regresi maka dapat dilakukan dengan cara memplot antara residual dengan nilai prediksi Y. Jika terdapat satu atau beberapa data yang terletak jauh dari pola kumpulan data keseluruhan maka hal ini mengindikasikan adanya *outlier*.

Keuntungan dari metode ini adalah mudah untuk dipahami karena menampilkan data secara grafis dan tanpa melibatkan perhitungan yang rumit. Sedangkan kelemahan pada metode ini adalah keputusan yang memperlihatkan data yang merupakan *outlier* atau bukan hanya tergantung pada kebijakan peneliti, karena hanya mengandalkan visualisasi melalui gambar.

2. Boxplot

Metode ini merupakan yang paling umum yakni dengan mempergunakan kuartil dari jangkauan. Kuartil 1, 2, dan 3 akan membagi sebuah urutan data menjadi empat bagian. Jangkauan (IQR, *Interquartil Range*) didefinisikan sebagai selisih kuartil 3, atau $IQR = Q_3 - Q_1$.

Data-data *outlier* dapat ditentukan yaitu nilai dengan kuartil yang kurang dari 1.5 $\times IQR$ terhadap kuartil 1 dan nilai dengan kuartil yang lebih dari 1.5 $\times$ IQR terhadap kuartil 3.



Gambar 2.1 Identifikasi *Outlier*

3. Metode *DfFITS (Difference fitted value FITS)* atau *Standardized DfFITS*

Metode ini menampilkan nilai perubahan dalam harga yang diprediksi bilamana*case* tertentu dikeluarkan, yang sudah distandarkan. Perhitungan *DfFITS* adalah sebagai berikut:

$$(DfFITS)_I = t_i \left(\frac{h_{ii}}{1 - h_{ii}}\right)^{\frac{1}{2}} \tag{2.13}$$

Dimana $t_i$adalah*studentized deleted* untuk kasus ke-i dan h_{ii} adalah nilai *leverage* untuk kasus ke-i, dengan:

$$t_i = e_i \sqrt{\frac{n-p-1}{JKG(1-h_{ii}-e_i{}^2}} \qquad (2.14)$$

Dimana$e_i$adalah residual ke-i dan JKG adalah jumlah kuadrat *error*dalam matriks adalah sebagai berikut:

$$H = X(X'X)^{-1}X' \qquad (2.15)$$

dengan H adalah matriks$n \times n$,

Elemen diagonal h_{ii} dalam matriks dapat diperoleh langsung dari:

$$h_{ii} = X_i(X'X)^{-1}X_i' \qquad (2.16)$$

Dengan X_i adalah matriks $p \times 1$, (X$(X'X)^{-1}$ adalah matriks $p \times p$, dan $X_i{}'$ adalah matriks $1 \times p$.

Suatu data yang mempunyai nilai *absolute DfFITS* lebih besar dari$\sqrt[2]{\frac{p}{n}}$. Maka diidentifikasikan sebagai *outlier*, dengan p banyaknya variabel independent dan banyaknya observasi.

2.6 Spatial Outlier

Spatial outlier menurut Shekhar, dkk (2003) adalah suatu titik *spatial* tertentu dimana nilai-nilai atribut *non-spatial*nya berbeda nyata dari titik-titik lain yang masih menjadi tetangganya.

Spatial outlier merupakan objek yang tereferensi secara *spatial* dimana atribut *non-spatial*nya sangat berbeda dengan lingkungannya. Tujuan *spatial outlier detection* untuk mencari ketidakstabilan lokal yang melanggar *spatial autocorrelation* dan kontinuitas. *Spatial outocorrelation* adalah korelasi antara nilai-nilai dari variabel tunggal disebabkan oleh posisi yang dekat pada permukaan dua dimensi. *Spatial autocorrelation* mengukur dan menganalisis tingkat ketergantungan pengamatan di antara geografis. *Spatial autocorrelation* membandingkan *spatial weighted* untuk hubungan kovarian (ukuran dari seberapa banyak dua set data yang berbeda-beda) di sejumlah lokasi.

Berbeda dengan *outlier* tradisional, *spatial outlier* adalah *anomaly local* yang ekstrim dibandingkan dengan lingkungannya, tetapi tidak selalu menyimpang dari sisa semua dataset yang ada. Secara tidak langsung *spatial outlier* dapat disebut juga "*local outlier"* karena *spatial outlier* selalu memperhatikan perbedaan-perbedaan lokal, sedangkan *traditional outlier* kita sebut dengan "*global outlier*" karena fokus terhadap perbandingan-perbandingan global.

Data-data yang digunakan dalam mendeteksi *spatial outlier* disebut data *spatial*, data *spatial* adalah data-data yang terdiri dari geometri dan topologi, misalnya bentuk, lokasi, ukuran, dll. Ada beberapa hal signifikan yang membedakan antara data *spatial*dengan data *non-spatial* di antaranya sebagai berikut:

1. Data *spatial* terdiri dari stuktur yang kompleks seperti titik, garis, daerah, bahkan objek 3-D.
2. Memiliki dataset yang lebih besar dari pada *non-spatial* dataset.
3. Data *spatial* digunakan untuk mekanisme tertentu seperti *storage*, *indexing*, dan *querying*.

Spatial data bisa dikategorikan dalam 2 grup, *macro spatial* dan *microspatial*. *Macro spatial* data terdiri dari geospasial data, sedangkan *micro spatial* data lebih kecil seperti lokasi dan bentuk.

Sekumpulan data *spatial* dapat dimodelkan sebagai kumpulan objek yang tereferensi secara *spatial*. Objek *spatial* memiliki dua kategori dimensi yang sangat berbeda sesuai dengan atribut mana yang diukur. Kategori tersebut terdiri atas:

1. Atribut *spatial* terdiri dari objek yang tereferensi secara *spatial* seperti lokasi, bentuk, dan geometrik atau topologi lainnya.
2. Atribut *non-spatial* dari objek yang tereferensi secara *spatial* seperti hasil suara, umur, dan pemilik.

2.7 Fungsi Objektif

Menurut Fox (2002) fungsi objektif adalah fungsi yang digunakan untuk mencari pembobot pada regresi *robust*. Adapun fungsi pembobot yang digunakan antara lain sebagai berikut:

1. Fungsi pembobot *Tukey* memakai fungsi objektif

$$\rho(e_i) = f(x) = \begin{cases} \dfrac{c^2}{6}\left\{1 - \left[1 - \left(\dfrac{e_i}{c}\right)^2\right]^3\right\}, & |e_i| \leq c \\ \dfrac{c^2}{6}, & |e_i| > c \end{cases} \tag{2.17}$$

dengan

$$\rho'^{(e_i)} = \psi(e_i) = \frac{\partial(\rho(e_i))}{\partial e_i} = \begin{cases} e_i\left[1 - \left(\dfrac{e_i}{c}\right)^2\right]^2, & |e_i| \leq c \\ 0, & |e_i| > c \end{cases} \tag{2.18}$$

Setelah didapatkan $\rho'^{(e_i)}$ maka didapatkan fungsi pembobot:

$$w_i = w(e_i) = \frac{\psi(e_i)}{e_i} = \begin{cases} \left[1 - \left(\frac{e_i}{c}\right)^2\right]^2, & |e_i| \leq c \\ 0, & |e_i| > c \end{cases} \tag{2.19}$$

2. FungsipembobotHuber memakai fungsi objektif

$$\rho(e_i) = \begin{cases} \frac{1}{2} {e_i}^2, & |e_i| \leq c \\ c|e_i| - \frac{1}{2} c^2, & |e_i| > c \end{cases} \tag{2.20}$$

dengan

$$\rho'^{(e_i)} = \psi(e_i) = \frac{\partial(\rho(e_i))}{\partial e_i} = \begin{cases} e_i, & |e_i| \leq c \\ c, & e_i > c \\ -c, & e_i < -c \end{cases} \tag{2.21}$$

Setelah didapatkan $\rho'^{(e_i)}$ maka didapatkan fungsi pembobot:

$$w_i = w(e_i) = \frac{\psi(e_i)}{e_i} = \begin{cases} 1, & |e_i| \leq c \\ \frac{c}{|e_i|}, & |e_i| > c \end{cases} \tag{2.22}$$

Dimana kostanta c adalah konstanta yang menghasilkan efisiensi tinggi dengan residual berdistribusi normal dan dapat memberikan perlindungan terhadap *outlier*. Untuk fungsi pembobot *Huber* nilai c =1,345 dan c = 4,685 untuk fungsi pembobot *Tukey Bisquare* (Fox, 2002).

2.8 Regresi *Robust*

Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari sisaan tidak normal atau ada beberapa *outlier* yang berpengaruh pada model. Model ini merupakan alat penting untuk menganalisis data yang dipengaruhi oleh *outlier* sehingga dihasilkan model yang kekar terhadap *outlier* (Draper dan Smith, 1998). Ketika peneliti menyusun model regresi dan melakukan uji asumsi sering ditemui bahwa asumsi regresi dilanggar, transformasi yang dilakukan tidak akan menghilangkan atau melemahkan pengaruh dari pencilan yang akhirnya prediksi menjadi bias. Dalam keadaan ini, regresi *robust* yang tahan terhadap pengaruh pencilan adalah metode yang terbaik. Regresi *robust* digunakan untuk mendeteksi pencilan dan memberikan hasil yang resisten terhadap adanya pencilan (Chen, 2002).

Dalam regresi *robust* salah satu metode estimasi yang terkenal adalah estimasi-M. Huruf M menunjukkan bahwa estimasi-M adalah estimasi tipe maksimum *likelihood.* Estimasi-M memenuhi sifat sebagai estimator tak bias dan memiliki *varians* minimum dalam kumpulan estimator. Jadi estimator M memiliki *varians* terkecil dibandingkan dengan *varians* estimator yang lain.

Berikut ini merupakan estimasi koefisien regresi *robust*menggunakan estimasi-M:

1. Menghitung parameter $\hat{\beta}^0$ dengan metode kuadrat terkecil.
2. Menghitung nilai sisaan $e_i = y_i - \hat{y}_i$.
3. Menghitung nilai $\hat{\sigma}_i = \frac{MAD}{0,6745} = \frac{median\,|e_i - median\,(e_i)|}{0.6745}$.
4. Menghitung nilai $u_i = \frac{e_i}{\partial_i}$.
5. Menghitung pembobot.
6. Menghitung parameter $\hat{\beta}_M$ dengan metode *Weighted Least Squares* (WLS). dengan pembobot w_i.
7. Mengulangi langkah 2-6 sampai diperoleh nilai $\hat{\beta}_M$ yang konvergen.

2.9 Metode Pendeteksian *Spatial Outlier*

2.9.1. DefinisiAlgoritma*Spatial Outlier*

DefinisiAlgoritma*spatialoutlier* dinyatakan sebagai berikut:

1. Set data *spatial*, $X = (x_1, x_2, \ldots, x_n)$ dalam ruang dengan dimensi $p \geq 1$.

2. Tentukanfungsi atribut f digambarkan sebagai suatu pemetaan dari X untuk R (set dari angka riil). Fungsi atribut $f(x_i)$ menyatakan nilai atribut dari titik *spatial*x_i.
3. Untuk titik yang ditentukan x_i, $NN_k(x_i)$ menandakan $k$tetangga terdekat x_i, dengan$k = k(x_i)$ tergantung pada nilai x_i, untuk $i = 1,2, \ldots, n$.
4. Fungsi tetangga g digambarkan sebagai suatu pemetaan dari X untuk R, dimana untuk masing-masing x_i, $g(x_i)$ mengembalikan suatu ringkasan statistik nilai atribut dari semua titik *spatial* di dalam $NN_k(x_i)$. Sebagai contoh, $g(x_i)$ berupa rata-rata nilai atribut k tetangga terdekat x_i. untuk mendeteksi *spatialoutlier*, kita bandingkan nilai atribut dari tiap titik x_i dengan nilai atribut tetangganya $NN_k(x_i)$.
5. Perbandingannilai atribut dari tiap titik x_i dengan nilai atribut tetangganya $NN_k(x_i)$ dilaksanakan melalui suatu fungsi pembanding h, yang merupakan suatu fungsi dari f dan g. Ada banyak pilihan untuk format h. Sebagai contoh, h dapat berupa selisih, $f - g$, atau rasio $\frac{f}{g}$.
6. Misalkan$y_i = h(x_i)$dengan$i = 1,2, \ldots, n$,y_i merupakan standarisasi dari fungsi $h(x_i)$ yang diberi tanda mutlak.
7. Jika y_i adalah suatu nilai yang ekstrim pada set $(y_1, y_2, \ldots, y_n)$, nyatakan titik x_i sebagai *spatialoutlier.*

Dapat digarisbawahi definisi tersebut tergantung pada pemilihan fungsi $k, g,$ dan $h$yang digunakan (Lu, dkk, 2003).

2.9.2. *k* Tetangga Terdekat

Metode *k-nearest neighbor* atau yang sering disebut sebagai k tetangga terdekat ($k - NN$ atau KNN) adalah sebuah pendekatan pengklasifikasian data dengan prinsip kerja yang mudah dan efisien. Metode tersebut mengasumsikan data yang saling berdekatan sebagai tetangga terdekat dan dimasukkan dalam kelas yang sama. Tetangga terdekat tersebut ditentukan berdasarkan perhitugan *vector* dan jarak. Jarak yang sering digunakan untuk menghitung *k* tetangga terdekat adalah jarak Euclid (Jiang Sheng, 2002).

Menurut Shekhar, dkk (2003), metode *Connectivity Clustered Acces Method* (CCAM) disarankan sebagai alternative metode clustering yang digunakan untuk menentukan *k* tetangga terdekat. Pada metode CCAM menggunakan *edge* yaitu garis koneksi dari titik-titik *spatial* sebagai ukuran kedekatan dalam menentukan cluster.

Edge dapat dihitung dengan menggunakan jarak Euclid. Prosesalgoritma CCAM diawali dengan menganggap seluruh titik pada dataset sebagai satucluster. Kemudian, titik dengan konektivitas terdekat membentuk cluster tersendiri. *Edge* tersebut merupakan ukuran *spatial* (Lin dan Ye, 2008).

2.9.3. Ukuran keruangan (Ukuran *Spatial*)

Keterikatan atau hubungan *spatial* adalah salah satu sifat dalam lingkungan geografi, karakter dari lokasi yang saling berdekatan akan menjadi berkolerasi. Terdapat beberapa alasan terkait dengan adanya suatu hubungan *spatial*. Kemungkinan pertama, karena terdapat hubungan korelasi *spatial* sederhana: apapun yang menyebabkannya, objek pada suatu lokasi memiliki apapun yang menyebabkannya, objek pada suatu lokasi memiliki kemiripan dengan objek lain yang berdekatan. Kemungkinan kedua adalah hubungan kuasalitas *spatial*, objek pada suatu lokasi tertentu akan secara langsung mempengaruhi objek lain disekitarnya. Kemungkinan ketiga adalah adanya interaksi *spatial*, mobilitas masyarakat, barang, dan informasi menciptakan hubungan yang nyata antara beberapa lokasi (Upton dan Fingelton, 1985).

Ukuran kedekatan didapatkan dari perhitungan hubungan *spatial* antara titik yang berbeda. Setiap objek *spatial* memiliki posisi dalam sistem koordinat tertentu dan masing-masing memiliki keterkaitan atau hubungan *spatial*. Hubungan *spatial*tersebut dapat diukur dengan tipologi, jarak, dan hubungan langsung (Schewering dan Raubal, 2005). Jarak yang dapat menggambarkan kedekatan tersebut adalah jarak *Euclidean* dan *Manhattan*.

Jarak *Euclidean* adalah Jarak lurus antara dua titik dengan menggunakan fungsi *Pythagoras*. Jika titik $P = (x_1, x_2)$ pada suatu permukaan, jarak *Euclidean*, $d = (O, P)$ dari titik P ke titik pusat $O = (0,0)$ berdasarkan pada teorema *Pythagoras* adalah:

$$d(O,P) = \sqrt{x_1^2 + x_2^2} \tag{2.23}$$

Secara umum, jika titik P memiliki koordinat sehingga $P = (x_1, x_2, \dots, x_p)$ maka jarak *Euclidean* dari titik P ke titik pusat $O = (0,0,\dots,0)$ adalah:

$$d(O,P) = \sqrt{x_1^2 + x_2^2 + \cdots + x_p^2} \tag{2.24}$$

Sedangkan jarak *Euclidean* antara dua titik P dan Q dengan koordinat $P = (x_1, x_2, \dots, x_p)$ dan $Q = (y_1, y_2, \dots, y_p)$ untuk n-dimensi adalah:

$$d(P,Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + \left(x_p - y_p\right)^2} \tag{2.25}$$

(Johnson dan Winchern, 2002).

Jarak *Euclidean* antara dua titik P dan Q dengan koordinat $P = (x_1, x_2)$ dan $Q = (y_1, y_2)$ untuk dua dimensi adalah:

$$d(P,Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \tag{2.26}$$

Jarak *Manhattan* adalah jarak yang dibatasi untuk berjalan sejajar dengan sumbu (Upton dan Fingelton, 1985). Pada jarak *Manhattan* jarak antara dua objek adalah jumlah dari harga mutlak selisih koordinat kedua objek. Jarak *Manhattan* antara dua objek (Pdan Q) pada dimensi-p adalah sebagai berikut (Krause dalam Wikipedia, 1987):

$$d(P,Q) = \sum_{i=1}^{p} |x_i - y_i| \tag{2.27}$$

dengan:

$d(P,Q)$: Jarak *Manhattan* antara objek P dengan objek Q

x_i : objek P pada dimensi-p dimana $x_i = (x_1, x_2, \dots, x_p)$

y_i : objek Q pada dimensi-p dimana $y_i = (y_1, y_2, \dots, y_p)$

Sehingga pada ruang dimensi, jarak *Manhattan* antara objek p dengan objek q adalah (Krause dalam Wikipedia, 1987):

$$\mathrm{d}(\mathrm{p},\mathrm{q}) = |\mathrm{p}_1 - \mathrm{q}_1| + |\mathrm{p}_2 - \mathrm{q}_2| \quad ` \tag{2.28}$$

Dibandingkan dengan jarak *Euclidean*,jarak *Manhattan* lebih baik digunakan didaerah yang memiliki banyak batasan yang tidak dapat dilanggar seperti di perkotaan yang banyak dibatasi oleh gedung. Oleh karena itulah jarak *Manhattan* dapat disebut dengan jarak *City Blok* (Upton dan Fingelton, 1985).

2.9.4. Algoritma Z

Spatialoutlier mengindikasikan bahwa dapat atau tidaknya suatu titik dikatakan *spatialoutlier* tergantung pada selisih antara nilai atribut suatu titik tersebut dengan rata-rata nilai atribut dari seluruh titik yang merupakan tetangganya (Shekhar, dkk, 2003). Algoritma yang diusulkan Shekhar, dkk (2003) menggunakan atribut *non-spatial* tunggal yang membandingkan selisih antara tetangga-tetangga suatu titik dan mengidentifikasi *spatialoutlier* melalui perhitungan yang efisien dengan cara menghitung fungsi *algebraic aggregate* umum. Selain itu, Shekhar, dkk (2003) juga mempertimbangkan struktur grafik data *spatial* dan memanfaatkan suatu metode grafis untuk mendeteksi *spatialoutlier*. Secara rinci, metode ini membandingkan selisih antara nilai atribut dari suatu titik data dan rata-rata atribut tetangganya, dan memeriksa keseluruhan data ke dalam distribusi normal. Jika nilai pengujian pada titik-titik tersebut lebih besar dari suatu taraf kepercayaan yang ditetapkan maka titik tersebut dinyatakan sebagai *spatialoutlier*.

Algoritma pendeteksian *spatialoutlier* dengan metode pendeteksian nilai z dinyatakan sebagai berikut (Shekhar, dkk, 2003):

1. Untuk masing masing titik *spatial*x_i, hitung k tetangga terdekat $NN_k(x_i)$, fungsi rata-rata nilai atribut tetangga dengan persamaan:

$$\mathrm{g}(\mathrm{x_i}) = \frac{1}{\mathrm{k}} \sum_{\mathrm{x} \in \mathrm{NN_k}(\mathrm{x_i})} \mathrm{f}(\mathrm{x}) \qquad (2.29)$$

   Dan fungsipembandingdenganpersamaan:

$$h_i = h(x_i) = f(x_i) - g(x_i) \qquad (2.30)$$

2. $\mu_h$dan σ_h secara berturut-turut merupakan rata-rata dan standar deviasi dari fungsi pembanding $h(x_i)$, yang kemudian digunakan untuk menghitung statistik uji z dimana:

$$\mathrm{Z}_{\mathrm{h}(\mathrm{x_i})} = \left| \frac{\mathrm{h}(\mathrm{x_i}) - \mu_\mathrm{h}}{\sigma_\mathrm{h}} \right| \qquad (2.31)$$

3. Objek *spatial* yang dinyatakan sebagai *spatialoutlier* adalah objek yang memiliki stasistik uji $Z_{h(x_i)} > \theta$ dimana $\theta = 2$ untuk selang kepercayaan 95%.

2.9.5. *WeightedZAlgorithm*

Weighted z algorithm merupakan pengembangan dari *z algorithm*. Perbedaan *Weighted z algorithm* dengan *z algorithm* terletak pada fungsi tetangga (*neighborhood function*) g. jika pada *z algorithm*, fungsi tetangga (*neighborhood function*) yang digunakan merupakan rata-rata nilai atribut *non-spatial* dari masing-masing lingkungan *spatial*, sedangkan pada *weighted z algorithm* menggunakan *neighborhood function* berupa*weighted mean* (rata-rata terboboti) dari masing masing lingkungan *spatial* (Kou, dkk, 2006).

Weighted *mean* (rata-rata terboboti) dalah salah satu ukuran pemusatan. Rata-rata terboboti (M_w) dari n objek pengamatan $(x_1, x_2, \dots, x_n)$ adalah (Glodberg dan Kercheval, 2002):

$$M_w(x_1, x_2, \dots, x_n) = \frac{\sum_{j=1}^{n} w_j x_j}{\sum_{j=1}^{n} w_j} \qquad (2.32)$$

atau

$$M_w(x_1, x_2, \dots, x_n) = \frac{w_1 x_1 + w_2 x_2 + \cdots + w_n x_n}{w_1 + w_2 + \cdots + w_n} \qquad (2.33)$$

Dimana w_i adalah *weight* (bobot) dari objek x_i. Semakin besar bobot w_i, maka semakin besar pula kontribusi objek lain terhadap rata-rata. Rata-rata terboboti sama dengan rata-rata aritmatik pada saat masing-masing objek x_i memberikan kontribusi yang sama pada rata-rata, dengan kata lain jika semua bobot w_i memiliki nilai yang sama, maka rata-rata terboboti sama dengan rata-rata aritmatik. Rata-rata terboboti pada umumnya digunakan pada saat akan mengkombinasikan beberapa nilai rata-rata dari suatu populasi dengan banyak sampel yang berbeda (Hardy, dkk, 1988).

Menurut Kou, dkk (2006), objek *spatial* pada data *spatial* bukan merupakan suatu objek yang terisolasi dari lingkungan sekitarnya. Masing-masing objek *spatial* memiliki sifat saling mempengaruhi objek *spatial* lain yang masih berada dalam lingkungan *spatial* yang sama. Semakin dekat objek *spatial* dengan objek *spatial* lain, maka semakin besar pengaruh objek *spatial* tersebut bagi objek *spatial* yang lain. Berdasarkan sifat tersebut, maka dibuat suatu algoritma pendeteksi *spatialoutlier*. Bobot ditentukan berdasarkan hubungan *spatial* yaitu jarak *Euclidean* antara dua objek *spatial*.

Penentuan bobot melalui hubungan *spatial* antara objek x_i dan objek tetangganya x_j. Nilai bobot untuk x_j terhadap x_i sebesar 0 sampai 1 dan jumlah dari bobot semua objek tetangga x_i adalah 1. Jika x_j adalah tetangga ke r dari x_i, maka bobot dari x_j adalah (Kou, dkk, 2006):

$$w_j = \sum_{p=1}^{q} \alpha_p \cdot \frac{S_{pr}}{\sum_{l=1}^{k} S_{pl}} \qquad (2.34)$$

dengan:

w_j : bobot dari objek x_j terhadap objek x_i

q : banyaknya sifat *spatial* yang menentukan bobot

S_{pr} : *particular spatial property* atau atribut *spatial*S_p untuk objek tetangga ke r dari x_i, dalam hal ini S_{pr} merupakan invers jarak antara x_i dengan x_r.

S_{pl} : particular *spatial property* atau atribut *spatial*S_p untuk objek tetangga ke l dari x_i, dalam hal ini S_{pl} merupakan invers jarak antara x_i dengan x_l.

α_p : faktor yang menyatakan tingkat kepentingan sifat *spatial*S_p dan $\sum_{p=1}^{q} \alpha_p = 1$

Rata-rata terboboti (M_w) didapatkan dengan mengalikan bobot dengan nilai atribut *non-spatial*y_j untuk masing-masing neighbor x_j (Kou, dkk, 2006).

$$M_w(x_i) = \sum_{j=1}^{k} f(x_j) * w_j \tag{2.35}$$

dengan:

$M_w(x_i)$	:	rata-rata terboboti untuk objek *spatial*x_i
k	:	banyak tetangga terdekat dengan objek x_i
$f(x_j)$	:	fungsi atribut tetangga, x_j
w_j	:	bobot dari objek tetangga (x_j) terhadap objek x_i

Berikut ini algoritma pendeteksian *spatialoutlier* dengan *weighted z algorithm* (Kou, dkk, 2006):

1. Untuk masing-masing objek, menentukan $k$tetangga terdekat menggunakan jarak *Euclidean* antara dua objek berdasarkan persamaan (2.1).
2. Untuk masing-masing objek x_i, menghitung *neighborhood functiong*(x_i)yang merupakan rata-rata terboboti (M_w) dari atribut *non-spatial* untuk semua objek *spatial*x_i berdasarkan persamaan (2.35).
3. Menghitung fungsi pembanding yaitu selisih antara nilai atribut *non-spatial*x_i dengan *neighborhood functiong*(x_i)atau dilambangkan dengan $h(x_i)$.

$$h(x_i) = f(x_i) - g(x_i) \tag{2.36}$$

4. Menghitung $h(x_i)$yang telah distandarisasi berdasarkan rata-rata dan standar deviasi dari $h(x_i)$ yang kemudian akan dijadikan faktor untuk menentukan *spatialoutlier*.
5. Mengidentifikasi objek dengan nilai $Z(x_i) > \theta$ sebagai *spatialoutlier*, di mana $\theta = 2$ untuk selang kepercayaan *95%*.

2.10 Kajian *Outlier* danEstimasidalam Al-Quran

Al-Quran diturunkan Allah untuk menjadi pegangan bagi mereka yang ingin mencapai kebahagiaan hidup di dunia maupun di akhirat. Al-Quran tidak diturunkan hanya untuk kepentingan suatu umat tertentu atau untuk satu abad tertentu, akan tetapi untuk seluruh umat manusia dan berlaku sepanjang masa. Karena itu luas ajarannya meliputi luasnya persoalan-persoalan atau masalah-masalah yang dihadapi oleh seluruh umat manusia. Seperti halnya dalam al-Quran suratal-A’raaf/7:52, yaitu:

وَلَقَدۡ جِئۡنَٰهُم بِكِتَٰبٖ فَصَّلۡنَٰهُ عَلَىٰ عِلۡمٍ هُدٗى وَرَحۡمَةٗ لِّقَوۡمٖ يُؤۡمِنُونَ ٥٢

“Dan Sesungguhnya kami Telah mendatangkan sebuah Kitab (Al Quran) kepada mereka yang kami Telah menjelaskannya atas dasar pengetahuan Kami[546]; menjadi petunjuk dan rahmat bagi orang-orang yang beriman”(QS. al-A’raaf/7:52).

Al-Quran bukan hanya mencakup tentang ilmu agama, melainkan banyak hal yang berkaitan dengan masalah ekonomi, hukum, politik, sosial, dan sains dan teknologi. Oleh karena itu, disini akan dibuktikan bahwa al-Quran bukan hanya menjelaskan tentang ilmu agama akan tetapi membahas tentang ilmu statistik juga

2.10.1 *Outlier* dalam Al-Quran

Dalam al-Quran telah disinggung terkait dengan permasalahan *outlier* yaitu terdapat pada surat al-Jin/72:14, yaitu:

وَأَنَّا مِنَّا ٱلۡمُسۡلِمُونَ وَمِنَّا ٱلۡقَٰسِطُونَۖ فَمَنۡ أَسۡلَمَ فَأُوْلَٰٓئِكَ تَحَرَّوۡاْ رَشَدٗا ١٤

“Dan Sesungguhnya di antara kami ada orang-orang yang taat dan ada (pula) orang-orang yang menyimpang dari kebenaran. barangsiapa yang yang taat, Maka mereka itu benar-benar telah memilih jalan yang lurus”(QS. Al-Jin/72:14).

Asal turunnya surat al-Jin/72:14 yaitu untuk menampik dugaan bahwa semua jin baik yang mendengar langsung ayat-ayat al-Quran maupun yang belum atau tidak mendengarnya kesemuanya telah patuh kepada Allah. Kemudian pada ayat tersebut diterangkan bahwa *dan sesungguhnya diantara kami* masyarakat jin *ada orang-orang muslim* yakni yang benar-benar taat dan penuh kepatuhan kepada Allah dan ada pula *para penyimpang* yakni mereka yang telah sangat jauh dari kebenaran lagi sangat mantap kekufurannya. *Barang siapa yang patuh, maka mereka itu telah* bersungguh-sungguh *memilih arah* yang mengantar ke jalan *kebenaran* (Shihab, 2003).

Outlier juga dijelaskan dalam al-Quran pada surat al-A'raaf/7:180, yaitu:

وَلِلَّهِ ٱلۡأَسۡمَآءُ ٱلۡحُسۡنَىٰ فَٱدۡعُوهُ بِهَاۖ وَذَرُواْ ٱلَّذِينَ يُلۡحِدُونَ فِيٓ أَسۡمَٰٓئِهِۦۚ سَيُجۡزَوۡنَ مَا كَانُواْ يَعۡمَلُونَ ١٨٠

"Hanya milik Allah asmaa-ulhusna[585], maka bermohonlah kepada-Nya dengan menyebut asmaa-ulhusna itu dan tinggalkanlah orang orang yang menyimpang dari kebenaran dalam (menyebut) nama-nama-Nya[586], nanti mereka akan mendapat balasan terhadapapa yang telah mereka kerjakan" (QS. Al-A'raaf/7:180)

Dalam surat ini dijelaskan, *Dan Allah memiliki Asmaul husna* karena nama-nama tersebut menunjukkan nama-nama yang agung. Contohnya:

1. *Al'Aliim* (Maha Mengetahui) yang menunjukkan bahwa Dia memiliki ilmu yang meliputi segala sesuatu, tidak keluar dari pengetahuan-Nya seberat biji dzarrah pun dilangit maupun di bumi.
2. *Ar Rahiim* yang menunjukkan bahwa Dia memiliki kekuasaan yang menyeluruh, tidak dapat dikalahkan oleh sesuatu.

dan tinggalkanlah orang-orang menyimpang dari kebenaran dalam (menyebut) nama-nama-Nya, misalnya berkata *"Yaa Razzaq, urzuqnaa*"(artinya:Wahai Pemberi rezeki, berilah kami rezeki)*, "Yaa ghafuur, ighfir lii"* (artinya:Wahai

Maha Pengampun, ampunilah aku), *"Yaa rahiiim, irhamnii" (*artinya:Wahai Maha Penyayang, sayangilah aku), dan sebagainya. *Mereka kelak akan mendapat balasan terhadap apa yang telah mereka kerjakan* (Marwan bin Musa, Tafsir Hidayatul Insan).

2.10.2 Estimasidalam Al-Quran

Dalam al-Quran pada surat ash-Shaffat/3:147 terdapat ayat yang mengandung arti tentang estimasi/taksiran, yaitu :

وَأَرْسَلْنَٰهُ إِلَىٰ مِا۟ئَةِ أَلْفٍ أَوْ يَزِيدُونَ ﴿١٤٧﴾

"Dan kami utus dia kepada seratus ribu orang atau lebih" (QS. ash-Shaffat/3:147)

Pada surat ash-Shaffat/3:147 tersebut dijelaskan bahwa nabi Yunus di utus kepada umatnya yang jumlahnya 100.000 orang atau lebih. Pada ayat tersebut terdapat ketidakpastian dalam menentukan jumlah umat Nabi Yunus. Mengapa harus menyatakan 100.000 atau lebih? Mengapa tidak menyatakan dengan jumlah sebenarnya? Bukankah Allah mengetahui yang ghaib dan yang nyata? Bukankah Allah Maha mengetahui segala sesuatu termasuk jumlah umat Nabi Yunus (Abdussakir, 2007)

Dalam ayat tersebut menjelaskan bahwa Nabi Yunus mengutus seratus ribu atau lebih umatnya. Kata seratus ribu atau lebih terkesan tidak pasti atau berupa taksiran. Taksiran inilah dalam matematika disebut estimasi.

Dalam ayat lain juga dijelaskan yaitu pada surat Ali Imron/3:191 yaitu:

ٱلَّذِينَ يَذْكُرُونَ ٱللَّهَ قِيَٰمًا وَقُعُودًا وَعَلَىٰ جُنُوبِهِمْ وَيَتَفَكَّرُونَ فِى خَلْقِ ٱلسَّمَٰوَٰتِ
وَٱلْأَرْضِ رَبَّنَا مَا خَلَقْتَ هَٰذَا بَٰطِلًا سُبْحَٰنَكَ فَقِنَا عَذَابَ ٱلنَّارِ ﴿١٩١﴾

”(yaitu) orang-orang yang mengingat Allah sambil berdiri atau duduk atau dalam keadan berbaring dan mereka memikirkan tentang penciptaan langit dan bumi (seraya berkata): "Ya Tuhan kami, tiadalah Engkau menciptakan Ini dengan sia-sia, Maha Suci Engkau, Maka peliharalah kami dari siksa neraka”(QS.Ali Imron/3:191).

Dijelaskan dalam Tafsir Hidayatul Insan yaitu “*(yaitu) orang-orang yang mengingat Allah sambil berdiri atau duduk atau dalam keadan berbaring”* yakni dalam setiap keadaan. Menurut Ibnu Abbas, bahwa maksudnya mereka melakukan shalat sesuai kemampuan, yakni jika tidak sanggup berdiri, maka sambil duduk, namun demikian, ayat ini mencakup semua dikir lainnya dengan lisan maupun hati. *“dan mereka memikirkan tentang penciptaan langit dan bumi”* yakni memikirkan kekuasaan penciptanya atau memikirkan maksudnya, Ayat ini menunjukkan bawa berpikir merupakan ibadah dan termasuk sifat wali-wali Allah yang mengenal-Nya. Setelah mereka memikirkannya, mereka pun tahu bawa Allah tidak menciptakannya sia-sia.*“(seraya berkata): "Ya Tuhan kami, tiadalah Engkau menciptakan Ini dengan sia-sia”* yakni bahkan disana terdapat dalil sempurnanya kekuasaan-Mu. *“Maha Suci Engkau”* yakni dari menciptakan sesuatu secara main-main. *“Maka peliharalah kami dari siksa neraka”*Termasuk juga di dalamnya meminta surga, karena ketika mereka meminta dilindungi dari neraka, maka secara langsung mereka juga meminta surga, akan tetapi karena besarnya rasa takut dalam hati mereka, maka mereka menyebut sesuatu yang paling merisaukan mereka.

BAB III

PEMBAHASAN

Spatial outlier menurut Shekhar, dkk (2003) adalah suatu titik *spatial* tertentu dimana nilai-nilai atribut *non-spatial*nya berbeda nyata dari titik-titik lain yang masih menjadi tetangganya.

Langkah-langkah untuk mengestimasi parameter pada model regresi *spatial lag* yang mengandung *outlier* dengan menggunakan Metode *Weighted Z Algorithm* adalah sebagai berikut:

3.1 Menentukan Model Regresi *Spatial Lag* yang Mengandung *Outlier*

Model *spatial lag* didapatkan dari persamaan (2.1), jika nilai $\rho \neq 0$ dan $\lambda = 0$ serta diasumsikan bahwa proses autoregressive hanya pada variabel dependen, maka model regresi *spatial Mixed Regressive-Autoregressive* atau model *Spatial lag* adalah sebagai berikut:

$$y = \rho W_1 y + X\beta + \varepsilon \tag{3.1}$$

dengan:

$y_{(n\times 1)}$: vektor variabel dependen

$X_{(n\times(k+1))}$: matriks variabel independen

$\beta_{((k+1)\times 1)}$: vektor koefisien parameter regresi

$W_{1(n\times n)}$: matriks pembobot *spatial*

ρ : parameter koefisien *spatial lag* variabel dependen

λ : parameter koefisien *spatial error*

$\varepsilon_{(n\times 1)}$: vektor *error* yang diasumsikan tidak mengalami autokorelasi, yang berdistribusi normal dengan mean nol dan varians $I\sigma^2$

dari persamaan tersebut di dapatkan model *spatial lag* yang mengandung *outlier* dengan mengasumsikan φ sebagai *outlier*

$$f(x) = \varphi y = \varphi(\rho W_1 y) + \varphi X\beta + \varphi\varepsilon \qquad (3.2)$$

3.2 Menghitung *Neighborhood Function* $g(x_i)$

Untuk menghitung *neighborhood function* $g(x_i)$ yang merupakan rata-rata terboboti dari atribut *non-spatial* untuk semua objek *spatial* x_i dari persamaan (2.35) sehingga didapat sebagai berikut:

$$\mathrm{g(x_i) = M_w(x_i)} = \sum_{\mathrm{j=1}}^{\mathrm{k}} \mathrm{f(x_j) * w_j}$$

sehingga persamaannya menjadi:

$$g(x_i) = \sum_{j=1}^{k} (\varphi(\rho W_1 y) + \varphi X\beta + \varphi\varepsilon) * \mathrm{w_j}$$

$$g(x_i) = \sum_{j=1}^{k} [\varphi(\rho W_1 w_j y) + \varphi X\beta w_j + \varphi w_j \varepsilon] \qquad (3.3)$$

Misal $W_1 \cdot w_j = w^*$, sehingga persamaan (3.4) dapat dituliskan sebagai berikut:

$$g(x_i) = \varphi(\rho w^* y) + \varphi X\beta w_j + \varphi w_j \varepsilon \qquad (3.4)$$

3.3 Menghitung Fungsi Pembanding $h(x_i)$

Menentukan fungsi pembanding $h(x_i)$ yaitu selisih antara nilai atribut *non-spatial* x_i dengan *neighborhood function* x_i melalui persamaan (2.36) yaitu:

$$h(x_i) = f(x_i) - g(x_i)$$

sehingga persamaan (3.5) dapat ditulisakan sebagai berikut:

$$h(x_i) = \big((\rho W_1 y\varphi) + X\beta\varphi + \varepsilon\varphi\big) - \big((\rho w^* y)\ \varphi + X\beta w_j\varphi + \varepsilon w_j\varphi\big)$$

$$= (\rho W_1 y)\varphi + X\beta\varphi + \varepsilon\varphi - (\rho w^* y)\varphi - X\beta w_j\varphi - \varepsilon w_j\varphi$$

$$= (\rho W_1 y)\varphi - (\rho w^* y)\varphi + X\beta\varphi - X\beta w_j\varphi + \varepsilon\varphi - \varepsilon w_j\varphi$$

$$= (\rho W_1 y)\varphi - \big(\rho W_1 w_j y\big)\varphi + X\beta\big(\varphi - \varphi w_j\big) + \varepsilon\big(\varphi - \varphi w_j\big)$$

$$= (\rho W_1 y)\big(\varphi - \varphi w_j\big) + X\beta\big(\varphi - \varphi w_j\big) + \varepsilon\big(\varphi - \varphi w_j\big)$$

$$h(x_i) = (\rho W_1 y)\big(I - w_j\big)\varphi + X\beta\big(I - w_j\big)\varphi + \varepsilon\big(I - w_j\big)\varphi \tag{3.5}$$

$$y = (\rho W_1 y)\big(I - w_j\big)\varphi + X\beta\big(I - w_j\big)\varphi + \varepsilon\big(I - w_j\big)\varphi$$

$$y - (\rho W_1 y)\big(I - w_j\big)\varphi = X\beta\big(I - w_j\big)\varphi + \varepsilon\big(I - w_j\big)\varphi$$

$$y(I - \rho W_1)\big(I - w_j\big)\varphi = X\beta\big(I - w_j\big)\varphi + \varepsilon\big(I - w_j\big)\varphi$$

$$\varepsilon\big(I - w_j\big)\varphi = y(I - \rho W_1)\big(I - w_j\big)\varphi - X\beta\big(I - w_j\big)\varphi \tag{3.6}$$

$$\varepsilon = y(I - \rho W_1) - X\beta$$

3.4 Menghitung Jumlah Kuadrat *Error*

Untuk menyelesaikan fungsi di atas, dilakukan dengan cara meminimumkan fungsi objektif (meminimumkan residual φ) dengan persamaan berikut:

$$\sum_{i=1}^{n} \varepsilon\big(I - w_j\big)\varphi = 0$$

Berdasarkan persamaan (3.7) dapat dituliskan sebagai berikut:

$$\textstyle\sum_{i=1}^{n} y(I - \rho W_1)\big(I - w_j\big)\varphi - X\beta\big(I - w_j\big)\varphi = 0 \tag{3.7}$$

Berdasarkan persamaan (3.6) dan (3.7), maka fungsi jumlah kuadrat *error* yang mengandung *outlier* adalah sebagai berikut:

$$SSE = \left(y(I-\rho W_1)\left(I-w_j\right)\varphi - X\beta\left(I-w_j\right)\varphi\right)^T\left(y(I-\rho W_1)\left(I-w_j\right)\varphi - X\beta\left(I-w_j\right)\varphi\right)$$

$$= \left(\left(I-w_j\right)\varphi(y(I-\rho W_1)-X\beta)\right)^T\left(\left(I-w_j\right)\varphi(y(I-\rho W_1)-X\beta)\right)$$

$$= (y(I-\rho W_1)-X\beta)^T\left(\left(I-w_j\right)\varphi\right)^T\left(\left(I-w_j\right)\varphi\right)(y(I-\rho W_1)-X\beta)$$

$$= (y(I-\rho W_1)-X\beta)^T\left(\left(I-w_j\right)\varphi\right)(y(I-\rho W_1)-X\beta)$$ sifat Idempoten $M'M$ (Aziz, 2010)

$$= ((y(I-\rho W_1)-X\beta)^T)\left(y(I-\rho W_1)\left(I-w_j\right)\varphi - X\beta\left(I-w_j\right)\varphi\right)$$

$$= y^T(I-\rho W_1)^T - X^T\beta^T\left(y(I-\rho W_1)\left(I-w_j\right)\varphi - X\beta\left(I-w_j\right)\varphi\right)$$

$$= y^T(I-\rho W_1)^T y(I-\rho W_1)\left(I-w_j\right)\varphi - X^T\beta^T y(I-\rho W_1)\left(I-w_j\right)\varphi - y^T(I-\rho W_1)^T X\beta\left(I-w_j\right)\varphi + X^T\beta^T X\beta\left(I-w_j\right)\varphi$$

$$= y^T(I-\rho W_1)^T y(I-\rho W_1)\left(I-w_j\right)\varphi - X^T\beta^T y(I-\rho W_1)\left(I-w_j\right)\varphi - \left(y^T(I-\rho W_1)^T X\beta\left(I-w_j\right)\varphi\right)^T + X^T\beta^T X\beta\left(I-w_j\right)\varphi$$

$$= y^T(I-\rho W_1)^T y(I-\rho W_1)\left(I-w_j\right)\varphi - X^T\beta^T y(I-\rho W_1)\left(I-w_j\right)\varphi - X^T\beta^T y(I-\rho W_1)\left(I-w_j\right)\varphi + X^T\beta^T X\beta\left(I-w_j\right)\varphi$$

$$= y^T(I-\rho W_1)^T y(I-\rho W_1)\left(I-w_j\right)\varphi - 2X^T\beta^T y(I-\rho W_1)\left(I-w_j\right)\varphi + X^T\beta^T X\beta\left(I-w_j\right)\varphi$$

Sehingga didapatkan:

$$\varepsilon^T(I-w_j)\varphi\varepsilon = y^T(I-\rho W_1)^T y(I-\rho W_1)(I-w_j)\varphi \qquad (3.8)$$

$$-2X^T\beta^T y(I-\rho W_1)(I-w_j)\varphi + X^T\beta^T X\beta(I-w_j)\varphi$$

3.5 Menghitung $\hat{\boldsymbol{\beta}}$ dan Pembobot $\boldsymbol{\omega_i}$

Untuk meminimumkan persamaan (3.8) dapat dilakukan dengan cara mencari turunan pertama $\varepsilon^T(I-w_j)\varphi\varepsilon$ terhadap β^T

$$\frac{\partial}{\partial\beta^T}\varepsilon^T(I-w_j)\varphi\varepsilon = 0$$

$$\frac{\partial}{\partial\beta^T}y^T(I-\rho W_1)^T y(I-\rho W_1)(I-w_j)\varphi - 2X^T\beta^T y(I-\rho W_1)(I-w_j)\varphi$$

$$+X^T\beta^T X\beta(I-w_j)\varphi = 0$$

$$0 = 0 - 2X^T y(I-\rho W_1)(I-w_j)\varphi + X^T(I-w_j)\varphi X\beta + X^T\beta^T X(I-w_j)\varphi$$

$$0 = -2X^T y(I-\rho W_1)(I-w_j)\varphi + X^T(I-w_j)\varphi X\beta + \left(X^T\beta^T X(I-w_j)\varphi\right)^T$$

$$0 = -2X^T y(I-\rho W_1)(I-w_j)\varphi + X^T(I-w_j)\varphi X\beta + X^T(I-w_j)\varphi X\beta$$

$$0 = -2X^T y(I-\rho W_1)(I-w_j)\varphi + 2X^T(I-w_j)\varphi X\beta$$

$$-2X^T(I-w_j)\varphi X\beta = -2X^T y(I-\rho W_1)(I-w_j)\varphi$$

$$X^T(I-w_j)\varphi X\beta = X^T y(I-\rho W_1)(I-w_j)\varphi$$

$$\hat{\beta}_{OLS} = \left(X^T\left(I - w_j\right)\varphi X\right)^{-1} X^T y(I - \rho W_1)\left(I - w_j\right)\varphi \quad (3.9)$$

Pada persamaan (3.9) karena terdapat φ yang merupakan parameter yang mengandung *outlier*, dan φbersifat skalar, maka φ dapat dicari dengan memisalkan $\varphi_i = \psi_i$ sebagai fungsi *influence*, sehingga persamaan (3.9) dapat diubah menjadi:

$$\hat{\beta}_{OLS} = \left(X^T\left(1 - w_j\right)\psi_i X\right)^{-1} X^T y(I - \rho W_1)\left(1 - w_j\right)\psi_i \quad (3.10)$$

Menurut Draper dan Smith (1988), fungsi *influence* dari fungsi pembobot dinyatakan sebagai berikut:

$$\omega_i = \omega(\varepsilon_i^*) = \frac{\psi_i(\varepsilon_i^*)}{\varepsilon_i^*} \quad (3.11)$$

Dimana ε_i^* merupakan residual yang distandarisasi terhadap estimasi simpangan baku ($\hat{\sigma}$), maka diperoleh

$$\varepsilon_i{}^* = \frac{\varepsilon_i}{\hat{\sigma}} \quad (3.12)$$

Untuk mendapatkan nilai $\varepsilon_i{}^*$ maka terlebih dahulu menghitung *standard deviation residual* $\hat{\sigma}$. Menurut Marona, dkk (2006) nilai dari $\hat{\sigma}$ dapat diperoleh dengan cara sebagai berikut:

$$\hat{\sigma} = \frac{MAD(x)}{0{,}6745} \quad (3.13)$$

Di mana $\mathrm{MAD}(x) = med\{|x - med(x)|\}$ dan pemilihan konstanta 0.6745 membuat $\hat{\sigma}$ suatu *estimator* yang mendekati *unbias* dari σ untuk *n* besar dan *residual* berdistribusi normal (Montgomery, dkk, 2006).

Sehingga dari persamaan (3.12) di atas dapat diubah menjadi:

$$\varepsilon_i{}^* = \frac{y(I - \rho W_1) - X\beta}{\frac{MAD(x)}{0{,}6745}} \quad (3.14)$$

Berdasarkan persamaan (3.14), maka fungsi pembobot pada persamaan (3.11) dapat diubah menjadi:

$$\omega_i = \frac{\psi_i \frac{y(I-\rho W_1)-X\beta}{\frac{MAD(x)}{0,6745}}}{\frac{y(I-\rho W_1)-X\beta}{\frac{MAD(x)}{0,6745}}} \qquad (3.15)$$

Dari proses pembobotan pada persamaan (3.11) maka diharapkan diperoleh taksiran yang *unbias* karena fungsi *influence* telah distandarisasi, selain itu dari (3.11) dapat juga dinyatakan sebagai:

$$\psi_i(\varepsilon_i{}^*) = \frac{\omega(\varepsilon_i^*)}{\varepsilon_i^*}$$

atau

$$\psi_i = \frac{\omega_i}{\varepsilon_i^*}$$

Sehingga (3.10) dapat diubah menjadi:

$$\hat{\beta}_{OLS} = \left(X^T(1-w_j)\psi_i X\right)^{-1} X^T y(I-\rho W_1)(1-w_j)\psi_i$$

$$= \left(X^T(1-w_j)\frac{\omega_i}{\varepsilon_i^*} X\right)^{-1} X^T y(I-\rho W_1)(1-w_j)\frac{\omega_1}{\varepsilon_i^*}$$

$$= \left(\frac{1}{\varepsilon_i}\right)^{-1} \left(X^T(1-w_j)\omega_i X\right)^{-1} \frac{1}{\varepsilon_i} X^T y(I-\rho W_1)(1-w_j)\omega_i$$

$$= \varepsilon_i\left(X^T(1-w_j)\omega_i X\right)^{-1} \frac{1}{\varepsilon_i} X^T y(I-\rho W_1)(1-w_j)\omega_i$$

$$\hat{\beta}_{OLS} = \left(X^T(1-w_j)\omega_i X\right)^{-1} X^T y(I-\rho W_1)(1-w_j)\omega_i \qquad (3.16)$$

dengan ω_i adalah matrik pembobot yang berukuran *n x n* dengan elemen-elemen diagonal yang berisi pembobot ω_1, ω_2, ω_3, … ω_n. Persamaan tersebut dikenal dengan persamaan *Weighted Least Square* (WLS). Pada pembahasan ini fungsi pembobot yang digunakan adalah fungsi pembobot *Tukey Bisquare* sebagai berikut:

$$\omega_i = \begin{cases} \left[1 - \left(\frac{\varepsilon_i{}^*}{c}\right)^2\right]^2, & |u_i| < c \\ 0, & |u_i| \geq c \end{cases} \tag{3.17}$$

Dengan c adalah *tunning constant* yang besarnya $c = 4.685$ dan berfungsi sebagai pengatur pembobot pada *outlier* agar $\hat{\sigma}$ sebagai penduga yang mampu mendekati keadaan *unbias*.

Jika fungsi ψ tidak linier, maka estimasi parameter dapat diselesaikan dengan metode iterasi kuadrat terkecil terboboti yaitu dengan metode IRLS (*Iteratively Reweighted Least Square*) (Fox, 2002). Pada iterasi ini nilai $\omega_i$akan berubah nilainya di setiap iterasinya sehingga diperoleh, $\hat{\beta}^0, \hat{\beta}^1, \dots, \hat{\beta}^m$. Untuk parameter dengan m adalah jumlah iterasi yang akan mengestimasi, maka *estimator* awal $\hat{\beta}^0$ adalah

$$\hat{\beta}^0 = \left(X^T\left(1 - w_j\right)\omega_i^0 X\right)^{-1} X^T y(I - \rho W_1)\left(1 - w_j\right)\omega_i^0 \tag{3.18}$$

Dengan ω_i^0 adalah matriks pembobot pertama yang berukuran $n \times n$ yang berisi pembobot $\omega_1^0, \omega_2^0, \omega_3^0, \dots \omega_n^0$. Maka untuk estimator selanjutnya

$$\hat{\beta}^1 = \left(X^T\left(1 - w_j\right)\omega_i^0 X\right)^{-1} X^T y(I - \rho W_1)\left(1 - w_j\right)\omega_i^0 \tag{3.19}$$

Kemudian dihitung kembali pembobot dari ω_i^1, tetapi menggunakan $\hat{\beta}^1$ sebagai pengganti $\hat{\beta}^0$, sehingga didapatkan:

$$\omega_i^{*1} = \frac{\psi\left(\frac{y(I-\rho W_1) - X\hat{\beta}^1}{\hat{\sigma}}\right)}{\left(\frac{y(I-\rho W_1) - X\hat{\beta}^1}{\hat{\sigma}}\right)} \tag{3.20}$$

Sehingga diperoleh

$$\hat{\beta}^1 = \left(X^T\left(1-w_j\right)\omega_i^1 X\right)^{-1} X^T y(I-\rho W_1)\left(1-w_j\right)\omega_i^1 \tag{3.21}$$

Seterusnya hingga didapatkan:

$$\omega_i^{*m-1} = \frac{\psi\left(\frac{y(I-\rho W_1)-X\hat{\beta}^{m-1}}{\hat{\sigma}}\right)}{\left(\frac{y(I-\rho W_1)-X\hat{\beta}^{m-1}}{\hat{\sigma}}\right)} \tag{3.22}$$

Dari persamaan di atas diperoleh:

$$\hat{\beta}^m = \left(X^T\left(1-w_j\right)\omega_i^{m-1} X\right)^{-1} X^T y(I-\rho W_1)\left(1-w_j\right)\omega_i^{m-1} \tag{3.23}$$

Kemudian untuk ω_i^m pembobot yang diberikan, maka diperoleh estimator sebagai berikut:

$$\hat{\beta}^{m+1} = \left(X^T\left(1-w_j\right)\omega_i^m X\right)^{-1} X^T y(I-\rho W_1)\left(1-w_j\right)\omega_i^m \tag{3.24}$$

Perhitungan seperti tersebut dilakukan berulang sampai diperoleh estimator yang konvergen, yakni ketika selisih nilai $\hat{\beta}^{m+1}$ dan $\hat{\beta}^m$ mendekati 0, dengan m merupakan banyaknya iterasi.

Kemudian akan ditunjukkan estimator $\hat{\beta}^{m+1}$ adalah *unbias*, estimator $\hat{\beta}^{m+1}$ dikatakan *unbias* jika $E\left(\hat{\beta}^{m+1}\right) = \beta$.

$$E\left(\hat{\beta}^{m+1}\right) = E\left(X^T\left(1-w_j\right)\omega_i^m X\right)^{-1} X^T y(I-\rho W_1)\left(1-w_j\right)\omega_i^m$$

$$= E(\left(X^T\left(1-w_j\right)\omega_i^m X\right)^{-1} X^T\left(1-w_j\right)\omega_i^m) E(y(I-\rho W_1))$$

$$= \left(X^T\left(1-w_j\right)\omega_i^m X\right)^{-1} [(X^T\left(1-w_j\right)\omega_i^m)(X\beta)]$$

$$= \left(X^T\left(1-w_j\right)\omega_i^m X\right)^{-1} \left(X^T\left(1-w_j\right)\omega_i^m X\right)\beta$$

$$= I\beta$$

$$= \beta$$

Dari uraian tersebut maka terbukti bahwa $\hat{\beta}^{m+1}$ merupakan estimator *unbias.*

Sehingga algoritma estimasi parameter dengan metode Weighted Z algorithm dapat dituliskan sebagai berikut.

3.6 Algoritma Estimasi Parameter dengan Metode *Weighted Z Algorithm*

Algoritma yang diusulkan menggunakan 3 *input* parameter. y adalah model regresi *spatial lag*. φ adalah nilai fungsi objektif sebagai *outlier*. Dan β adalah vektor koefisien parameter regresi.

Sehingga algoritma dari estimasi parameter dengan metode *Weighted Z Algorithm* adalah sebagai berikut:

Input : y adalah model regresi *spatial lag*

φ adalah nilai fungsi objektif sebagai *outlier*

β adalah vektor koefisien parameter regresi

Output : $\hat{\beta}_m$ konvergen adalah parameter *spatial lag* yang mengandung *outlier* sebagai estimator yang konvergen

```
for(y=ρW₁y + Xβ + ε1){
    /*calculate the spatial lag with outlier/*
     y*(x_i)=Get spatial lag outlier
     y*(x_i)=y*φ
    /*calculate the neighborhood relationship*/
    g(x_i)=Get Neighbors
        NbrAvg(x_i)=0
  NbrAvg(x_i)= NbrAvg(x_i)+ y_i*w_j
  }
    /*calculate fungsi pembanding*/
    h(x_i)=Get fungsi pembanding
  h(x_i)= y*(x_i)− g(x_i)
    ε=Get error

/*calculate the sum square error*/
    ε^T ε=Get sum square error
  }
```

```
for (β, β̂_m konvergen){
        /*calculate β̂⁰*/
        β̂⁰=Get parameter β̂⁰(Diff)
        /*calculate the standarized β̂⁰(Diff)
        σ=Get Std(Diff)
        ε_i=Get error(Diff)
    σ= MAD(x)/0,6745
    ε_i=ε_i/σ
        /*calculate the weighted for spatial lag*/
        weight=Get weight
        }
    }
        β̂_m=get β̂_m konvergen
```

Untuk model regresi *spatial lag* (y), langkah pertama memberikan φ sebagai fungsi objektif sebagai *outlier* sehingga model regresi mengandung *outlier*$\left(y^*(x_i)\right)$. Selanjutnya menghitung *neighborhood function* $g(x_i)$ yang merupakan rata-rata terboboti. Setelah itu menghitung fungsi pembanding $h(x_i)$, yaitu selisih antara nilai atribut *non-spatial* dengan *neighborhood function* dengan persamaan:

$$h(x_i) = y^*(x_i) - g(x_i)$$

sehinga didapatkan *error* (ε). Kemudian, menghitung jumlah kuadrat *error* $(\varepsilon^T\varepsilon)$.

Untuk langkah selanjutnya, yaitu menghitung β sebagai vektor koefisien parameter regresi dengan mencari turunan pertama dari $\varepsilon^T\varepsilon$. Setelah itu, menghitung nilai σ dan nilai ε_i. Dilanjutkan dengan menghitung nilai pembobot (weight) dan menghitung parameter $\hat{\beta}_m$. dan melakukan langkah tersebut sampai didapatkan $\hat{\beta}_m$ yang konvergen.

3.7 Kajian *Outlier* dalam Al-Quran

Secara umum *outlier* dapat diartikan data yang tidak mengikuti pola umum model, atau dapat dikatakan sebagai data yang menyimpang. Dalam kehidupan, *outlier* dapat dikatakan sebagai sesuatu yang menyimpang dalam kebenaran.

Menyimpang dari kebenaran berhubungan dengan amal perbuatan manusia. Amal perbuatan manusia terbagi menjadi dua yaitu amal perbuatan baik dan amal perbuatan yang buruk. Pada al-Quran dijelaskan pada surat Fushilat/41:46, yaitu:

مَّنْ عَمِلَ صَٰلِحًا فَلِنَفْسِهِۦ ۖ وَمَنْ أَسَآءَ فَعَلَيْهَا ۗ وَمَا رَبُّكَ بِظَلَّٰمٍ لِّلْعَبِيدِ ﴿٤٦﴾

"Barangsiapa yang mengerjakan amal shaleh maka (pahalanya) untuk dirinya sendiri dan barang siapa mengerjakan perbuatan jahat, maka (dosanya) untuk dirinya sendiri; dan sekali-kali tidaklah Rabb-mu menganiaya hamba-hambaNya. (QS. Fushilat/41:46).

Pada al-Quran surat Fushilat/41:46 ini dijelaskan *barang siapa mengerjakan amal saleh* yaitu amal yang diperintahkan Allah dan Rasul-Nya *maka pahala dan* manfaatnya *untuk dirinya sendiri dan barang siapa berbuat jahat maka* (dosa dan hukumannya) *menjadi tanggungan dirinya sendiri* dalam ayat ini terdapat dorongan untuk mengerjakan kebaikan dan meninggalkan keburukan, adanya akibat dari amal yang dilakukan, bahwa seseorang tidak dapat memikul dosa orang lain, *dan tuhanmu sama sekali tidak menzalimi hamba-hamba-Nya* seperti memikul kepada hamba dosa-dosa diluar dosa mereka (Marwan bin Musa, Tafsir Hidayatul Insan).

Outlier merupakan salah satu faktor yang dapat mempengaruhi pendugaan parameter. Yang dapat mengakibatkan data tidak konsisten. Seperti halnya perbuatan manusia, semua amalan baik dan buruk akan mendapatkan balasannya. Hal ini di bahas dalam surat al-Mukmin/41:46, yaitu:

مَنْ عَمِلَ سَيِّئَةً فَلَا يُجْزَىٰٓ إِلَّا مِثْلَهَا ۖ وَمَنْ عَمِلَ صَٰلِحًا مِّن ذَكَرٍ أَوْ أُنثَىٰ وَهُوَ مُؤْمِنٌ
فَأُوْلَٰٓئِكَ يَدْخُلُونَ ٱلْجَنَّةَ يُرْزَقُونَ فِيهَا بِغَيْرِ حِسَابٍ ﴿٤٠﴾

"Barangsiapa mengerjakan perbuatan jahat, maka dia tidak akan dibalasi melainkan sebanding dengan kejahatan itu. Dan barangsiapa mengerjakan amal shaleh baik laki-laki maupun perempuan sedang ia dalam keadaan beriman. Maka mereka akan masuk surga, mereka diberi rezki di dalamnya tanpa hisab." (QS. Al-Mukmin/41:46).

Dari surat al-Mukmin/41:46 ini dijelaskan bahwa *barangsiapa mengerjakan perbuatan jahat, maka dia akan dibalas sebanding dengan kejahatan itu. Dan barangsiapa mengerjakan amal shaleh* baik yang berkait dengan hati, lisan maupun anggota badan, *baik laki-laki maupun perempuan sedangkan dia dalam keadaan beriman, maka mereka akan masuk surga, mereka diberi rezeki dalamnya tak terhingga*, Allah akan memberikan rezeki kepada mereka yang tidak dicapai oleh amal mereka (Marwan bin Musa, Tafsir Hidayatul Insan).

BAB IV

PENUTUP

5.1 Kesimpulan

Berdasarkan dari pembahasan yang telah dilakukan, maka dapat disimpulkan bahwa bentuk estimasi parameter regresi *spatial lag* yang mengandung *outlier* dengan metode *Weighted Z Algorithm* adalah sebagai berikut:

$$\hat{\beta}_{OLS} = \left(X^T(1-w_j)\omega_i X\right)^{-1} X^T y(I-\rho W_1)(1-w_j)\omega_i$$

dengan $\omega_i$adalah matrik pembobot yang berukuran *n x n* dengan elemen-elemen diagonal yang berisi pembobotω_1, ω_2, ω_3, … ω_n. Persamaan tersebut dikenal dengan persamaan *Weighted Least Square* (WLS). Dan terbukti estimator parameter regtresi *spatial lag* yang mengandung *outlier* dengan metode *weighted z algorithm* adalah estimator yang bersifat *unbias.*

5.2 Saran

Dari hasil penelitian ini ada beberapa saran yang dapat digunakan untuk penelitian selanjutnya antara lain adalah sebagai berikut:

1. Perlu dilakukan penelitian dengan metode lain, agar *outlier* pada model spasial dapat diselesaikan dengan lebih baik.
2. Perlu dilakukan dengan model *spasial* yang lain, contohnya dengan model spasial *Spatial Error Model* (SEM)

DAFTAR PUSTAKA

Abdussyakir. 2007. *Ketika Kyai Mengajar Matematika.* Malang: UIN PRESS.

Anselin, L.1988. *Spatial Econometrics: Method and Models*. Netherlands: Kluwer Academic Publisher.

Aziz, A. 2010. *Ekonometrika Teori dan Praktek Eksperimen dengan Matlab*. Malang: UIN MALIKI PRESS.

Chen, C.2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. *Paper Statistics and Data Analysis*, SUGI 27, Hal.265-267.

Draper, Norman dan Harry Smith. 1992.*AnalisisRegresiTerapan, EdisiKedua.* Jakarta: PT GramediaPustakaUtama.

Draper, N.R. dan Smith. H. 1998. *Applied Regression Analysis, Three Edition*. John Wiley and sons, Inc. New York.

Fox, J. 2002. *Robust Regression*. New York.

Goldberg, R. Lisa dan Alec N. Kercheval. 2002. *t-Statistics for Weighted Means with Application to Risk Factor Models.* Department of Mathematics Florida State University.

Harini, Sri danTurmudi. 2008.*Metode Statistika: Pendekatan Teoritis dan Aplikatif*. Malang: UIN Malang Press.

Hartoyo, E., Nugroho, Y., Bhirowo, A, danKhalil,B. 2010. *Modul Pelatihan Sistem Informasi Geografis (GIS) Tingkat Dasar.* Tropenbos International Indonesia Programme (TBI Indonesia).

Hasan, M. Iqbal. 2002. *Pokok-Pokok Materi Statistik 1(Statistik Deskriptif).* Jakarta: PT Bumi Aksara.

Kissling, W.D., Carl, G. 2007. *Spatial Autocorrelation and the Selection of Simultaneous Autoregressive Model, A Jurnal Macroecologi*. 17(1): 59-71.

Kou, Y., Lu, C,T, dan Chen, D. 2006. *Spatial Weighted Outlier Detection*. IEEE Computer Society

LeSage, James P.1998. Spatial econometrics.Department of Economics University of Toledo

Marwan bin Musa. *Tafsir Al-Quran Hidayatul Insan,* (Online), (http://www.tafsir.web.id). diakses 24 Juni 2016.

Marona, R.A, Martin. D & Yohai, V.J. 2006. *Robust Statistics: Theory and Methods*, England: John Wiley & Sons Ltd.

Montgomery, D.C., Peck, E.A., and Vining, G.G. 2006. *Introduction to Linier Regression Analysis*. 4th Ed. Canada: John Wiley & Sons.

Sembiring. 1995. *AnalisisRegresi*. Bandung: ITB.

Shekhar, S., Tien Lu, C, danZhang,P. 2002. *Detecting Graph Based Spatial Outlier*.University of Minnesota. Minneapolis.

Shekhar, S., Tien Lu, C, danZhang,P. 2003. *A Unified Approach to Detecting Spatial Outlier*.University of Minnesota. Minneapolis.

Shihab, M Quraish. 2003. *Tafsir Al-Misbah volume 14*. Jakarta: Lentera Hati.

Soemartini. 2007. *Outlier (Pencilan)*. Bandung: Universitas Padjajaran.

Yitnosumarto, Suntoyo. 1990. *Dasar-Dasar Ekonometrika*. Yogyakarta: BPFE.