PEMODELAN *ROBUST SPATIAL DURBIN MODEL* PADA DATA KASUS TUBERKULOSIS DI PROVINSI SULAWESI SELATAN DENGAN METODE MM-*ESTIMATOR*

SKRIPSI

UMMUL AULIYAH SYAM

H051191017

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

PEMODELAN ROBUST *SPATIAL DURBIN MODEL* PADA DATA KASUS TUBERKULOSIS DI PROVINSI SULAWESI SELATAN DENGAN METODE *MM-ESTIMATOR*

SKRIPSI

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

UMMUL AULIYAH SYAM

H051191017

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

DESEMBER 2022

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

Pemodelan Robust *Spatial Durbin Model* pada Data Kasus Tuberkulosis di Provinsi Sulawesi Selatan dengan Metode *MM-Estimator*

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 28 Desember 2022

Ummul Auliyah Syam

NIM H051191017

PEMODELAN ROBUST *SPATIAL DURBIN MODEL* PADA DATA KASUS TUBERKULOSIS DI PROVINSI SULAWESI SELATAN DENGAN METODE MM-*ESTIMATOR*

Disetujui Oleh:

Pembimbing Utama	Pembimbing Pertama
Siswanto, S.Si., M.Si.	**Dr. Nurtiti Sunusi, S.Si., M.Si.**
NIP. 19920107 201903 1 012	**NIP. 19720117 199703 2 002**

Ketua Program Studi

Dr. Nurtiti Sunusi, S.Si., M.Si.
NIP. 19720117 199703 2 002

Pada 28 Desember 2022

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama	: Ummul Auliyah Syam
NIM	: H051191017
Program Studi	: Statistika
Judul Skripsi	: Pemodelan *Robust Spatial Durbin Model* pada Data Kasus Tuberkulosis di Provinsi Sulawesi Selatan dengan Metode MM-*Estimator*

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Siswano, S.Si., M.Si. (.....)
2. Sekretaris : Dr.Nurtiti Sunusi, S.Si., M.Si. (.....)
3. Anggota : Andi Kresna Jaya, S.Si., M.Si. (.....)
4. Anggota : Dr. Dr. Georgina Maria Tinungki, M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 28 Desember 2022

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabaraktuh

Puji syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas segala limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan para sahabatnya. *Alhamdulillahirobbil'alamin*, berkat nikmat kemudahan dan pertolongan yang diberikan oleh Allah *Subhanahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi yang berjudul "**Pemodelan *Robust Spatial Durbin Model* pada Data Kasus Tuberkulosis di Provinsi Sulawesi Selatan dengan Metode MM-*Estimator*"** yang disusun sebagai salah satu syarat akademik untuk memperoleh gelar sarjana pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis menyadari bahwa dalam penyelesaian skripsi ini tidak lepas dari bantuan dan dorongan dari berbagai pihak yang senantiasa turut membantu dalam bentuk moril maupun materil sehingga dengan segala keterbatasan kemampuan dan pengetahuan, penulis dapat menyelesaikan skripsi ini. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang setulus-tulusnya serta penghargaan yang setinggi-tingginya kepada Ayahanda **Syamsul Jafar** dan Ibunda **Rismawati** yang telah memberikan dukungan penuh, pengorbanan luar biasa, limpahan cinta dan kasih sayang, kesabaran hati, serta dengan ikhlas telah menemani setiap langkah penulis dengan doa dan restu mulianya. Ucapan terima kasih juga penulis haturkan kepada Kakak tersayang **Nisa Anugerah Syam** yang senantiasa memberikan bantuan dan mendengarkan segala keluh kesah penulis. Terima kasih juga kepada adik-adik tersayang penulis **Khiran Abdullah Syamsul**, **Rabiatul Adawiah Syamsul**, dan **Aqilah Ulyah Syamsul** yang selalu menghibur dan memberi semangat kepada penulis dalam menyelesaikan skripsi ini, serta keluarga besar penulis, terima kasih atas doa mulia dan dukungannya selama ini.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan dan ketulusan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.

2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurtiti Sunusi S.Si., M.Si.**, selaku Ketua Departemen Statistika sekaligus Pembimbing Pertama penulis yang dengan penuh kesabaran telah memberikan arahan, dorongan semangat dan motivasi kepada penulis selama menjadi mahasiswa di Departemen Statistika.
4. **Bapak Siswanto, S.Si., M.Si.**, selaku Pembimbing Utama sekaligus Penasehat Akademik penulis yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya untuk senantiasa memberikan arahan, dorongan semangat, dan motivasi kepada penulis dari awal hingga selesainya penulisan tugas akhir ini.
5. **Bapak Andi Kresna Jaya, S.Si., M.Si.** dan **Ibu Dr. Dr. Georgina Maria Tinungki, M.Si.**, selaku Tim Penguji yang telah meluangkan waktu dalam memberikan motivasi serta kritikan yang membangun kepada penulis dalam penyempurnaan tugas akhir ini.
6. Segenap Dosen Pengajar dan Staf yang telah memberikan ilmu dan kemudahan kepada penulis dalam berbagai hal selama menempuh pendidikan sarjana di Departemen Statistika.
7. Sahabat tercinta penulis, **Alya Safira Irtiqa Miolo** yang telah menjadi sahabat dan saudara terbaik serta senantiasa memberikan semangat, mendengarkan segala keluh kesah penulis dalam hal apapun, dan menemani perjalanan suka dan duka penulis selama ini.
8. Sahabat terbaik penulis sejak kecil, **Azizah Nurul Waqiah**, **Sarmila**, dan **Nurul Eka Pratiwi**, terima kasih selalu mendengarkan keluh kesah penulis, memberikan semangat, dan menghibur penulis selama ini.
9. Sahabat terbaik penulis sejak di bangku SMA, **Andini Sulfitra**, **Melinnia Yulanda** dan **Ni'ma Haifa Amin**, serta keluarga besar **XII IPA 1** yang senantiasa memberi semangat dan menenani perjuangan pendidikan penulis.
10. Sahabat seperjuangan di Statistika 2019, **Muhammad Yusran**, **Wahyu Dwi Rahmawati**, **Vinaya Rifqi Anandari**, **Muh. Nur Iskandar Zulkarnain**, **Muhammad Rayhan Rifaldi**, **A. Ahmad Qeis Tenridapi**, **Fadhila Febriyanti Najamuddin**, **Fatiyah Utami**, **Diah Lestari**, **Alfiyah Salsa Dila Sabir** dan **Inda Mutmainna**. Terima kasih atas kebersamaan, kebahagiaan,

kesedihan, serta kebaikannya menjadi sosok guru bagi penulis. Terima kasih telah mengukir kenangan indah bersama penulis selama masa perkuliahan.

11. Sahabat berorganisasi penulis, **Aulia, Anita, Kak Sonya, Rosha, Andini,** dan **Paisa**. Terima kasih untuk dukungan, motivasi, canda dan tawa yang diberikan serta selalu menjadi pendengar yang baik bagi penulis.
12. Sepupu tercinta penulis, **Alivia Nurul Ilmi**, **Hani Sabaniah**, dan **Muhammad Alian Amnur**, terima kasih atas semangat dan bantuannya kepada penulis.
13. Kakak-kakak di Laboratorium Statistika, **Kak Samsir Aditya Ania**, **Kak Andi Isna Yunita**, **Kak Hedi Kuswanto**, **Kak Ratmila**, dan **Kak Alimatun Najiha**. Terima kasih atas bimbingan, dorongan semangat, dan nasehatnya kepada penulis selama semester 7 penulis. Terima kasih untuk segala canda dan tawa yang sangat menghibur penulis selama berada di lab.
14. Teman seperjuangan di **Statistika 2019**. Terima kasih atas ilmu, kebersamaan, suka dan duka dalam menjalani perkuliahan di Departemen Statistika. Terima kasih sudah menerima kehadiran penulis. Kalian hebat dan luar biasa.
15. **Keluarga Besar UKM Radio Kampus EBS FM Unhas**, terkhusus untuk **REKAMAN 2020**. Terima kasih telah menjadi "rumah kedua" bagi penlis, memberikan kesempatan kepada penulis untuk mendapatkan pengetahuan dan pengalaman yang sangat berkesan. Teruslah mengudara EBS-ku tercinta.
16. **Keluarga Besar Himastat FMIPA Unhas**, terima kasih atas ilmu, pengalaman, dan telah menjadi tempat belajar bagi penulis.
17. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih setinggi-tingginya untuk segala dukungan, partisipasi, dan apresiasi yang diberikan kepada penulis.

Penulis menyadari bahwa masih banyak kekurangan dalam penyusunan skripsi ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini dapat memberikan manfaat untuk berbagai pihak.

Makassar, 28 Desember 2022

Ummul Auliyah Syam

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama	: Ummul Auliyah Syam
NIM	: H051191017
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-eksklusif** (***Non-exclusive Royalty- Free Right***) atas tugas akhir saya yang berjudul:

"Pemodelan *Robust Spatial Durbin Model* pada Data Kasus Tuberkulosis di Provinsi Sulawesi Selatan dengan Metode MM-*Estimator*"

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar tanggal 28 Desember 2022.

Yang menyatakan,

(Ummul Auliyah Syam)

ABSTRAK

Regresi spasial merupakan model yang digunakan untuk mengetahui hubungan antara variabel respon dan prediktor dengan mempertimbangkan efek spasial. Apabila terdapat dependensi spasial pada variabel respon dan prediktor maka model yang dapat digunakan adalah model *Spatial Durbin Model* (SDM). Namun, dalam pemodelan regresi spasial terkadang ditemukan beberapa permasalahan, salah satunya adalah adanya pencilan spasial. Salah satu cara untuk mengatasi adanya pencilan pada model SDM adalah dengan menggunakan regresi *robust* yang disebut *Robust Spatial Durbin Model* (RSDM). Penelitian ini bertujuan untuk membentuk model RSDM dan mengidentifikasi faktor-faktor yang memengaruhi jumlah kasus TBC di Provinsi Sulawesi Selatan. Metode estimasi regresi *robust* yang digunakan adalah MM-*Estimator*. Penerapan MM-*Estimator* dilakukan dalam mengestimasi koefisien parameter regresi spasial yang resisten terhadap pencilan. Data yang digunakan dalam penelitian ini adalah data jumlah kasus tuberkulosis di Provinsi Sulawesi Selatan tahun 2020. Hasil analisis menunjukkan bahwa model RSDM dapat mengakomodir keberadaan pencilan dalam model regresi spasial. Hal ini ditunjukkan dengan adanya perubahan nilai estimasi koefisien parameter regresi RSDM yang dapat menurunkan nilai MSE menjadi 6461.734 dan meningkatkan nilai *adjusted* R^2 menjadi 99.52%. Faktor-faktor yang memengaruhi jumlah kasus TBC di Provinsi Sulawesi Selatan adalah kepadatan penduduk, persentase rumah tangga yang ber-PHBS, persentase imunisasi BCG dan persentase gizi buruk.

Kata Kunci: Regresi spasial, RSDM, Tuberkulosis, Pencilan, MM-*Estimator*.

ABSTRACT

Spatial regression is a model used to determine the relationship between response variables and predictors variables by considering spatial effects. If there is a spatial dependency on the response and predictor variables, the model that can be used is the Spatial Durbin Model (SDM). However, in spatial regression modeling, some problems are sometimes found, one of which is the presence of spatial outliers. One way to overcome the existence of outliers in the SDM model is to use a robust regression called the Robust Spatial Durbin Model (RSDM). The aim of this research is to obtain the RSDM model and identify the factors that influence the number of tuberculosis cases in South Sulawesi Province. The robust regression estimation method used is the MM-Estimator. The application of the MM-Estimator is carried out in estimating the coefficients of spatial regression parameters that are resistant to outliers. The data used in this study is data on the number of tuberculosis cases in South Sulawesi Province in 2020. The results of the analysis show that the RSDM model can accommodate the existence of outliers in the spatial regression model. This is indicated by changes in the estimated value of the RSDM regression parameter coefficient, which can reduce the MSE value to 6461.734 and increase the adjusted R^2 value to 99.52%. Factors that influence the number of Tuberculosis cases in South Sulawesi Province are population density, the percentage of households with a lively and healthy behaviour, the percentage of BCG immunization and the percentage of malnutrition.

Keywords: *Spatial Regression, RSDM, Tuberculosis, Outlier,* MM-*Estimator*.

DAFTAR ISI

HALAMAN SAMPUL ..... **i**
HALAMAN JUDUL ..... **ii**
LEMBAR PERNYATAAN KEOTENTIKAN ..... **iii**
HALAMAN PENGESAHAN ..... **v**
KATA PENGANTAR ..... **vi**
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI ..... **ix**
ABSTRAK ..... **x**
ABSTRACT ..... **xi**
DAFTAR ISI ..... **xii**
DAFTAR GAMBAR ..... **xv**
DAFTAR TABEL ..... **xvi**
DAFTAR LAMPIRAN ..... **xvii**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 4
BAB II TINJAUAN PUSTAKA **5**
2.1 Regresi Spasial 5
2.2 Matriks Pembobot Spasial 6
2.3 Identifikasi Efek Spasial 7
2.3.1 Heterogenitas Spasial (*Spatial Heterogeneity*) 7
2.3.2 Dependensi Spasial (*Spatial Dependency*) 7
2.4 *Spatial Durbin Model* 9
2.5 Pengujian Model Regresi 10
2.5.1 Uji Kecocokan Model 10
2.5.2 Uji Signifikansi Parameter 10
2.5.3 Uji Asumsi Regresi 11
2.6 Ukuran Kebaikan Model 13

2.7 Pencilan Spasial 14
2.8 Regresi *Robust* 15
2.8.1 *Robust* M-*Estimator* 15
2.8.2 *Robust* S-*Estimator* 16
2.8.3 *Robust* MM-*Estimator* 16
2.8.4 Fungsi Objektif 17
2.9 Tuberkulosis 17
BAB III METODOLOGI PENELITIAN 18
3.1 Data 18
3.2 Metode Analisis 18
BAB IV HASIL DAN PEMBAHASAN 22
4.1 Estimasi Parameter 22
4.1.1 Estimasi Parameter *Spatial Durbin Model* 22
4.1.2 Estimasi Parameter *Robust Spatial Durbin Model* 27
4.2 Eksplorasi Data 31
4.3 Matriks Pembobot Spasial 32
4.4 Efek Spasial 33
4.4.1 Efek Heterogenitas Spasial 33
4.4.2 Efek Dependensi Spasial 34
4.5 Pemodelan *Spatial Durbin Model* 35
4.6 Pengujian Asumsi Regresi Model SDM 37
4.6.1 Asumsi Normalitas 37
4.6.2 Asumsi Homoskedastisitas 37
4.6.3 Asumsi Autokorelasi 37
4.6.4 Uji Multikolinieritas 38
4.7 Pendeteksian Pencilan Spasial 38
4.8 Pemodelan *Robust Spatial Durbin Model* Menggunakan MM-*Estimator* .. 39
4.9 Uji Kecocokan Model Regresi RSDM 41
4.10 Uji Signifikansi Parameter Model RSDM 42
4.11 Ukuran Kebaikan Model 43
4.12 Interpretasi Model RSDM 43
BAB V KESIMPULAN DAN SARAN 46

5.1 Kesimpulan 46
5.2 Saran 47
DAFTAR PUSTAKA 48
LAMPIRAN 53

DAFTAR GAMBAR

Gambar 2.1. Ilustrasi *Moran's Scatterplot* 14
Gambar 4.1. Peta Sebaran Kasus Tuberkulosis di Sulawesi Selatan..... 31
Gambar 4.2. Plot Korelasi Antar Variabel 32
Gambar 4.3. *Moran's Scatterplot* 38

DAFTAR TABEL

Tabel 3.1. Variabel Respon dan Prediktor 18
Tabel 4.1. Pembobot Spasial IDW Terstandarisasi 33
Tabel 4.2. Pengujian Heterogenitas Spasial 33
Tabel 4.3. Hasil Pengujian Dependensi Spasial 34
Tabel 4.4. Estimasi Parameter Model SDM dengan Matriks Pembobot IDW..... 35
Tabel 4.5. Uji Normalitas Model SDM 37
Tabel 4.6. Uji Homoskedastisitas Model SDM..... 37
Tabel 4.7. Uji Homoskedastisitas Model SDM..... 37
Tabel 4.8. Uji Homoskedastisitas Model SDM..... 38
Tabel 4.9. Hasil Estimasi Parameter *Robust* S-*Estimator* 40
Tabel 4.10. Hasil Estimasi Parameter RSDM dengan MM-*Estimator* 41
Tabel 4.11. Uji Signifikansi Parameter Model RSDM..... 42
Tabel 4.12. Perbandingan Ukuran Kebaikan Model 43
.....

DAFTAR LAMPIRAN

Lampiran 1. Koefisien *Shapiro-Wilk*..... 54
Lampiran 2. Jarak *Euclidean* antar Lokasi..... 55
Lampiran 3. Matriks Pembobot Spasial (*Inverse Distance Weight*) 56
Lampiran 4. Pendeteksian Pencilan Berdasarkan Nilai LISA..... 57
Lampiran 5. Hasil Perhitungan Nilai Estimasi Seluruh Iterasi (S-*Estimator*) 58
Lampiran 6. Hasil Perhitungan Nilai Estimasi Seluruh Iterasi (MM-*Estimator*) 59

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis data spasial merupakan salah satu metode statistik untuk menganalisis data yang memuat informasi atribut dan lokasi geografis yang memiliki sistem koordinat (lintang dan bujur) (Paramita dkk., 2021). Data spasial dapat dibedakan menjadi tiga tipe berdasarkan jenis datanya, yaitu data geostatistik, data area, dan pola titik. Berdasarkan tipe data tersebut sehingga pemodelan data spasial juga dibedakan menjadi pemodelan dengan pendekatan titik dan area (Wardani dkk., 2018). Pemodelan pada data spasial perlu memperhatikan efek spasial, yaitu heterogenitas spasial dan dependensi spasial (Djuraidah, 2020).

Heterogenitas spasial menunjukkan adanya keragaman antar lokasi sedangkan dependensi spasial menunjukkan adanya keterkaitan antar lokasi objek penelitian (Agustina dkk., 2022). Jika terdapat efek heterogenitas maka diselesaikan menggunakan pendekatan titik, tetapi jika terdapat efek dependensi spasial maka diselesaikan menggunakan pendekatan area (Pratiwi dkk., 2018). Metode statistika yang dapat digunakan apabila ditemukan adanya efek dependensi spasial adalah regresi spasial. Regresi spasial merupakan analisis statistik yang mengukur hubungan antara variabel prediktor dan variabel respon dalam model regresi dengan mempertimbangkan efek spasial (Ramadhani dkk., 2020).

Salah satu pemodelan dengan menggunakan regresi spasial adalah model *Spatial Durbin Model* (SDM). Model SDM adalah model regresi spasial yang menunjukkan adanya keterkaitan spasial pada variabel respon dan prediktor (Pigitha, 2021). Namun, pada pemodelan regresi spasial seperti SDM terkadang ditemukan beberapa permasalahan, salah satunya adalah adanya pencilan spasial (*spatial outlier*). Pencilan spasial merupakan amatan yang memiliki nilai yang jauh dari nilai amatan lainnya dengan mempertimbangkan bobot spasial. Oleh karena itu, untuk mengatasi adanya pencilan pada model regresi spasial dikembangkan suatu metode regresi *robust* yang dapat menghasilkan estimasi model yang resisten terhadap pencilan (Rahman dan Widodo, 2018).

Beberapa metode estimasi yang dikembangkan dalam regresi *robust* antara lain adalah estimasi-M, *Least Median of Square* (LMS), *Least Trimmed Square*

(LTS), estimasi-S, dan estimasi-MM. Estimasi-M mempunyai efisiensi yang tinggi, tetapi nilai *breakdown point* 0. LMS, LTS dan estimasi-S mempunyai *breakdown* yang tinggi yaitu 0.5, tetapi efisiensinya sangat rendah. Sedangkan, estimasi-MM merupakan gabungan efisiensi tinggi dari estimasi-M dengan *breakdown point* tinggi dari estimasi-S (Shodiqin dkk., 2018). Regresi *robust* yang digunakan untuk mengatasi adanya pencilan spasial pada model SDM disebut *Robust Spatial Durbin Model* (RSDM).

Khofifah (2022) telah melakukan penelitian mengenai pemodelan tingkat pengangguran terbuka (TPT) di Provinsi Jawa Barat menggunakan RSDM dengan estimasi-M. Hasil penelitian menunjukkan RSDM merupakan model yang baik untuk menjelaskan TPT di Provinsi Jawa Barat. Selain itu, RSDM juga telah diterapkan pada Angka Harapan Hidup (AHH) Provinsi Jawa Tengah menggunakan estimasi-M yang menunjukkan hasil bahwa model yang lebih baik digunakan untuk menjelaskan AHH di Jawa Tengah adalah model RSDM (Hakim dkk., 2020). Penelitian tersebut berfokus pada penerapan RSDM dalam bidang ekonomi dan sosial sehingga pada penelitian ini akan dilakukan analisis di bidang lainnya, yaitu bidang kesehatan dalam hal ini adalah kasus penyakit menular Tuberkulosis (TBC).

TBC merupakan penyakit menular yang diakibatkan oleh kuman *mycobacterium tuberculosis*. Setiap tahunnya, 1,5 juta orang meninggal karena TBC sehingga menjadikannya sebagai pembunuh menular teratas di dunia (WHO, 2021). Sulawesi Selatan menduduki peringkat ke-9 dari 34 provinsi di Indonesia dengan jumlah kasus TBC tertinggi (Kemenkes, 2021). Jumlah penderita TBC per kabupaten/kota di Sulawesi Selatan tahun 2020 sebanyak 18.863 kasus dengan jumlah BTA+ sebesar 11.476 orang (60.83%) yang terdaftar dan diobati (Dinkes Sulsel, 2021). Salah satu cara untuk menurunkan jumlah kasus TBC di Sulawesi Selatan adalah dengan mengetahui faktor-faktor yang diperkirakan dapat memengaruhi jumlah kasus tersebut.

Beberapa faktor yang dapat memengaruhi jumlah kasus TBC dan mempermudah penularannya adalah kepadatan penduduk, rumah tangga yang berperilaku hidup bersih dan sehat, dan keluarga miskin (Hartanto dkk., 2019). Selain itu, penelitian lainnya juga menyebutkan bahwa jumlah kasus TBC juga

dipengaruhi oleh status gizi dan pemberian vaksinasi BCG (Sjahriani dan Sari, 2018; Widyastuti dkk., 2021). Faktor-faktor tersebut juga memungkinkan adanya keragaman jumlah kasus TBC di setiap kabupaten/kota. Berdasarkan data yang dihimpun oleh Dinas Kesehatan Provinsi Sulawesi Selatan (2021), Kota Makassar mempunyai penderita TBC terbanyak sebesar 5.421 penderita, disusul Kabupaten Gowa sebanyak 1.810 penderita, kemudian Kabupaten Bone sebanyak 1.288 penderita. Sementara Kabupaten Selayar mempunyai penderita TBC terkecil yaitu sebanyak 220 penderita.

Kondisi jumlah kasus TBC di Provinsi Sulawesi Selatan yang beragam tersebut dapat menyebabkan beberapa nilai memiliki nilai yang berbeda dari wilayah lainnya sehingga memungkinkan ditemukan pencilan spasial pada data. Berdasarkan hal tersebut, maka penelitian ini difokuskan dalam pembentukan model RSDM untuk mengatasi adanya pencilan pada model SDM dan mengidentifikasi faktor-faktor yang memengaruhi jumlah kasus TBC di Provinsi Sulawesi Selatan. Metode estimasi pada regresi *robust* yang digunakan adalah metode MM-*estimator*. Hasil penelitian ini dapat menjadi acuan bagi pemerintah maupun yayasan masyarakat dalam mewujudkan Sulawesi Selatan bebas TBC.

1.2 Rumusan Masalah

Rumusan masalah yang dibahas pada penelitian ini adalah sebagai berikut:

1. Bagaimana mengestimasi parameter pada model SDM dan RSDM?
2. Bagaimana perbandingan hasil model SDM dan RSDM pada kasus TBC di Provinsi Sulawesi Selatan tahun 2020?
3. Faktor apa saja yang memengaruhi jumlah kasus TBC di Provinsi Sulawesi Selatan?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Matriks pembobot spasial didasarkan pada informasi jarak antar lokasi (*distance*) menggunakan metode *inverse distance weight* (IDW).
2. Data yang digunakan adalah data kasus TBC di Provinsi Sulawesi Selatan tahun 2020.

3. Metode estimasi pada regresi *robust* yang digunakan adalah metode MM-*estimator* yang merupakan gabungan dari metode dengan nilai *breakdown* yang tinggi yaitu S-*estimator* dan efisiensi tinggi dari M-*estimator*.
4. Penentuan model regresi terbaik berdasarkan nilai *Mean Square Error* (MSE) dan *Adjusted* R^2.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mendapatkan *estimator* parameter model SDM dan RSDM
2. Membandingkan hasil model SDM dengan model RSDM pada kasus TBC di Provinsi Sulawesi Selatan tahun 2020.
3. Mengidentifikasi faktor-faktor yang memengaruhi jumlah kasus TBC di Provinsi Sulawesi Selatan.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut:

1. Memberikan informasi tentang pemodelan regresi SDM dengan pembobot invers jarak pada kasus TBC di Provinsi Sulawesi Selatan tahun 2020.
2. Memberikan informasi tentang analisis regresi RSDM dengan MM-*estimator* sebagai salah satu metode regresi *robust spatial* untuk menganalisis data spasial yang mengandung pencilan.
3. Memberikan inspirasi dan gagasan baru bagi pembaca dalam mengembangkan analisis regresi spasial dengan menggunakan model lain ataupun analisis data spasial dengan kasus lain yang ditemukan dalam kehidupan sehari-hari.
4. Sebagai sumber informasi bagi Dinas Kesehatan Provinsi Sulawesi selatan ataupun masyarakat umum terkait faktor-faktor yang memengaruhi jumlah kasus TBC di Provinsi Sulawesi Selatan sehingga dapat mengambil langkah antisipatif untuk menekan penyebaran penyakit TBC.

BAB II
TINJAUAN PUSTAKA

2.1 Regresi Spasial

Regresi spasial adalah metode regresi yang digunakan untuk tipe data spasial atau data yang memiliki efek lokasi (*spatial effect*). Dasar berkembangnya metode regresi spasial adalah metode regresi linier klasik (regresi linier berganda). Pengembangan itu berdasarkan adanya pengaruh tempat atau spasial pada data yang dianalisis (Anselin, 1988).

LeSage (1999) menjelaskan bahwa model umum regresi spasial dapat dituliskan sebagai berikut.

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim \mathbf{N}(0, \boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^{\mathbf{2}}\mathbf{I}_{\boldsymbol{n}})$$

dengan:

$\boldsymbol{y}$ = vektor variabel respon berukuran $n \times 1$

ρ = koefisien parameter spasial *lag* dari variabel respon

$\mathbf{W}$ = matriks pembobot spasial yang berukuran $n \times n$

$\mathbf{X}$ = matriks variabel prediktor berukuran $n \times (p + 1)$

$\boldsymbol{\beta}$ = vektor koefisien parameter regresi berukuran $(p + 1) \times 1$

λ = koefisien parameter spasial residual

$\mathbf{u}$ = vektor residual yang mempunyai efek spasial dengan ukuran $n \times 1$

$\boldsymbol{\varepsilon}$ = vektor residual dengan uran $n \times 1$

Berdasarkan persamaan model umum regresi spasial, menurut Anselin (1988) persamaan (2.1) dapat dibentuk beberapa model lain sebagai berikut:

Jika $\rho = 0$ dan $\lambda = 0$ maka disebut model regresi linier klasik dengan persamaan yang terbentuk adalah:

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

Jika $\rho \neq 0$ dan $\lambda = 0$ disebut regresi *Spatial Autoregressive Model* (SAR) dengan persamaan yang terbentuk adalah:

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

Jika $\rho = 0$ dan $\lambda \neq 0$ disebut regresi *Spatial Error Model* (SEM) dengan persamaan yang terbentuk adalah:

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

$$\mathbf{u} = \lambda\mathbf{Wu} + \boldsymbol{\varepsilon}$$

Jika $\rho \neq 0$ dan $\lambda \neq 0$ disebut *Spatial Autoregressive Moving Average* (SARMA) dengan persamaan yang terbentuk adalah:

$$\mathbf{y} = \boldsymbol{\rho}\mathbf{Wy} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda\mathbf{Wu} + \boldsymbol{\varepsilon}$$

2.2 Matriks Pembobot Spasial

Matriks pembobot spasial merupakan elemen penting dalam pemodelan spasial (Feng dkk., 2018). Matriks pembobot spasial ($\boldsymbol{W}_{n\times n}$) adalah matriks yang menjelaskan hubungan spasial antara n lokasi yang berisi konstanta tertentu. Setiap bobot spasial, w_{ij}, biasanya mencerminkan pengaruh spasial unit j pada unit i. Penentuan nilai bobot w_{ij} ditentukan dengan dua cara yaitu berdasarkan informasi jarak antar lokasi (*distance*) dan informasi ketetanggan (*contiguity*) (Smith, 2008).

Semakin dekat jarak kedua lokasi maka bobot yang diberikan semakin besar sesuai dengan hukum geografi pertama Tobler. Matriks yang digunakan berdasarkan pada jarak *centroid* (d_{ij}) antar unit spasial i dan j umumnya didefinisikan sebagai jarak *Euclidean*. Jarak *Euclidean* dapat diperoleh dengan rumus sebagai berikut:

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{2.1}$$

dengan u_i adalah koordinat garis lintang atau *latitude* pada lokasi ke-i; u_j adalah koordinat garis lintang atau *latitude* pada lokasi ke-j; v_i adalah koordinat garis bujur atau *longitude* pada lokasi ke-i; dan v_j adalah koordinat garis bujur atau *longitude* pada lokasi ke-j (Agustina dkk., 2022).

Nilai dari bobot lokasi *inverse distance* diperoleh dari hasil invers jarak sebenarnya yang kemudian distandarisasi. Nilai dari bobot lokasi *invers distance* dapat dihitung dengan rumus berikut:

$$w_{ij}^* = \frac{1}{d_{ij}}$$

dengan d_{ij} adalah jarak antar lokasi ke-i terhadap lokasi ke-j (Maulina dkk., 2019). Sedangkan nilai dari pembobot *inverse distance* yang terstandarisasi dapat dihitung menggunakan rumus berikut (Prasetya, 2017).

$$w_{ij} = \begin{cases} \frac{w_{ij}^*}{\sum_{j=1}^{n} w_{ij}^*}, & i \neq j \\ 0 & i = j \end{cases} \tag{2.2}$$

Matriks pembobot pada umumnya distandarisasi sehingga tiap barisnya akan berjumlah 1 yang menjamin bahwa bobot tiap daerah observasi akan berkisar antara 0-1 (Elhorst, 2014).

2.3 Identifikasi Efek Spasial

Efek spasial terbagi menjadi dua jenis yaitu heterogenitas spasial dan dependensi spasial (Yasin dkk., 2020).

2.3.1 Heterogenitas Spasial (*Spatial Heterogeneity*)

Efek heterogenitas adalah efek yang menunjukkan adanya keragaman antar lokasi. Jadi setiap lokasi mempunyai struktur dan parameter hubungan yang berbeda. Pengujian efek heterogenitas spasial dilakukan dengan uji *Breusch-Pagan* (BP). Hipotesis yang digunakan pada uji ini yaitu:

H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2$ (Homoskedastisitas)

H_1: minimal ada satu $\sigma_i^2 \neq \sigma_j^2$ dengan $i \neq j$ (Heteroskedastisitas)

Statistik uji yang digunakan pada uji *Breusch-Pagan* sebagai berikut:

$$BP = \frac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{f} \tag{2.3}$$

dengan

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$$

e_i^2 adalah residual untuk observasi ke-i dari hasil estimasi regresi dengan menggunakan OLS dan $\mathbf{Z}$ merupakan matriks variabel prediktor berukuran $n \times (p+1)$. Pengambilan keputusan pada uji ini adalah tolak H_0 jika nilai $BP > \chi^2_{(\alpha,p)}$ atau jika p-*value* $< \alpha$ dengan p adalah banyaknya variabel prediktor (Agustina dkk., 2022).

2.3.2 Dependensi Spasial (*Spatial Dependency*)

Dependensi spasial atau ketergantungan spasial terjadi akibat adanya ketergantungan dalam data wilayah (Agustina dkk., 2022). Ketergantungan antara

suatu lokasi dengan lokasi lain yang letaknya berdekatan menyebabkan terjadinya autokorelasi spasial (Grekousis, 2020). Salah satu uji yang digunakan dalam pemeriksaan keberadaan autokorelasi spasial atau dependensi spasial yaitu *Morans'I* (Gong dkk., 2020). Hipotesis yang digunakan dalam uji ini adalah sebagai berikut:

H_0: $I = 0$ (tidak terdapat dependensi spasial)

H_1: $I \neq 0$ (terdapat dependensi spasial)

Sedangkan indeks Moran (*Moran's I*) dapat dirumuskan sebagai berikut (Lee dan Wong, 2001):

$$I = \frac{n \sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.4}$$

dengan

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \; ; S_1 = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(w_{ij} + w_{ji}\right)^2 \; ;$$

$$S_2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji}\right)^2$$

Statistik uji yang digunakan yaitu:

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \approx N(0,1) \tag{2.5}$$

dengan

$$E(I) = -\frac{1}{n-1} \qquad Var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2-1) S_0^2} - [E(I)]^2$$

Keterangan:

I = Indeks Moran

n = banyaknya lokasi kejadian

x_i = nilai pengamatan pada lokasi ke-i

x_j = nilai pengamatan pada lokasi ke-j

$\bar{x}$ = rata-rata nilai pengamatan

w_{ij} = elemen pada pembobot terstandarisasi antara daerah i dan j

$E(I)$ = nilai ekspektasi dari I

$Var(I)$ = nilai variansi dari I

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika $|Z(I)| > Z_{\frac{\alpha}{2}}$ atau jika *p-value* $< \alpha$ yang berarti terdapat dependensi spasial dalam data.

Indeks Moran memiliki nilai yang berkisar antara -1 sampai 1. Nilai nol pada indeks moran menyatakan data tidak berkelompok berdasarkan wilayahnya. Indeks moran bernilai positif berarti terdapat autokorelasi spasial yang positif dan menunjukkan daerah-daerah yang berdekatan memiliki nilai efek spasial yang mirip. Sedangkan apabila indeks moran bernilai negatif maka terdapat autokorelasi spasial negatif, yang artinya daerah-daerah yang berdekatan memiliki nilai efek spasial yang berbeda (Kim dkk., 2022; Sun dkk., 2022).

2.4 *Spatial Durbin Model*

SDM merupakan model regresi spasial yang memiliki bentuk seperti *Spatial Autoregressive* (SAR) yang memiliki spasial *lag* pada variabel respon (y). Akan tetapi, SDM memiliki ciri khas yaitu adanya spasial *lag* pada variabel prediktor (X) (Anselin, 1988). Menurut (LeSage dan Pace, 2009) model SDM memiliki bentuk persamaan sebagai berikut:

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \alpha \mathbf{1}_{\boldsymbol{n}} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

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

atau dapat dituliskan sebagai berikut:

$$y_i = \rho \sum_{j=1}^{n} w_{ij} y_j + \alpha + \sum_{k=1}^{p} \beta_k x_{ik} + \sum_{k=1}^{p} \sum_{j=1}^{n} \theta_k \left(w_{ij} x_{kj} \right) + \varepsilon_i$$

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

dengan:

$\mathbf{y}$ = vektor variabel respon, berukuran $n \times 1$

$\mathbf{X}$ = matriks variabel prediktor, berukuran $n \times p$

$\boldsymbol{Z}$ = $[\mathbf{1}_{\boldsymbol{n}} \quad \boldsymbol{X} \quad \boldsymbol{W}\boldsymbol{X}]_{n \times (2p+1)}$

$\boldsymbol{\delta}$ = $\begin{bmatrix} \alpha \\ \boldsymbol{\beta} \\ \boldsymbol{\theta} \end{bmatrix}_{(2p+1) \times 1}$

ρ = koefisien *lag* spasial variabel respon (Y)

α = parameter konstan

$\boldsymbol{\beta}$ = vektor parameter regresi, berukuran $p \times 1$

$\boldsymbol{\theta}$ = vektor parameter *lag* spasial variabel prediktor berukuran $p \times 1$

$\mathbf{W}$ = matriks pembobot, berukuran $n \times n$

$\boldsymbol{\varepsilon}$ = vektor residual, berukuran $n \times 1$

$\mathbf{1}_{\boldsymbol{n}}$ = vektor yang berisi angka 1 berukuran $n \times 1$

2.5 Pengujian Model Regresi

2.5.1 Uji Kecocokan Model

Uji kecocokan model SDM menggunakan prosedur uji dengan hipotesis sebagai berikut (Mukrom dkk., 2021).

H_0: $\rho = \beta_j = \theta_j = 0$, dengan $j = 1, 2, \dots, p$

H_1: $\rho \neq 0$ atau setidaknya ada satu $\beta_j \neq 0$, $\theta_j \neq 0$, dengan $j = 1, 2, \dots, p$

Statistik Uji:

$$F_{hitung} = \frac{SSR/p}{SSE/(n-p-1)} \tag{2.6}$$

dengan SSR adalah jumlah kuadrat regresi, SSE adalah jumlah kuadrat residual, p adalah variabel prediktor, dan n adalah jumlah sampel. Kriteria keputusannya adalah H_0 ditolak jika nilai $F_{hitung} > F_{\alpha,p,n-p-1}$ atau $p\text{-}value < \alpha$.

2.5.2 Uji Signifikansi Parameter

Pengujian signifikansi parameter pemodelan spasial menggunakan uji *Wald* (Anselin, 1988). Pengujian parameter ρ, β_j, dan θ_j dilakukan sebagai berikut (Mukrom dkk., 2021).

a. Pengujian parameter ρ digunakan hipotesis dan statistik uji sebagai berikut:

$$\mathrm{H}_0: \rho = 0 \text{ ; } \mathrm{H}_1: \rho \neq 0$$

$$Wald_\rho = \frac{\hat{\rho}^2}{var(\hat{\rho})} \tag{2.7}$$

dengan $\hat{\rho}$ menyatakan penduga parameter ρ, $var(\hat{\rho})$ menyatakan elemen diagonal dari matriks varians yang berkorespondensi terhadap ρ.

b. Pengujian parameter β digunakan hipotesis dan statistik uji sebagai berikut:

$$\mathrm{H}_0: \beta_j = 0 \text{ ; } \mathrm{H}_1: \beta_j \neq 0 \text{ , } j = 1, 2, \dots, p$$

$$Wald_\beta = \frac{\hat{\beta}_j^{\,2}}{var(\hat{\beta}_j)} \tag{2.8}$$

$\hat{\beta}_j$ menyatakan penduga parameter β ke-j dan $var(\hat{\beta}_j)$ menyatakan elemen

diagonal dari matriks varians yang berkorespondensi terhadap β.

c. Pengujian parameter θ digunakan hipotesis dan statistik uji sebagai berikut:

$$\mathrm{H}_0: \theta_j = 0 \text{ ; } \mathrm{H}_1: \theta \neq 0, j = 1, 2, \dots, p$$

$$Wald_\rho = \frac{\hat{\theta}_j^{\,2}}{var(\hat{\theta}_j)} \tag{2.9}$$

$\hat{\theta}_j$ menyatakan penduga parameter θ ke-j dan $var(\hat{\theta}_j)$ menyatakan elemen diagonal dari matriks varians yang berkorespondensi terhadap θ.

Kriteria pengambilan keputusan adalah H_0 ditolak jika nilai $Wald > \chi^2_{(\alpha,1)}$.

2.5.3 Uji Asumsi Regresi

1. Uji Normalitas

Salah satu uji yang dapat digunakan adalah uji *Shapiro-Wilk* dengan hipotesis uji yang digunakan adalah sebagai berikut:

H_0: Residual berdistribusi normal

H_1: Residual tidak berdistribusi normal

Statistik uji *Shapiro-Wilk* dirumuskan sebagai berikut:

$$W = \frac{1}{s^2}\left[\sum_{i=1}^{n} a_i(e_{n-i+1} - e_i)\right]^2$$

dengan

$$s^2 = \sum_{i=1}^{n}(e_i - \bar{e})^2$$

Keterangan:

a_i = Koefisien *Shapiro-Wilk test* (Lampiran 1)

e_{n-i+1} = Residual ke $n - i + 1$

e_i = Residual ke-i

$\bar{e}$ = Rata-rata residual

s^2 = Varians residual

Jika nilai W hitung lebih besar dari nilai kritis pada tabel $W(n, \alpha)$ atau $W > W(n, \alpha)$ maka asumsi normalitas terpenuhi (Syamsudin dan Wachidah, 2020).

2. Uji Multikolinieritas

Salah satu cara untuk mengetahui adanya multikolinearitas dengan *Variance Inflation Factors* (VIF), yang didefinisikan dengan rumus (Montgomery dkk., 2012).

$$VIF = \frac{1}{\left(1 - R_j^2\right)}; \quad j = 1, 2, 3, \dots, p$$

dengan R_j^2 adalah nilai koefisien determinasi pada variabel ke-j. Nilai $VIF > 10$ menunjukkan multikolinieritas yang kuat.

3. Uji Homoskedastisitas

Asumsi homoskedastisitas disebut juga dengan asumsi kesamaan varians, artinya seluruh nilai-nilai dari variabel prediktor memiliki varians residual ε_i yang sama (Celik, 2017). Uji homoskedastisitas yang digunakan yaitu uji *Breusch-Pagan* dengan hipotesis (Mustari dan Zain, 2017).

H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2$ (Terjadi homoskedastisitas)

H_1: minimal ada satu $\sigma_i^2 \neq \sigma_j^2$ dengan $i \neq j$ (Tidak terjadi homoskedastisitas)

Nilai dari *Breusch-Pagan test* adalah sebagai berikut:

$$BP = \frac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{f} \tag{2.3}$$

dengan elemen vektor f adalah

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$$

dengan

e_i : residual observasi ke-i, $i = 1,2,..,n$

σ^2 : $\frac{\boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{e}}{n}$

$\boldsymbol{Z}$: matriks $n \times (p + 1)$ dari observasi dengan elemen kolom pertama yakni vektor satu

p : jumlah variabel prediktor

Pengambilan keputusan dalam uji *Breusch Pagan* adalah apabila nilai BP lebih besar dari $\chi^2_{(\alpha,p)}$ dan nilai *p-value* lebih kecil dari α maka H_0 ditolak yang artinya terjadi heteroskedastisitas.

4. Uji Autokorelasi

Salah satu cara untuk mendeteksi adanya autokorelasi dapat menggunakan uji *Durbin-Watson* dengan statistik uji (Nurdin dkk., 2018).

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

dengan

d : nilai *Durbin-Watson*

e_i : residual pada obeservasi ke-i

e_{i-1} : residual pada observasi $i - 1$

n : banyaknya observasi

Kriteria yang digunakan dalam pengujian adalah: jika $0 < d < dL$ atau $4 - dl < d < 4$, maka terjadi autokorelasi; jika $dU < d < 4 - dU$, maka tidak terjadi autokorelasi; jika $dL < d < dU$ atau $4 - dU < d < 4 - dL$, maka tidak ada keputusan (Taryono dkk., 2018).

2.6 Ukuran Kebaikan Model

Metode yang digunakan untuk menentukan ukuran kebaikan model dilakukan dengan menghitung nilai *Mean Square Error* (MSE). Perhitungan MSE dilakukan menggunakan persamaan berikut (Montgomery dkk., 2012):

$$MSE = \frac{SSE}{n - p - 1} \tag{2.10}$$

dengan SSE adalah jumlah kuadrat residual, n adalah jumlah sampel, dan p adalah banyaknya variabel prediktor. Jika nilai MSE semakin kecil hingga mendekati nol maka dapat dikatakan bahwa model regresi semakin baik.

Selain metode tersebut, untuk menentukan ukuran kebaikan model juga dapat menggunakan koefisien determinasi (R^2). Kecocokan model lebih baik jika R^2 semakin mendekati satu (Gujarati, 2003). Kebanyakan peneliti lebih suka menggunakan koefisien determinasi yang disesuaikan (*Adjusted* R^2) (Montgomery dkk., 2012), rumusnya sebagai berikut:

$$R_{adj,k}^2 = 1 - \frac{SSE/(n - p - 1)}{SSR/(n - 1)} \tag{2.11}$$

dengan SSR adalah jumlah kuadrat regresi, SSE adalah jumlah kuadrat residual, p adalah banyaknya variabel prediktor, dan n adalah jumlah sampel.

2.7 Pencilan Spasial

Pencilan adalah pengamatan yang tampak berbeda dengan pengamatan lainnya pada sekumpulan data yang ada (Barnett dan Lewis, 1994). Salah satu metode yang dapat digunakan untuk mendeteksi adanya pencilan spasial adalah *Moran's Scatterplot* (Atikah dkk., 2021). *Moran's Scatterplot* merupakan grafik yang menunjukkan hubungan antara nilai amatan pada suatu lokasi dengan rata-rata nilai amatan tetangganya. *Moran's Scatterplot* dibagi atas empat kuadran, pembagiannya adalah sebagai berikut:

Kuadran II *Low-High* (*Coldspot)*	Kuadran I *High-High*
Kuadran III *Low-Low*	Kuadran IV *High-Low* (*Hotspot*)

Gambar 2.1. Ilustrasi *Moran's Scatterplot*

Plot data yang berada di kuadran bagian kiri atas dan kanan bawah mengindikasikan adanya autokorelasi spasial negatif yang menunjukkan bahwa daerah pengamatan rendah dikelilingi oleh daerah pengamatan tinggi maupun sebaliknya. Oleh karena itu, dapat diidentifikasi bahwa titik-titik tersebut dikategorikan sebagai pencilan spasial atau titik-titik data yang berada pada kuadran II dan kuadran IV merupakan pencilan spasial (Shekhar dkk., 2003).

Selain dilihat dari grafik, deteksi pencilan spasial menggunakan *Moran's Scatterplot* juga dapat diidentifikasi secara matematik berdasarkan LISA (*Local Indicator of Spatial Autocorrelation*). LISA digunakan untuk mengidentifikasi autokorelasi secara lokal pada setiap lokasi. LISA dinyatakan dengan persamaan berikut (Habinuddin, 2021):

$$L_i = \left(Z_i \times \left(\sum_j \left(w_{ij} Z_j\right)\right)\right)$$

$$Z_i = \frac{x_i - \mu}{\sigma}\ ;\ Z_j = \frac{x_j - \mu}{\sigma};\ I_i = \sum_j \left(\mathrm{w}_{ij} Z_j\right)$$

dengan,

Z_i = nilai pengamatan yang distandarisasi pada lokasi ke-*i*

Z_j = nilai pengamatan yang distandarisasi pada lokasi ke-*j*

I_i = nilai pengamatan yang distandarisasi dikalikan dengan elemen matriks pembobot yang distandarisasi antara lokasi *i* dan *j*

w_{ij} = elemen matriks pembobot yang distandarisasi antara lokasi *i* dan *j*

μ = rata-rata dari nilai pengamatan

σ = standar deviasi dari nilai pengamatan

Jika nilai $L_i < 0$, maka dapat dikatakan bahwa titik tersebut termasuk pencilan spasial (Khofifah, 2022).

2.8 Regresi *Robust*

Regresi *robust* adalah metode regresi yang digunakan ketika residual tidak berdistribusi normal atau ada beberapa pencilan yang memengaruhi model (Draper dan Smith, 1998). Metode ini adalah alat yang penting untuk menganalisis data yang terkontaminasi oleh pencilan dan dapat memberikan hasil yang resisten akan kehadiran pencilan (Chen, 2002). Beberapa jenis estimasi parameter yang digunakan untuk mengestimasi regresi *robust*, yaitu *Least Median of Square* (LMS), *Least Trimmed Square* (LTS), estimasi *Maximum Likelihood Type* (M-*estimator*), estimasi *scale* (S-*estimator*), dan estimasi *Method of Moment* (MM-*estimator*) (Shodiqin dkk., 2018).

2.8.1 *Robust* M-*Estimator*

Regresi *robust* M-*estimator* pada umumnya dilakukan dengan meminimumkan fungsi objektif (Montgomery dkk., 2012):

$$min_\beta \sum_{i=1}^{n} \rho(e_i) = min_\beta \sum_{i=1}^{n} \rho\left(y_i - \sum_{j=0}^{p} x_{ij}\beta_j\right)$$

Untuk mendapatkan skala *invariant* pada *estimator* ini, biasanya dengan menyelesaikan persamaan (2.25) berikut:

$$min_\beta \sum_{i=1}^{n} \rho(u_i) = min_\beta \sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = min_\beta \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{p} x_{ij}\beta_j}{s}\right)$$

dengan s adalah skala estimasi *robust*. Estimasi s yang sering digunakan adalah

$$s = \frac{median|e_i - median(e_i)|}{0.6745}$$

Konstan 0.6745 membuat s mendekati *estimator* tak bias dari σ, dan residual berdistribusi normal.

2.8.2 *Robust* S-*Estimator*

Metode *robust* S merupakan metode *high breakdown value* yang diperkenalkan pertama kali oleh Rousseeuw pada tahun 1984. Selain meminimumkan varians dari residual, estimasi-S juga meminimumkan skala residual dalam estimasi-M (Abonazel dan Rabie, 2019). Estimasi-S didefinisikan sebagai berikut:

$$\hat{\beta}_s = \min \hat{\sigma}_s[e_1(\beta), e_2(\beta), \dots, e_n(\beta)]$$

dengan menentukan nilai *estimator* skala *robust* $(\hat{\sigma}_s)$ yang minimum dan memenuhi:

$$\min \sum_{i=1}^{n} \rho\left(\frac{y_i \sum_{j=0}^{p} X_{i,j}\beta_j}{\hat{\sigma}}\right)$$

dengan:

$$\hat{\sigma} = \sqrt{\frac{n\sum_{i=0}^{n}(e_i^2) - \left(\sum_{i=0}^{n} e_i\right)^2}{n(n-1)}}$$

Estimator $\hat{\beta}$ pada metode regresi *robust* estimasi S diperoleh dengan cara melakukan iterasi hingga diperoleh hasil yang konvergen. Proses ini dikenal sebagai MKT terboboti secara iterasi yang selanjutnya disebut sebagai *Iteratively Reweighted Least Square* (IRLS) (Fox dan Weisberg, 2010).

2.8.3 *Robust* MM-*Estimator*

MM-*estimator* merupakan gabungan dari metode estimasi yang mempunyai nilai *breakdown* tinggi (metode LTS-*estimator* atau S-*estimator*) dan metode M-*estimator* (Yohai, 1987). MM-*estimator* didefinisikan melalui tiga tahap. Tahap pertama yaitu menghitung estimasi parameter awal regresi dengan metode yang memiliki *breakdown* tinggi seperti S-*estimator* dan LTS-*estimator*. Tahap kedua, menghitung residual dan skala estimasi *robust* dengan menggunakan M-*estimator* (Sari dkk., 2020). Ketiga, menghitung estimasi parameter akhir dengan M-*estimator*. MM-*estimator* didefinisikan sebagai berikut:

$$\hat{\beta}_{MM} = \min \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}}\right) = \min \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{p} X_i\beta_j}{\hat{\sigma}}\right)$$

2.8.4 Fungsi Objektif

Fungsi yang digunakan untuk mencari fungsi pembobot pada regresi *robust* adalah fungsi objektif. Salah satu fungsi pembobot yang dapat digunakan adalah fungsi pembobot *Tukey Bisquare* sebagai berikut:

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & untuk\ |u_i| \leq c \\ 0 & untuk\ |u_i| > c \end{cases} \tag{2.31}$$

Nilai c pada metode *Tukey Bisquare* disebut *tuning constant*. Diketahui bahwa *tuning constant* untuk fungsi pembobot *Tukey Bisquare* pada analisis regresi *robust* S-*estimator* adalah $c = 1.547$ dan *tunning constant* untuk M-*Estimator* adalah $c = 4{,}685$ (Fox dan Weisberg, 2010).

2.9 Tuberkulosis

TBC merupakan penyakit menular yang diakibatkan oleh kuman *mycobacterium tuberculosis*. Penyakit merupakan salah satu dari 20 penyebab utama kematian di seluruh dunia dunia. Indonesia berada pada peringkat ke-3 dengan penderita TBC tertinggi di dunia setelah India dan China. Jumlah kematian akibat TBC secara global pada tahun 2020 sebesar 1,3 juta (WHO Global Tuberculosis Report, 2021). Sedangkan, angka notifikasi semua kasus TBC di Sulawesi Selatan menduduki peringkat ke-9 dari 34 provinsi di Indonesia (Kemenkes, 2021).

Oleh karena itu, TBC saat ini masih merupakan masalah kesehatan masyarakat baik di Sulawesi Selatan maupun Indonesia sehingga menjadi salah satu tujuan pembangunan kesehatan berkelanjutan (SDGs). Penanganan TBC sejak dini di Indonesia telah dilakukan melalui pemberian paket imunisasi *Bacillus Calmette-Guerin* (BCG) pada balita. Selain itu, tuberkulosis juga dapat disembuhkan dengan perbaikan sanitasi lingkungan dan melalui pengobatan (Dinkes Sulsel, 2021).