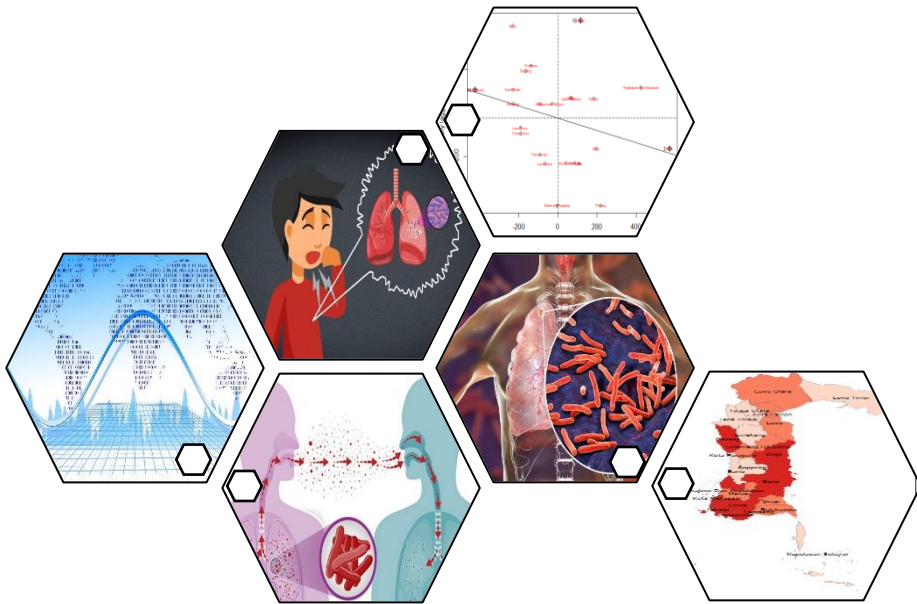


**PEMODELAN *ROBUST SPATIAL AUTOREGRESSIVE* DENGAN
METODE *MM-ESTIMATOR* PADA DATA KASUS TUBERKULOSIS
DI SULAWESI SELATAN**



**MARYANA MAHARANI
H051201046**



**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

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

**MARYANA MAHARANI
H051201046**



**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

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

MARYANA MAHARANI
H051201046

Skripsi

sebagai salah satu syarat untuk mencapai gelar sarjana

Program Studi Statistika

pada

**PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

SKRIPSI

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

MARYANA MAHARANI

H051201046

Skripsi,

telah dipertahankan di depan Panitia Ujian Sarjana Statistika pada 13 Agustus 2024
dan dinyatakan telah memenuhi syarat kelulusan
pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:
Pembimbing Tugas Akhir,

Siswanto, S.Si., M.Si.
NIP. 19920107 201903 1 012

Mengetahui:
Ketua Program Studi,

Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19770808 200501 2 002



PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Pemodelan *Robust Spatial Autoregressive* dengan Metode *MM-Estimator* Pada Data Kasus Tuberkulosis di Sulawesi Selatan" adalah benar karya saya dengan arahan dari pembimbing Siswanto, S.Si., M.Si. sebagai Pembimbing Utama. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 13 Agustus 2024



MARYANA MAHARANI
NIM H051201046

UCAPAN TERIMA KASIH

Alhamdulillahirobbil'alamin, puji syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas segala limpahan rahmat, hidayah, dan karunia-Nya beserta nikmat kesehatan dan kemudahan sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul "**Pemodelan *Robust Spatial Autoregressive* dengan Metode *MM-Estimator* Pada Data Kasus Tuberkulosis di Sulawesi Selatan**". Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan para sahabatnya.

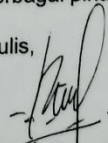
Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada **Bapak Siswanto, S.Si., M.Si.** selaku pembimbing utama yang telah memberikan arahan, bimbingan, dan motivasi dalam proses penyelesaian skripsi ini. Ucapan terima kasih juga penulis sampaikan kepada **Ibu Anisa, S.Si., M.Si.** dan **Bapak Dr. Nirwan, M.Si.** selaku tim penguji yang telah memberikan saran dan kritik yang membangun dalam penyempurnaan skripsi ini. Tak lupa pula penulis menyampaikan terima kasih kepada **Rektor Universitas Hasanuddin, Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam** dan **Kepala Departemen Statistika** beserta jajarannya beserta **Dosen-Dosen** karena telah memberikan fasilitas, ilmu, dan kemudahan kepada penulis selama menempuh pendidikan.

Ucapan terima kasih yang setulus-tulusnya penulis sampaikan kepada kedua orang tua tercinta penulis Ibunda **Hj. Juriyana** dan Ayahanda **H. Abd. Malik** (almarhum) atas segala kepercayaan, dukungan, inspirasi, cinta dan kasih sayang yang telah mengiringi setiap langkah penulis dalam setiap proses yang penulis alami hingga saat ini dengan doa dan restunya. Terima kasih juga kepada Bapak **Umar** yang telah menjadi ayah sambung bagi penulis yang telah memberikan kasih sayang dan dukungan yang sangat berarti. Tak lupa, penulis menyampaikan terima kasih juga kepada adik-adik tersayang penulis dan keluarga besar atas segala dukungan dan perhatian yang diberikan kepada penulis.

Terima kasih kepada teman-teman **STATISTIKA 2020** atas kebersamaan dan dukungan yang telah diberikan selama menempuh pendidikan. Terima kasih juga kepada keluarga besar **Himastat FMIPA Unhas** khususnya **POIS20N** dan **KM FMIPA Unhas** khususnya **Pengurus BEM Periode 2023/2024** atas segala ilmu dan pengalaman berharga. Kepada **Shella, Opi, Time, Muntas, Pila, Tasya, Nayah, Azal, Uci, Liza, Kur, Razy, Azhar, Fadlan, Izzul, Sabil, Ayuni, Nahdah, Ayu, Linda, Heri, Faldi, Naje, Kak Juni, Kak Snuf** beserta sahabat/teman penulis di **GAU22IAN, Seperbimbingan, KKN 110 Posko Bonsal, Magang MakSel, The Ortols**, dan lainnya yang tidak sempat penulis sebutkan satu persatu terima kasih atas segala doa, dukungan, bantuan, momen kebersamaan, tempat berbagi cerita, teman diskusi dan pembelajaran yang akan senantiasa penulis ingat sampai kapan pun.

Penulis menyadari bahwa masih banyak terdapat kekurangan dalam penyusunan skripsi ini. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf. Semoga skripsi ini dapat bermanfaat bagi berbagai pihak.

Penulis,


Maryana Maharani

ABSTRAK

MARYANA MAHARANI. **Pemodelan *Robust Spatial Autoregressive* dengan Metode *MM-Estimator* Pada Data Kasus Tuberkulosis di Sulawesi Selatan** (dibimbing oleh Siswanto).

Latar belakang. Regresi spasial adalah pengembangan dari analisis regresi biasa dengan melibatkan efek spasial untuk memodelkan dan melihat seberapa besar pengaruh variabel independen terhadap variabel dependen. Efek spasial yang biasa terjadi adalah dependensi spasial. *Spasial Autoregressive Model* (SAR) adalah salah satu model regresi spasial yang memiliki efek spasial berupa dependensi spasial pada variabel dependen. Namun, pada pemodelan SAR terkadang ditemukan permasalahan yaitu adanya pencilan spasial sehingga digunakan regresi *robust* dengan *MM-Estimator* yang memiliki *breakdown point* tinggi dan efisiensi tinggi. *Robust Spatial Autoregressive* (RSAR) dengan *MM-Estimator* digunakan untuk memodelkan data spasial berupa kasus tuberkulosis di Sulawesi Selatan yang memiliki efek dependensi spasial pada variabel dependen dan terdapat pecilan spasial. **Tujuan.** Penelitian ini bertujuan untuk mendapatkan model SAR dan RSAR serta faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di Sulawesi Selatan tahun 2021. **Metode.** Penelitian ini menggunakan estimasi parameter model SAR dengan *Maximum Likelihood* (MLE) dan model RSAR dengan *MM-Estimator*. **Hasil.** Hasil estimasi parameter SAR diterapkan pada data sehingga menghasilkan model SAR dengan beberapa variabel yang signifikan. Model SAR tersebut dirobustkan menggunakan estimasi parameter dengan *MM-Estimator* sehingga menunjukkan jumlah kasus tuberkulosis dipengaruhi oleh kepadatan penduduk dan persentase gizi buruk sebesar 94,55% dan 5,45% dipengaruhi oleh faktor lain diluar model. **Kesimpulan.** Model RSAR dengan *MM-Estimator* yang diperoleh dari estimasi parameter lebih baik daripada model SAR dan dapat menunjukkan faktor-faktor yang berpengaruh secara signifikan terhadap jumlah tuberkulosis di Sulawesi Selatan tahun 2021.

Kata Kunci: Regresi Spasial; Regresi Robust; Dependensi Spasial; Pencilan; Tuberkulosis.

ABSTRACT

MARYANA MAHARANI. **Robust Spatial Autoregressive Modeling with MM-Estimator Method on Tuberculosis Case Data in South Sulawesi** (supervised by Siswanto)

Background. Spatial regression is a development of ordinary regression analysis by involving spatial effects to model and see how much influence the independent variable has on the dependent variable. The common spatial effects is spatial dependency. Spatial Autoregressive Model (SAR) is one of the spatial regression models that has a spatial effect in the form of spatial dependency on the dependent variable. However, in SAR modeling, problems are sometimes found, namely the presence of spatial outliers, so robust regression is used with MM-Estimator which has a high breakdown point and high efficiency. Robust Spatial Autoregressive (RSAR) with MM-Estimator is used to model spatial data in the form of tuberculosis cases in South Sulawesi which have a spatial dependency effect on the dependent variable and there are spatial outliers. **Aim.** This study aims to obtain SAR and RSAR models and factors that influence the number of tuberculosis cases in South Sulawesi in 2021. **Method.** This study uses parameter estimation of the SAR model with Maximum Likelihood (MLE) and the RSAR model with MM-Estimator. **Results.** The results of the SAR parameter estimation are applied to the data to produce a SAR model with several significant variables. The SAR model is refuted using parameter estimation with MM-Estimator, showing that the number of tuberculosis cases is influenced by population density and the percentage of malnutrition by 94,55% and 5,45% is influenced by other factors outside the model. **Conclusion.** The RSAR model with MM-Estimator obtained from parameter estimation is better than the SAR model and can show factors that significantly influence the number of tuberculosis in South Sulawesi in 2021.

Keywords: Spatial Regression; Robust Regression; Spatial Dependence; Outliers; Tuberculosis.

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Variabel dependen	Variabel yang nilainya dipengaruhi dengan variabel lain
Variabel independen	Variabel yang mempengaruhi nilai variabel dependen
Efek spasial	Pengaruh yang dimiliki lokasi geografis dan hubungan antara lokasi tersebut terhadap variabel atau fenomena yang sedang diamati
Heterogenitas spasial	Keragaman antar lokasi atau perbedaan karakteristik diberbagai lokasi pengamatan
Dependensi spasial	Ketergantungan antara lokasi yang satu dengan lokasi yang lain di sekitarnya
Pencilan	Amatan yang memiliki nilai yang jauh dari amatan lainnya
Regresi	Metode untuk menganalisis hubungan antara variabel dependen dan satu atau lebih variabel independen
<i>Lag</i>	Tingkat keterkaitan antara lokasi
<i>Lagrange multiplier</i>	Konsep yang digunakan sebagai dasar memilih model regresi spasial yang sesuai
Bobot	Nilai yang diberikan untuk mengukur besar pengaruh dari lokasi amatan
Model	Representasi statistik dari hubungan antara variabel-variabel
Parameter	Nilai atau koefisien yang mendeskripsikan model statistik atau populasi
Residual	Selisih antara nilai yang sebenarnya dari variabel dependen dengan nilai dugaan dari model
<i>Estimator</i>	Suatu fungsi atau metode yang digunakan untuk menduga nilai dari parameter
<i>Breakdown point</i>	Ketahanan atau kekuatan suatu estimator terhadap pencilan
<i>Tuning constant</i>	Nilai yang telah ditetapkan untuk menentukan tingkat kerobustan suatu pembobot
<i>Robust</i>	Kekuatan atau kemampuan metode atau model untuk mengatasi data yang tidak ideal

DAFTAR LAMBANG DAN SINGKATAN

Lambang/Singkatan	Arti dan Penjelasan
y	Vektor variabel dependen
ρ	Koefisien parameter lag dari variabel dependen
W	Matriks pembobot spasial
X	Matriks variabel independen
β	Vektor koefisien parameter regresi
λ	Koefisien parameter spasial residual
u	Vektor residual yang mempunyai efek spasial
ε	Vektor residual
w_{ij}^*	Elemen matriks <i>power distance weights</i> antara lokai ke- i dan j
w_{ij}	Elemen matriks <i>power distance weights</i> terstandarisasi antara lokai ke- i dan j
d_{ij}	Jarak terdekat antara lokasi dalam kilometer
n	Banyaknya lokasi amatan
p	Banyaknya variabel independen
m	Banyaknya iterasi
$\hat{\sigma}_S$	Skala robust <i>S-Estimator</i>
$\hat{\sigma}_M$	Skala robust <i>M-Estimator</i>
$\hat{\beta}_S$	Penduga parameter <i>S-Estimator</i>
$\hat{\beta}_{MM}$	Penduga parameter <i>MM-Estimator</i>
SSR	Jumlah kuadrat regresi
SSE	Jumlah kuadrat residual
SST	Jumlah kuadrat total
SAR	<i>Spatial Autoregressive</i>
SEM	<i>Spatial Error Model</i>
$SARMA$	<i>Spatial Autoregressive Moving Average</i>
$RSAR$	<i>Robust Spatial Autoregressive</i>
AIC	<i>Akaike's Information Criterion</i>
LM	<i>Lagrange Multiplier</i>
OLS	<i>Ordinary Least Square</i>
MLE	<i>Maximum Likelihood Estimation</i>
$S - Estimator$	Estimasi regresi robust <i>Scale</i>
$M - Estimator$	Estimasi regresi <i>robust Maximum Likelihood type</i>
$MM - Estimator$	Estimasi regresi <i>robust Method of Moment</i>
MKT	Metode Kuadrat Terkecil
$IRLS$	Metode <i>Iteratively Reweightes Least Square</i>

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
PERNYATAAN PENGAJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
UCAPAN TERIMA KASIH.....	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISTILAH	viii
DAFTAR LAMBANG DAN SINGKATAN.....	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Batasan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Teori	4
1.5.1 Regresi Spasial	4
1.5.2 Matriks Pembobot Spasial	4
1.5.3 Pengujian Efek Spasial	5
1.5.4 Pengujian <i>Lagrange Multiplier</i>	7
1.5.5 <i>Spatial Autoregressive Model</i>	8
1.5.6 Pengujian Model Regresi	8
1.5.7 Ukuran Kebaikan Model.....	11
1.5.8 Pendeteksian Pencilan Spasial.....	11
1.5.9 Regresi <i>Robust</i>	12
1.5.10 Tuberkulosis	13
BAB II METODE PENELITIAN	15
2.1 Data.....	15
2.2 Analisis Data	15

BAB III HASIL DAN PEMBAHASAN.....	19
3.1 Estimasi Parameter.....	19
3.1.1 Estimasi Parameter Model SAR.....	19
3.1.2 Estimasi Parameter Model RSAR.....	23
3.2 Eksplorasi Data.....	26
3.3 Matriks Pembobot Spasial.....	28
3.4 Efek Spasial.....	28
3.4.1 Efek Heterogenitas Spasial.....	29
3.4.2 Efek Dependensi Spasial.....	29
3.5 Uji <i>Lagrange Multiplier</i>	29
3.6 Pemodelan <i>Spatial Autoregressive Model</i>	30
3.7 Pengujian Asumsi Model SAR.....	31
3.7.1 Uji Normalitas.....	31
3.7.2 Uji Homoskedastisitas.....	32
3.8 Pendeteksian Pencilan Spasial.....	33
3.9 Pemodelan <i>Robust Spatial Autoregressive</i>	33
3.10 Uji Kecocokan dan Signifikansi Parameter Model RSAR.....	35
3.10.1 Uji Kecocokan Model.....	35
3.10.2 Uji Signifikansi Parameter.....	35
3.11 Evaluasi Keباikan Model.....	36
3.12 Interpretasi.....	36
BAB IV KESIMPULAN.....	39
4.1 Kesimpulan.....	39
4.2 Saran.....	39
DAFTAR PUSTAKA.....	41
LAMPIRAN.....	45

DAFTAR TABEL

Tabel	Halaman
1. Variabel Penelitian	15
2. Pendeteksian Multikolinearitas	27
3. Jarak Antar Lokasi Pengamatan (KM)	28
4. Matriks Power Distance Weights	28
5. Hasil Pengujian Heterogenitas Spasial	29
6. Hasil Pengujian Dependensi Spasial	29
7. Hasil Pengujian Lagrange Multiplier	30
8. Hasil Estimasi Parameter Model SAR	30
9. Hasil Estimasi Parameter Model SAR Akhir	31
10. Hasil Pengujian Normalitas	32
11. Hasil Pengujian Homoskedastisitas	32
12. Hasil Iterasi S-Estimator	34
13. Hasil Iterasi MM-Estimator	35
14. Hasil Uji Kecocokan Model RSAR	35
15. Hasil Uji Signifikansi Parameter	36
16. Hasil Evaluasi Kebaikan Model	36

DAFTAR GAMBAR

Gambar	Halaman
1. Ilustrasi <i>Moran's Scatterplot</i>	11
2. Peta Sebaran Kasus Tuberkulosis di Sulawesi Selatan	26
3. Plot Korelasi Antar Variabel	27
4. <i>Moran's Scatterplot</i>	33

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Penelitian.....	46
2. Perhitungan Jarak Antar Lokasi (KM).....	47
3. Perhitungan Matriks <i>Power Distance Weights</i>	48
4. Hasil Iterasi RSAR <i>S-Estimator</i>	49
5. Hasil Iterasi RSAR <i>MM-Estimator</i>	50
6. Riwayat Hidup Penulis.....	51

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis data spasial adalah suatu metode statistik untuk mempelajari pola hubungan spasial pada data yang memuat informasi atribut dan lokasi geografis yang memiliki sistem koordinat (lintang dan bujur) (Paramita dkk., 2021). Data spasial terdiri dari tiga tipe yaitu data titik, garis, dan area (Mahading dkk., 2020). Berdasarkan tipe data tersebut, pemodelan data spasial dapat dibedakan menjadi pemodelan dengan pendekatan titik dan area (Aulele dkk., 2021). Memodelkan data spasial perlu memperhatikan efek spasial yaitu heterogenitas dan dependensi spasial (Yasin dkk., 2020).

Heterogenitas spasial terjadi apabila terdapat keragaman antar lokasi, sedangkan dependensi spasial terjadi apabila terdapat ketergantungan antara observasi di lokasi yang satu dengan lokasi yang lainnya (Mahdy, 2020). Jika terdapat heterogenitas spasial maka diselesaikan dengan menggunakan pendekatan titik, sedangkan jika terdapat dependensi spasial maka diselesaikan dengan menggunakan pendekatan area (Pratiwi dkk., 2018). Metode statistik yang digunakan dalam memodelkan data spasial adalah analisis regresi spasial (Wardani, 2018). Analisis regresi spasial adalah pengembangan dari analisis regresi biasa dengan melibatkan efek spasial suatu lokasi untuk memodelkan dan melihat seberapa besar pengaruh variabel independen terhadap variabel dependen. Salah satu pemodelan dengan menggunakan analisis regresi spasial adalah model *Spatial Autoregressive*.

Spatial Autoregressive Model (SAR) adalah model spasial yang memiliki efek spasial berupa dependensi pada variabel dependen. Namun, pada pemodelan regresi spasial seperti SAR terkadang ditemukan permasalahan, salah satunya adalah pencilan spasial (*spatial outlier*). Observasi yang memiliki nilai yang relatif sangat jauh dari observasi yang lainnya dan mempertimbangkan bobot spasial disebut *spatial outlier* (Musyarofah dkk., 2020). Hal tersebut jika tidak diatasi maka akan mempengaruhi nilai koefisien estimasi parameter dan membuat model tidak dapat diandalkan. Oleh karena itu, untuk mengatasi *outlier* dibutuhkan metode yang resisten terhadap *outlier* yaitu regresi *robust*.

Regresi *robust* adalah salah satu metode regresi untuk menganalisis data yang dipengaruhi adanya *outlier* sehingga model yang dihasilkan resisten terhadap *outlier*. Terdapat beberapa metode estimasi parameter pada regresi *robust* yaitu estimasi *Least Trimmed Square* (LTS), *Scale* (S), *Maximum Likelihood type* (M), dan *Method of Moment* (MM) (Rositawati & Fitri, 2022). Menurut Seprina dkk. (2023), *MM-Estimator* adalah estimasi yang sangat baik digunakan untuk mengestimasi parameter pada model regresi *robust* karena bersifat resisten dan memiliki efisiensi yang lebih besar. Hal ini dikarenakan *MM-Estimator* merupakan penggabungan antara *S-Estimator* yang memiliki *breakdown point* tinggi yaitu 0.5 tetapi memiliki efisiensi yang rendah dan *M-Estimator* yang memiliki *breakdown*

point rendah yaitu 0 tetapi memiliki efisiensi yang tinggi. Salah satu model regresi *robust* yang digunakan untuk mengatasi adanya *spatial outlier* adalah *Robust Spatial Autoregressive* (RSAR). Ini adalah metode yang digunakan untuk memodelkan data spasial yang memiliki efek dependensi spasial pada variabel dependen dan terdapat *outlier spatial* (Atikah dkk., 2021).

Robust Spatial Autoregressive menunjukkan hasil model terbaik pada data lokasi pendapatan pemerintah di Jawa tahun 2017 dengan menggunakan *S-Estimator* (Mastuti dkk., 2019). Selain itu, model RSAR dengan *S-Estimator* juga digunakan untuk memodelkan angka harapan hidup provinsi Jawa Timur yang menghasilkan model terbaik dibandingkan dengan SAR (Musyarofah dkk., 2020). Pemodelan RSAR juga dilakukan pada data harapan hidup pembangunan di Jawa Tengah tetapi menggunakan *M-Estimator* (Yasin dkk., 2020) dan Halilla (2022) juga melakukan pemodelan serupa tetapi pada faktor-faktor yang mempengaruhi kemiskinan di Provinsi Jawa Timur. Model terbaik yang menunjukkan adanya efek spasial pada variabel dependen dan terdapat *outlier spatial* adalah RSAR dengan *M-Estimator* yang mana telah dibuktikan oleh Abrari dkk. (2023) pada data gizi buruk balita di Indonesia dan Tho dkk. (2023) pada dataset *Nielsen Local Television View AS* dan hibah federal AS. Penelitian-penelitian tersebut rata-rata berfokus pada kasus sosial ekonomi yang berada di Jawa sehingga penelitian ini akan dilakukan pada kasus kesehatan yaitu Tuberkulosis yang terjadi di Sulawesi Selatan dengan menggunakan estimator yang berbeda.

Tuberkulosis (TBC) merupakan suatu penyakit menular yang paling sering menyerang paru-paru dan disebabkan oleh kuman *Mycobacterium Tuberculosis*. Ini menyebar melalui udara ketika orang yang terinfeksi batuk, bersin, atau meludah (WHO, 2023). Tahun 2021, Indonesia berada pada urutan ke-2 penderita TBC tertinggi di dunia setelah India dengan proporsi kasus baru sebesar 13% dibandingkan seluruh kasus di dunia (WHO, Global Tuberculosis Report, 2022).

Menurut data Kementerian Kesehatan (Kemenkes) RI pada tahun 2021, kasus TBC di Sulawesi Selatan ada sebanyak 14.303 kasus per kabupaten/kota yang mana berada pada urutan ke-7 dari 34 provinsi di Indonesia dengan jumlah kasus tertinggi. Untuk menurunkan jumlah kasus TBC di Sulawesi Selatan dapat dilakukan dengan mengetahui faktor-faktor yang mempengaruhi jumlah kasus tersebut. Adapun faktor-faktor yang mempengaruhi jumlah kasus TBC dan penyebarannya adalah keluarga miskin, kepadatan penduduk, dan sanitasi layak (Hartanto dkk., 2019). Selain itu, penelitian yang dilakukan oleh Syam dkk. (2024) menunjukkan faktor yang mempengaruhi jumlah kasus TBC dan penyebarannya adalah kepadatan penduduk, persentase rumah tangga yang berperilaku hidup bersih dan sehat, persentase imunisasi BCG (*Bacillus Calmette-Guerin*), dan persentase gizi buruk.

Berdasarkan data dari Dinas Kesehatan Provinsi Sulawesi Selatan (2022), Kota Makassar berada pada urutan pertama di Sulawesi Selatan dengan kasus tuberkulosis tertinggi sebesar 3.750 kasus, disusul Kabupaten Gowa sebanyak 1.270 kasus, dan Kabupaten Bone sebanyak 980 kasus. Sementara Kabupaten Tana Toraja merupakan kabupaten dengan kasus tuberkulosis terendah yaitu

sebanyak 154 kasus. Hal ini menunjukkan kondisi jumlah kasus TBC di Sulawesi Selatan beragam yang memungkinkan terdapat dependensi spasial dan *outlier spatial* pada data. Oleh karena itu, penelitian ini difokuskan untuk membentuk model RSAR dengan *MM-Estimator* dan mengetahui faktor-faktor yang berpengaruh secara signifikan pada data kasus TBC di Sulawesi Selatan tahun 2021. Hasil penelitian ini dapat menjadi referensi dan acuan bagi pemerintah, instansi atau pun yayasan kesehatan dalam menangani dan menentukan kebijakan untuk mengurangi kasus TBC di Indonesia khususnya Sulawesi Selatan.

1.2 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut.

1. Matriks pembobot spasial yang digunakan berdasarkan informasi jarak sesungguhnya yang dihitung melalui *Google Maps* dengan satuan kilometer dengan *power distance weights*.
2. Fungsi pembobot yang digunakan pada regresi robust adalah *Tukey Bisquare*.
3. Data yang digunakan adalah data kasus TBC di Sulawesi Selatan tahun 2021.
4. Penentuan model regresi terbaik berdasarkan nilai AIC dan R_{adj}^2 .

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut.

1. Mendapatkan model SAR dan RSAR.
2. Mengidentifikasi faktor-faktor yang mempengaruhi jumlah kasus TBC di Sulawesi Selatan tahun 2021.

1.4 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut.

1. Memberikan informasi tentang pemodelan RSAR dengan *MM-Estimator* menggunakan pembobot jarak pada kasus TBC di Sulawesi Selatan tahun 2021.
2. Memberikan pengetahuan baru bagi pembaca dalam mengembangkan analisis regresi spasial dengan menggunakan model RSAR pada kasus di berbagai bidang.
3. Memberikan informasi kepada instansi atau pun Yayasan seperti Dinas Kesehatan Provinsi Sulawesi Selatan dan masyarakat umum tentang faktor-faktor yang mempengaruhi jumlah kasus TBC di Sulawesi Selatan tahun 2021 sehingga dapat menjadi referensi dan acuan dalam menangani dan menentukan kebijakan untuk mengurangi kasus dan penyebaran TBC di Indonesia.

1.5 Teori

1.5.1 Regresi Spasial

Regresi spasial adalah metode yang digunakan untuk tipe data spasial atau data yang memiliki efek lokasi (*spatial effect*). Metode ini merupakan metode yang dikembangkan dari metode regresi linier klasik (regresi linier berganda). Pengembangan tersebut berdasarkan adanya pengaruh tempat atau spasial pada data yang dianalisis (Anselin, 1988).

Model umum regresi spasial yang dijelaskan Lesage (1999) dapat dituliskan

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim \mathbf{N}(0, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_n) \end{aligned} \quad (1)$$

dengan:

- \mathbf{y} : vektor variabel dependen berukuran $n \times 1$
- ρ : koefisien parameter spasial *lag* dari variabel dependen
- \mathbf{W} : matriks pembobot spasial berukuran $n \times n$
- \mathbf{X} : matriks variabel independen 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 berukuran $n \times 1$
- $\boldsymbol{\varepsilon}$: vektor residual berukuran $n \times 1$

Berdasarkan model umum regresi spasial, persamaan (1) dapat dibentuk beberapa model lain yaitu:

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

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

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

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3)$$

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

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \end{aligned} \quad (4)$$

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

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \end{aligned} \quad (5)$$

1.5.2 Matriks Pembobot Spasial

Matriks pembobot spasial sangat penting dalam pemodelan menggunakan regresi spasial (Ichsan dkk., 2022). Matriks pembobot spasial ($\mathbf{W}_{n \times n}$) adalah matriks yang menggambarkan kedekatan suatu lokasi dengan lokasi lainnya yang dapat menyatakan hubungan spasial antara n lokasi amatan (Abrari dkk., 2023). Matriks

pembobot spasial berisi konstanta tertentu yang digunakan sebagai bobot spasial (w_{ij}) antar lokasi yang diamati berdasarkan hubungan ketetanggaan antar lokasi (Nasir dkk., 2020). Setiap bobot spasial (w_{ij}) biasanya menggambarkan pengaruh spasial unit j dan unit i . Pembentukan matriks pembobot spasial (w_{ij}) dapat dilakukan dengan dua cara yaitu *contiguity* (ketetanggaan) dan *distance* (jarak) (Kosfeld, 2006).

Semakin dekat jarak dua lokasi semakin besar pengaruhnya satu sama lain sehingga bobot yang diberikan semakin besar. Hal ini sesuai hukum geografi pertama Tobler. Matriks yang digunakan berdasarkan jarak sesungguhnya antar lokasi dengan perhitungan jarak terdekat dalam satuan kilometer. Salah satu cara dalam menentukan matriks bobot berdasarkan jarak adalah dengan *power distance weights*. Nilai dari bobot lokasi berdasarkan *power distance weights matrix* dapat dihitung menggunakan persamaan (6).

$$w_{ij}^* = d_{ij}^{-1} \quad (6)$$

dengan d_{ij} adalah jarak antar lokasi ke- i terhadap lokasi ke- j (Amelia, 2021). Pada umumnya matriks pembobot akan distandarisasikan sehingga tiap barisnya akan berjumlah 1, hal ini menjamin bahwa bobot tiap lokasi observasi akan berkisar antar 0-1 (Elhorst, 2014). Nilai matriks pembobot *power distance* yang terstandarisasi dapat dihitung menggunakan persamaan (7) (Tho dkk., 2023).

$$w_{ij} = \begin{cases} \frac{w_{ij}^*}{\sum_{j=1}^n w_{ij}^*}, & i \neq j \\ 0, & i = j \end{cases} \quad (7)$$

1.5.3 Pengujian Efek Spasial

Untuk mengetahui ada atau tidaknya efek spasial dalam suatu data dilakukan pengujian efek spasial. Uji efek spasial terdiri dari dua jenis yaitu heterogenitas dan dependensi spasial (Wardhani & Yanti, 2021).

1. Uji Dependensi Spasial

Dependensi spasial terjadi akibat adanya ketergantungan dalam data wilayah yang mana nilai atribut pada lokasi tertentu bergantung pada nilai atribut lokasi lain yang letaknya berdekatan atau bertetanggaan (Agustina dkk., 2022). Suatu lokasi yang mengalami ketergantungan dengan lokasi lain yang letaknya berdekatan menyebabkan terjadi autokorelasi spasial (Grekousis, 2020). Salah satu uji yang digunakan dalam mengidentifikasi adanya autokorelasi spasial atau dependensi spasial yaitu *Morans'I* (Gong dkk., 2020). Hipotesis yang digunakan dalam uji ini adalah

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

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

Indeks moran (*Morans'I*) dapat dihitung dengan persamaan (8) (Lee & Wong, 2001):

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (8)$$

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

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \approx N(0,1) \quad (9)$$

dengan,

$$E(I) = -\frac{1}{n-1}; \quad Var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - [E(I)]^2; \quad S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2; \quad S_2 = \sum_{i=1}^n (\sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji})^2$$

Keterangan:

- I : Indeks Moran
 n : banyaknya lokasi amatan
 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 matriks pembobot spasial terstandarisasi antara lokasi 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 $p - value < \alpha$ yang artinya terdapat dependensi spasial pada data. Nilai *Morans'I* berada diantara -1 dan 1 yang mana nilai -1 menunjukkan adanya autokorelasi negatif dan 1 menunjukkan adanya autokorelasi positif (Atikah dkk., 2021). Adapun kriteria indeks I menurut Paradis (2010) yaitu:

- Jika $I > I_0$ maka terjadi autokorelasi positif, artinya lokasi-lokasi yang berdekatan memiliki nilai efek spasial yang serupa.
- Jika $I = I_0$ maka tidak terjadi autokorelasi, artinya data tidak berkelompok berdasarkan wilayahnya
- Jika $I < I_0$ maka terjadi autokorelasi negatif, artinya lokasi-lokasi yang berdekatan memiliki nilai efek spasial yang berbeda dengan I_0 adalah nilai harapan dari I .

2. Uji Heterogenitas Spasial

Heterogenitas spasial terjadi akibat adanya keragaman antar lokasi. Jadi setiap lokasi memiliki struktur dan parameter hubungan yang berbeda. Uji yang digunakan untuk mengidentifikasi adanya efek heterogenitas spasial adalah uji *Breusch-Pagan* (BP). Hipotesis yang digunakan pada uji ini adalah

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 \text{ (ragam antar lokasi sama)}$$

$$H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ dimana } i \neq j \text{ (ragam antar lokasi berbeda)}$$

Statistik uji yang digunakan pada uji *Breusch-Pagan* yaitu:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (10)$$

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 independen yang sudah

distandarisasi berukuran $n \times (p + 1)$. Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $BP > X_{\alpha,p}^2$ atau jika $p - value < \alpha$ dengan p adalah banyaknya variabel independen yang artinya terdapat heterogenitas spasial pada data (Agustina dkk., 2022).

1.5.4 Pengujian Lagrange Multiplier

Uji *Lagrange Multiplier* (LM) digunakan sebagai dasar untuk memilih model regresi spasial yang sesuai (LeSage & Pace, 2009). Uji LM terdiri dari LM_{lag} dan LM_{error} . Kelebihan uji ini adalah dapat mengidentifikasi dependensi spasial yang terjadi pada model apakah terjadi pada *lag*, *error*, atau keduanya. Apabila LM_{lag} signifikan maka model yang sesuai adalah *Spatial Autoregressive* (SAR), jika LM_{error} signifikan maka model yang sesuai adalah *Spatial Error Model* (SEM), dan jika keduanya signifikan maka model yang sesuai adalah *Spatial Autoregressive Moving Average* (SARMA). Namun, apabila LM_{lag} dan LM_{error} keduanya tidak signifikan, maka model yang sesuai adalah *Ordinary Least Square* (OLS) (Agustina dkk., 2022).

1. Lagrange Multiplier Lag

Uji *lagrange multiplier lag* digunakan untuk mengidentifikasi dependensi spasial pada *lag* yang mana jika hasil uji signifikan maka penelitian dapat dilanjutkan dengan model regresi spasial SAR. Hipotesis yang digunakan pada uji ini adalah (Tumanggor & Simamora, 2023).

$H_0: \rho = 0$ (tidak ada dependensi spasial *lag*)

$H_1: \rho \neq 0$ (ada dependensi spasial *lag*)

Statistik uji yang digunakan pada LM_{lag} seperti pada persamaan (11) (Soleha, 2023).

$$LM_{lag} = \frac{\left(\frac{e^T W y}{s^2}\right)^2}{nJ} \quad (11)$$

dengan,

$$nJ = T + \frac{(WX\beta)^T M (WX\beta)}{s^2}$$

$$T = tr(W^2 + W^T W)$$

$$M = I - X(X^T X)^{-1} X^T$$

$$s^2 = \frac{e^T e}{n}$$

Keterangan:

e : nilai residual dari hasil OLS

n : banyaknya lokasi amatan

W : matriks pembobot spasial terstandarisasi berukuran $n \times n$

X : matriks variabel independen dengan ukuran $n \times (p + 1)$

β : vektor koefisien parameter regresi dengan ukuran $(p + 1) \times 1$

I : matriks identitas

y : vektor variabel dependen berukuran $n \times 1$

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $LM_{lag} > X_{1,\alpha}^2$ atau jika $p - value < \alpha$ yang artinya terdapat dependensi spasial lag pada variabel dependen.

2. Lagrange Multiplier Error

Uji *lagrange multiplier error* digunakan untuk mengidentifikasi dependensi spasial pada *error* yang mana jika hasil uji signifikan maka penelitian dapat dilanjutkan dengan model regresi spasial SEM. Hipotesis yang digunakan pada uji ini adalah (Tumanggor & Simamora, 2023).

$H_0: \lambda = 0$ (tidak ada dependensi spasial *error*)

$H_1: \lambda \neq 0$ (ada dependensi spasial *error*)

Statistik uji yang digunakan pada LM_{error} seperti pada persamaan (12) (Soleha, 2023).

$$LM_{error} = \frac{\left(\frac{e^T W e}{S^2}\right)^2}{T} \quad (12)$$

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $LM_{error} > X_{1,\alpha}^2$ atau jika $p - value < \alpha$ yang artinya terdapat dependensi spasial *error* pada *error*.

1.5.5 Spatial Autoregressive Model

Spatial Autoregressive Model (SAR) adalah salah satu model regresi spasial dengan pendekatan area yang memperhitungkan pengaruh spasial *lag* pada variabel dependen saja. Model SAR menurut LeSage & Pace (2009) dapat dituliskan:

$$y = \rho W y + X \beta + \varepsilon \quad (13)$$

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

Keterangan:

- y : vektor variabel dependen berukuran $n \times 1$
- ρ : koefisien parameter spasial lag dari variabel dependen
- W : matriks pembobot spasial berukuran $n \times n$
- X : matriks variabel independen berukuran $n \times (p + 1)$
- β : vektor koefisien parameter regresi berukuran $(p + 1) \times 1$
- ε : vektor residual berukuran $n \times 1$

1.5.6 Pengujian Model Regresi

1. Uji Kecocokan Model

Uji kecocokan model SAR dilakukan dengan uji F. Uji ini digunakan untuk melihat apakah terdapat hubungan secara simultan variabel independen terhadap variabel dependen (Montgomery & Runger, 2011). Hipotesis yang digunakan pada uji ini adalah (Merdekawaty dkk., 2016).

$H_0: \beta_j = \rho = 0$

$H_1: \rho \neq 0$ dan terdapat minimal satu $\beta_j \neq 0$, $j = 1, 2, \dots, p$

Statistik uji yang digunakan yaitu (Musyarofah dkk, 2020):

$$F_{hitung} = \frac{\frac{SSR}{p}}{\frac{SSE}{n-p-1}} \quad (14)$$

Keterangan:

SSR : jumlah kuadrat regresi

SSE : jumlah kuadrat residual

p : jumlah variabel independen

n : jumlah sampel

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $F_{hitung} > F_{\alpha,p,n-p-1}$ atau jika $p - value < \alpha$.

2. Uji Signifikansi Parameter

Uji signifikansi parameter pada model SAR dilakukan dengan uji *Wald* (Anselin, 1988). Pengujian parameter ρ dan β dilakukan sebagai berikut (Mukrom dkk., 2021).

- a. Pengujian parameter ρ digunakan hipotesis dan statistik uji yaitu:

$$H_0: \rho = 0 ; H_1: \rho \neq 0$$

$$W_\rho = \frac{\hat{\rho}^2}{var(\hat{\rho})} \quad (15)$$

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

- b. Pengujian parameter β digunakan hipotesis dan statistik uji yaitu:

$$H_0: \beta_j = 0 ; H_1: \beta_j \neq 0, j = 1, 2, \dots, p$$

$$W_\beta = \frac{\hat{\beta}_j^2}{var(\hat{\beta}_j)} \quad (16)$$

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

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $Wald > X_{\alpha,1}^2$ atau jika $p - value < \alpha$.

3. Uji Asumsi

- a. Uji Normalitas

Salah satu uji yang dapat digunakan untuk normalitas residual adalah uji *Shapiro-Wilk* dengan hipotesis dan statistik uji yaitu:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik uji *Shapiro-Wilk* dirumuskan pada persamaan (17) (Syam dkk., 2024).

$$W_{hitung} = \frac{1}{s^2} \left[\sum_{i=1}^n a_i (e_{n-i+1} - e_i) \right]^2 \quad (17)$$

dengan,

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

Keterangan:

a_i : koefisien *Shapiro-Wilk test*

e_{n-i+1} : residual ke $n - i + 1$

e_i : residual ke- i

\bar{e} : rata-rata residual

e_i : residual ke- i

s^2 : varians residual

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

b. Uji Multikolinieritas

Untuk mengetahui adanya multikolinieritas, salah satu caranya dilakukan dengan menghitung *Variance Inflation Faktors* (VIF), yang dirumuskan dengan persamaan (18) (Montgomery dkk., 2012).

$$VIF = \frac{1}{(1 - R_j^2)}; j = 1, 2, 4, \dots, p \quad (18)$$

dengan R_j^2 adalah nilai koefisien determinasi pada variabel ke- j . Nilai VIF < 10 menunjukkan tidak ada multikolinieritas dan VIF > 10 menunjukkan adanya multikolinieritas yang kuat.

c. Uji Homoskedastisitas

Asumsi homoskedastisitas disebut juga dengan asumsi kesamaan varians, artinya seluruh nilai-nilai dari variabel independen memiliki varians residual ε_i yang sama (Çelik, 2017). Salah satu uji yang digunakan untuk mengidentifikasi adanya homoskedastisitas pada data yaitu uji *Breusch-Pagan* dengan hipotesis (Mustari & Zain, 2017).

$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$ (terjadi homoskedastisitas)

$H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ dengan } i \neq j$ (tidak terjadi homoskedastisitas)

Statistik uji *Breusch-Pagan* (Anselin, 1988) yaitu:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (19)$$

dengan,

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

$$\sigma^2 = \frac{\mathbf{e}^T \mathbf{e}}{n}$$

Keterangan:

e_i^2 : residual untuk observasi ke- i

\mathbf{Z} : matriks variabel independen berukuran $n \times (p + 1)$

p : jumlah variabel independen

n : jumlah sampel

\mathbf{e} : vektor residual

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $BP > X_{\alpha, p}^2$ atau jika $p - \text{value} < \alpha$ yang artinya tidak terjadi homoskedastisitas.

1.5.7 Ukuran Kebaikan Model

Metode yang digunakan untuk mengukur kebaikan model adalah koefisien determinasi (R^2) dan nilai *Akaike's Information Criterion* (AIC). Koefisien determinasi (R^2) adalah besaran yang digunakan untuk menilai kebaikan model regresi. Kebanyakan peneliti lebih menyukai menggunakan nilai koefisien determinasi yang disesuaikan (R_{adj}^2) (Montgomery & Runger, 2011), seperti pada persamaan (20).

$$R_{adj}^2 = 1 - \frac{\frac{SSE}{n-p-1}}{\frac{SST}{n-1}} \quad (20)$$

dengan SST adalah jumlah kuadrat total, SSE adalah jumlah kuadrat residual, p adalah banyaknya variabel independen, dan n adalah jumlah sampel.

Selain itu, nilai AIC dapat mengukur kebaikan dan kesesuaian model dalam menduga parameter. Suatu model regresi dikatakan model terbaik jika memiliki nilai AIC terkecil. Nilai AIC dapat dihitung menggunakan persamaan (21) (Anselin, 1988).

$$AIC = 2k - 2\ln L(\beta) \quad (21)$$

dengan k adalah jumlah parameter yang digunakan dan $L(\beta)$ adalah maksimum logaritma *likelihood*.

1.5.8 Pendeteksian Pencilan Spasial

Pencilan spasial merupakan observasi yang memiliki nilai yang relatif sangat jauh dari observasi yang lainnya dengan mempertimbangkan bobot spasial (Musyarofah dkk., 2020). Pencilan spasial dapat dideteksi menggunakan *Moran's Scatterplot* (Atikah dkk., 2021). *Moran's Scatterplot* adalah grafik yang menunjukkan hubungan antara nilai amatan pada suatu lokasi dengan rata-rata nilai amatan tetangganya. *Moran's Scatterplot* terbagi menjadi 4 kuadran.

Kuadran II <i>Low-High</i>	Kuadran I <i>High-High</i>
Kuadran III <i>Low-Low</i>	Kuadran IV <i>High-Low</i>

Gambar 1. Ilustrasi *Moran's Scatterplot*

Plot data yang berada pada kuadran II dan IV menunjukkan adanya autokorelasi spasial negatif yang menggambarkan lokasi pengamatan rendah dikelilingi oleh lokasi pengamatan tinggi maupun sebaliknya. Oleh karena itu, dapat dideteksi bahwa titik-titik yang berada pada kuadran tersebut dikategorikan sebagai pencilan spasial (Syam dkk., 2024).

1.5.9 Regresi *Robust*

Regresi *robust* adalah salah satu metode regresi untuk menganalisis data yang dipengaruhi adanya *outlier* sehingga model yang dihasilkan resisten terhadap *outlier*. Metode ini adalah alat yang penting untuk menganalisis data yang terkontaminasi oleh *outlier* dan dapat memberikan hasil yang resisten terhadap *outlier* (Chen, 2002). Terdapat beberapa metode estimasi parameter pada regresi *robust* yaitu estimasi *Least Trimmed Square* (LTS), *Scale* (S), *Maximum Likelihood type* (M), dan *Method of Moment* (MM) (Rositawati & Fitri, 2022).

1. Fungsi Objektif

Fungsi objektif adalah fungsi yang digunakan untuk mencari fungsi pembobot pada regresi *robust*. Fungsi pembobot yang dapat digunakan adalah fungsi pembobot *Tukey Bisquare* pada persamaan (22).

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & ; |u_i| \leq c \\ 0 & ; |u_i| > c \end{cases} \quad (22)$$

Nilai c pada metode *Tukey Bisquare* disebut *tuning constant*. *Tuning constant* untuk fungsi pembobot *Tukey Bisquare* pada analisis regresi *robust S-Estimator* adalah $c = 1,547$ dan *tuning constant* untuk *M-Estimator* adalah $c = 4,685$ (Fox & Weisberg, 2010).

2. *Robust S-Estimator*

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

$$\hat{\beta}_s = \min \hat{\sigma}_s [e_1(\beta), e_2(\beta), \dots, e_n(\beta)] \quad (23)$$

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

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

$$\hat{\sigma}_s = \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2}$$

dengan $K = 0,199$. Estimator $\hat{\beta}$ pada metode regresi *robust* estimasi S diperoleh dengan cara melakukan iterasi hingga diperoleh hasil yang konvergen. Skala *robust* yang digunakan pada iterasi awal pada regresi *robust* estimasi S adalah

$$\hat{\sigma}_s^{(1)} = \frac{\text{median}|e_i - e_m|}{k}$$

dengan $k = 0,6745$. Proses ini dikenal sebagai MKT terboboti secara iterasi yang selanjutnya disebut *Iteratively Reweighted Least Square* (IRLS) (Musyarofah dkk., 2020).

3. *Robust M-Estimator*

Metode *robust M-Estimator* umumnya dilakukan dengan meminimumkan fungsi objektif: (Montgomery dkk., 2012).

$$\min_{\beta} \sum_{i=1}^n \rho(e_i) = \min_{\beta} \sum_{i=1}^n \rho \left(y_i - \sum_{j=0}^p x_{ij} \beta_j \right) \quad (24)$$

Untuk mendapatkan skala *invariant* pada estimator ini, biasanya dengan menyelesaikan persamaan (25).

$$\min_{\beta} \sum_{i=1}^n \rho(u_i) = \min_{\beta} \sum_{i=1}^n \rho \left(\frac{e_i}{s} \right) = \min_{\beta} \sum_{i=1}^n \rho \left(\frac{y_i - \sum_{j=0}^p x_{ij} \beta_j}{s} \right) \quad (25)$$

dengan $\hat{\sigma}_M$ adalah skala estimasi *robust*. Estimasi $\hat{\sigma}_M$ yang sering digunakan adalah:

$$\hat{\sigma}_M = \frac{\text{median}|e_i - e_m|}{k}$$

dengan $k = 0,6745$ membuat $\hat{\sigma}_M$ mendekati *estimator* tak bias dari σ dan residual berdistribusi normal.

4. **Robust MM-Estimator**

Metode *MM-Estimator* merupakan gabungan dari metode *S-Estimator* yang mempunyai nilai *breakdown point* tinggi dan metode *M-Estimator* yang mempunyai efisiensi tinggi (Yohai, 1987). Ada tiga tahapan *MM-Estimator* yaitu.

- Menghitung estimasi parameter awal regresi dengan metode yang memiliki *breakdown point* tinggi (*S-Estimator*).
- Menghitung residual dan skala estimasi *robust* dengan menggunakan *M-Estimator*.
- Menghitung estimasi parameter akhir dengan *MM-Estimator* yang didefinisikan pada persamaan (26) (Sari dkk., 2020).

$$\hat{\beta}_{MM} = \min \sum_{i=1}^n \rho \left(\frac{e_i}{\hat{\sigma}} \right) = \min \sum_{i=1}^n \rho \left(\frac{y_i - \sum_{j=0}^p X_i \beta_j}{\hat{\sigma}} \right) \quad (26)$$

1.5.10 Tuberkulosis

Tuberkulosis (TBC) merupakan suatu penyakit menular yang paling sering menyerang paru-paru dan disebabkan oleh kuman *Mycobacterium Tuberculosis*. Ini menyebar melalui udara ketika orang yang terinfeksi batuk, bersin, atau meludah. Sekitar seperempat dari populasi global diperkirakan telah terinfeksi bakteri TBC. Sekitar 5-10% orang yang terinfeksi TBC pada akhirnya mengalami gejala dan mengembangkan penyakit TBC yang sangat mematikan apabila tidak segera dilakukan penanganan (WHO, 2023). Jumlah kematian akibat TBC secara global pada tahun 2021 sebesar 1,4 juta, hal ini meningkat dibandingkan dengan tahun 2020 sebesar 1,3 juta. Tahun 2021, Indonesia berada pada urutan ke-2 penderita TBC tertinggi di dunia setelah India dengan proporsi kasus baru sebesar 13% dibandingkan seluruh kasus di dunia (WHO, Global Tuberculosis Report, 2022).

Penyakit TBC menurut data Kementerian Kesehatan (Kemenkes) RI pada tahun 2021, kasus TBC di Sulawesi Selatan ada sebanyak 14.303 kasus per kabupaten/kota yang mana berada pada urutan ke-7 dari 34 provinsi di Indonesia dengan jumlah kasus tertinggi. Oleh karena itu, TBC saat ini masih menjadi salah

satu masalah kesehatan masyarakat, selain malaria dan HIV/AIDS dan menjadi target dari *Sustainable Development Goals* (SDGs) untuk diturunkan. Penanganan penyakit TBC telah dilakukan sejak dini di Indonesia dengan pemberian paket imunisasi *Bacillus Calmette-Gluerin* (BCG) pada balita (Dinkes Sulsel, 2022).

BAB II

METODE PENELITIAN

2.1 Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari Profil Kesehatan Provinsi Sulawesi Selatan Tahun 2022 yang dipublikasikan oleh Dinas Kesehatan Provinsi Sulawesi Selatan. Data penelitian dapat dilihat pada Lampiran 1. Adapun variabel penelitian yang digunakan dalam penelitian ini adalah variabel dependen (Y) dan variabel independen (X) yang dapat dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
Y	Jumlah kasus tuberkulosis di Provinsi Sulawesi Selatan
X_1	Kepadatan penduduk
X_2	Persentase penduduk miskin
X_3	Persentase Pemberian vaksinasi BCG
X_4	Persentase gizi buruk
X_5	Persentase rumah tangga yang memiliki sanitasi layak

2.2 Analisis Data

Analisis data yang digunakan pada penelitian ini adalah regresi spasial dengan metode *Robust Spatial Autoregressive (RSAR)* dengan *MM-Estimator* dan matriks *Power Distance Weights*. Dalam pengolahan data digunakan *Software R-Studio Version 4.2.1*. Adapun tahap analisis data yang dilakukan sebagai berikut.

1. Melakukan eksplorasi data untuk melihat gambaran awal data variabel dependen dari suatu lokasi dengan menggunakan peta tematik dan menghitung korelasi antar variabel.
2. Melakukan pemodelan regresi linier berganda dan pengujian asumsi pada model regresi.
3. Menentukan matriks pembobot spasial untuk mengidentifikasi kedekatan setiap lokasi. Matriks pembobot spasial yang digunakan adalah *power distance weights* berdasarkan informasi jarak sesungguhnya dengan satuan kilometer antar lokasi. Untuk menghitung matriks pembobot spasial pada penelitian ini secara berturut-turut digunakan persamaan (6) dan (7).

4. Mengidentifikasi efek heterogenitas spasial menggunakan uji *Breusch-Pagan* dengan hipotesis:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 \text{ (ragam antar lokasi sama)}$$

$$H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ dimana } i \neq j \text{ (ragam antar lokasi berbeda)}$$

Statistik uji yang digunakan pada uji ini dapat dihitung menggunakan persamaan (10). Tolak H_0 pada taraf signifikansi α jika nilai $BP > X_{\alpha,p}^2$ atau jika $p - \text{value} < \alpha$.

5. Mengidentifikasi efek dependensi spasial menggunakan *Morans'I*. Menghitung *Morans'I* menggunakan rumus pada persamaan (8). Hipotesis yang digunakan adalah

$$H_0: I = 0 \text{ (tidak terdapat dependensi spasial)}$$

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

Statistik uji yang digunakan pada uji ini dapat dihitung menggunakan persamaan (9). Tolak H_0 pada taraf signifikansi α jika $|Z(I)| > Z_{\frac{\alpha}{2}}$ atau $p - \text{value} < \alpha$.

6. Melakukan uji *Lagrange Multiplier* (LM) untuk menentukan model yang sesuai untuk digunakan dengan menggunakan persamaan (11) dan (12).
7. Mengestimasi parameter *Spatial Autoregressive Model* (SAR) dengan metode *Maximum Likelihood Estimation* (MLE) sehingga diperoleh model akhir yang sesuai.
8. Melakukan pengujian asumsi regresi pada model SAR.
9. Melakukan pendeteksian pencilan spasial dengan *Morans Scatterplot*. Apabila titik data berada pada kuadran II dan IV maka dikategorikan sebagai pencilan spasial.
10. Mengestimasi parameter *Robust Spatial Autoregressive* dengan *MM-Estimator*. Tahapan-tahapan estimasi parameter dengan *MM-Estimator* menurut Rahmiatun dkk. (2022) dan Shodiqin dkk. (2018) yaitu:

- a. Menghitung nilai residual $e_i = y_i - \hat{y}_i$

- b. Menghitung $\hat{\sigma}_s = \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2}$ dimana untuk iterasi awal digunakan

$$\hat{\sigma}_s^{(1)} = \frac{\text{median}|e_i - e_m|}{0,6745}$$

- c. Menghitung nilai $u_i = \frac{e_i}{\hat{\sigma}_s}$

- d. Menghitung nilai pembobot *Tukey Bisquare* untuk *S-Estimator* dengan menggunakan persamaan (22)

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & ; |u_i| \leq 1,547 \\ 0 & ; |u_i| > 1,547 \end{cases} \quad (22)$$

Mengestimasi parameter $\beta_s^{(1)}$ dengan metode *WLS* (*Weighted Least Squares*) menggunakan pembobot $w(u_i)$ sehingga diperoleh $e_i^{(1)}$

- e. Residual $e_i^{(1)}$ pada langkah pertama digunakan untuk menghitung skala residual *M-Estimator*, $\hat{\sigma}_M = \frac{\text{median}|e_i - e_m|}{0,6745}$

- f. Menghitung nilai $u_i = \frac{e_i}{\hat{\sigma}_M}$ lalu menghitung nilai pembobot *Tukey Bisquare* untuk *M-Estimator* menggunakan persamaan (22).

$$w(u_i)^{(1)} = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & ; |u_i| \leq 4,685 \\ 0 & ; |u_i| > 4,685 \end{cases} \quad (22)$$

- g. Residual $e_i^{(1)}$ dan skala estimasi $\hat{\sigma}_M$ dari langkah f digunakan dalam iterasi awal dengan metode WLS untuk menghitung $\hat{\beta}_{MM}$ menggunakan $w(u_i)^{(1)}$
- h. Langkah f, g, dan h diulang (reiterasi dengan skala estimate tetap konstan) sampai $\sum_{i=1}^n |e_i^m|$ konvergen yaitu selisih $\hat{\beta}_{MM}^{m+1}$ dengan $\hat{\beta}_{MM}^m$ kurang dari mendekati 0, dengan m adalah banyak iterasi.
11. Melakukan pengujian kecocokan model RSAR menggunakan uji F sesuai dengan statistik uji pada persamaan (14). Tolak H_0 pada taraf signifikansi α jika nilai $F_{hitung} > F_{\alpha, p, n-p-1}$ atau jika $p - value < \alpha$.
12. Melakukan uji signifikan parameter model RSAR menggunakan uji *Wald*. Pengujian parameter ρ dan β secara berturut-turut dapat dihitung menggunakan persamaan (15) dan (16). Tolak H_0 pada taraf signifikansi α jika nilai $Wald > X_{\alpha, 1}^2$ atau jika $p - value < \alpha$.
13. Menentukan dan mengevaluasi ukuran kebaikan model SAR dan RSAR berdasarkan koefisien determinasi yang disesuaikan (R_{adj}^2) dan nilai *Akaike's Information Criterion* (AIC). Perhitungan nilai R_{adj}^2 menggunakan persamaan (20) dan nilai AIC menggunakan persamaan (21). Model yang lebih baik memiliki nilai R_{adj}^2 lebih besar dan nilai AIC yang lebih kecil.
14. Menginterpretasi model dan membuat keputusan.