

**PEMODELAN ROBUST GEOGRAPHICALLY AND TEMPORALLY  
WEIGHTED AUTOREGRESSIVE DENGAN MM-ESTIMATOR  
PADA TINGKAT PENGANGGURAN TERBUKA  
DI SULAWESI SELATAN**

**FADLAN AMIN  
H051201036**



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

**PEMODELAN ROBUST GEOGRAPHICALLY AND TEMPORALLY  
WEIGHTED AUTOREGRESSIVE DENGAN MM-ESTIMATOR  
PADA TINGKAT PENGANGGURAN TERBUKA  
DI SULAWESI SELATAN**

FADLAN AMIN  
H051201036



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

**SKRIPSI**  
**PEMODELAN ROBUST GEOGRAPHICALLY AND TEMPORALLY**  
**WEIGHTED AUTOREGRESSIVE DENGAN MM-ESTIMATOR**  
**PADA TINGKAT PENGANGGURAN TERBUKA**  
**DI SULAWESI SELATAN**

**FADLAN AMIN**  
**H051201036**

Skripsi,

telah dipertahankan di depan Panitia Ujian Sarjana Statistika pada tanggal 6  
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 Geographically and Temporally Weighted Autoregressive* dengan *MM-Estimator* pada Tingkat Pengangguran Terbuka 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 pembuatan tersebut berdasarkan aturan yang berlaku.

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

Makassar, 6 Agustus 2024



Fadlan Amin  
NIM. H051201036

## UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas segala limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "**Pemodelan *Robust Geographically and Temporally Weighted Autoregressive* dengan *MM-Estimator* pada Tingkat Pengangguran Terbuka di Sulawesi Selatan**". Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan para sahabatnya.

Penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada **Bapak Siswanto, S.Si., M.Si.** selaku Pembimbing Utama yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya, memberikan motivasi, dorongan semangat, dan bimbingan kepada penulis selama proses penulisan tugas akhir ini. Terima kasih kepada **Ibu Anisa, S.Si., M.Si.** dan **Ibu Dr. Erna Tri Herdiani, S.Si., M.Si.**, selaku Tim Penguji yang senantiasa memberikan saran dan kritikan yang membangun dalam penyempurnaan penulisan tugas akhir ini. Terima kasih kepada **Pimpinan Universitas Hasanuddin, Fakultas MIPA, Departemen Statistika, Jajaran Dosen, dan Staf Departemen Statistika** yang telah memfasilitasi, memberikan ilmu bermanfaat, dan membantu penulis selama menempuh studi.

Penghargaan dan ucapan terima kasih yang tulus juga penulis ucapkan kepada kedua orang tua penulis yang tercinta, Ibunda **Hj. A. Haryati Yusuf, S.P.** dan Ayahanda **Dr. H. Muhammad Amin, S.Si., M.Eng.** yang telah memberikan kasih sayang, dukungan, semangat, dan doa yang senantiasa mengiringi langkah penulis. Terima kasih juga kepada **Nahdiah Amin, Fakhri Amin, dan Faqih Amin** sebagai saudara saya yang senantiasa menjadi tempat berdiskusi, bersunda-gurau, dan bercanda-tawa. Terima kasih kepada keluarga besar **Himastat FMIPA Unhas** khususnya untuk teman-teman **POIS20N** serta teman-teman seperjuangan **STATISTIKA 2020** untuk segala ilmu, cerita dan pengalamannya. Terima kasih teman-teman **Random, Seperbimbingan, dan Project Kelulusan** yang senantiasa menemani perjalanan penulis dalam menyelesaikan tugas akhir. Terima kasih kepada sahabat-sahabat **Mukhlis, Fahmi, Razy, Izzul, Kur, Ryan, Hakam, Reza, Azhar, Alif, Rais, Dzaky, Bahar, Rani, Sabil, Uci, Azal, Liza, Ara, Divia, Linda, Heriyani, Sri Rahayu, Laurine, Krisna, Sitti Aisyah, Nahla, Nur Aulia, dan Rahmi** yang setia menemani dalam menyelesaikan skripsi serta momen-momen kebersamaan belajar, diskusi yang terus membangun diri penulis.

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

Penulis,

Fadlan Amin

## ABSTRAK

FADLAN AMIN. **Pemodelan *Robust Geographically and Temporally Weighted Autoregressive* dengan *MM-Estimator* pada Tingkat Pengangguran Terbuka di Sulawesi Selatan** (dibimbing oleh Bapak Siswanto, S.Si., M.Si.).

**Latar belakang.** Regresi spasial memodelkan hubungan antara variabel dengan mempertimbangkan efek spasial. Akibat efek tersebut, estimasi parameter model regresi memberikan keragaman respon variabel yang berbeda antar lokasi dan waktu, disebut dengan heterogenitas spasial temporal. Dalam kasus spasial, data tidak hanya mengalami heterogenitas. Nilai suatu daerah dapat dipengaruhi oleh nilai daerah lain, sehingga menimbulkan sifat ketergantungan antar wilayah yang disebut autokorelasi spasial temporal. *Geographically and Temporally Weighted Autoregressive* (GTWAR) menjadi solusi dalam membangun model pada data yang mengalami heterogenitas dan autokorelasi. Pemodelan regresi GTWAR memberikan interpretasi yang akurat, namun ditemukan pencilan dalam data yang mengakibatkan kesalahan dalam *output* yang dihasilkan. Sehingga *Robust Geographically and Temporally Weighted Autoregressive* (RGTWAR) dengan *MM-Estimator* menjadi salah satu metode pengatasan pencilan yang memiliki keunggulan efisiensi yang tinggi dan nilai *breakdown point* yang tinggi. **Tujuan.** Penelitian ini bertujuan untuk mendapatkan estimasi parameter model RGTWAR dan faktor-faktor yang mempengaruhi data. **Metode.** Penelitian ini menerapkan model RGTWAR dengan *MM-Estimator* pada data Tingkat Pengangguran Terbuka (TPT) di Sulawesi Selatan tahun 2018-2022. Tahapan analisis data diawali dengan melakukan estimasi parameter model RGTWAR dengan *Maximum Likelihood Estimation* (MLE) kemudian dilakukan estimasi *robust* dengan *MM-Estimator*. **Hasil.** Model RGTWAR dengan *MM-Estimator* menghasilkan nilai  $R_{Adj}^2$  yang lebih baik sebesar 94,03% dibanding model GTWAR dengan nilai 93,40%. Selain itu, nilai RMSE dan *AIC* dari model RGTWAR menurun dengan nilai 0.62162 dan 140.24720 dari model GTWAR dengan nilai 0.65364 dan 156.46110. **Kesimpulan.** Pada signifikansi 5%, faktor yang mempengaruhi TPT Sulawesi Selatan 2018-2022 adalah PDRB, Tingkat Partisipasi Angkatan Kerja, Kepadatan Penduduk, dan Angka Melek Huruf.

**Kata kunci:** Regresi Spasial, Heterogenitas, Autokorelasi, Pencilan, Pengangguran.

## ABSTRACT

FADLAN AMIN. **Robust Geographically and Temporally Weighted Autoregressive Modeling with MM-Estimator on Open Unemployment Rates in South Sulawesi** (Supervised by Mr. Siswanto, S.Si., M.Si.)

**Background.** Spatial regression models the relationship between variables by considering spatial effects. As a result of this effect, the estimated parameters of the regression model provide a diversity of variable responses that differ between locations and time, which is called spatio-temporal heterogeneity. In the spatial case, the data is not only heterogeneous. The value of one area can be influenced by the value of other areas, giving rise to a dependency between regions called spatial-temporal autocorrelation. Geographically and Temporally Weighted Autoregressive (GTWAR) is a solution for building models on data that experiences heterogeneity and autocorrelation. GTWAR regression modeling provided an accurate interpretation, however, outliers were found in the data which resulted in errors in the resulting output. Robust Geographically and Temporally Weighted Autoregressive (RGTWAR) with MM-Estimator is a method for dealing with outliers which has the advantage of high efficiency and high breakdown point. **Objective.** This research aims to obtain estimates of the RGTWAR model parameters and factors that influence the data. **Method.** This research applies the RGTWAR model with MM-Estimator to Open Unemployment Rate in South Sulawesi for 2018-2022. The data analysis stage begins with estimating the parameters of the RGTWAR model with Maximum Likelihood Estimation (MLE) then robust estimation is carried out with MM-Estimator. **Results.** The RGTWAR model with MM-Estimator produces a better  $R_{Adj}^2$  value of 94.03% compared to the GTWAR model with a value of 93.40%. In addition, the RMSE and AIC values of the RGTWAR model decreased with values of 0.62162 and 140.24720 from the GTWAR model with values of 0.65364 and 156.46110. **Conclusion.** At a significance of 5%, the factors influencing Open Unemployment Rate in South Sulawesi in 2018-2022 are GRDP, Labor Force Participation Rate, Population Density, and Literacy Rate.

**Keywords:** Spatial Regression, Heterogeneity, Autocorrelation, Outliers, Unemployment

## DAFTAR ISTILAH

<b>Istilah</b>	<b>Arti dan Penjelasan</b>
Autokorelasi Spasial	Sifat ketergantungan suatu wilayah terhadap nilai wilayah lain
Autokorelasi Temporal	Sifat ketergantungan suatu wilayah terhadap nilai wilayahnya di waktu lain
Bobot	Nilai yang diberikan untuk menunjukkan besar pengaruh dari lokasi amatan, memiliki nilai dari 0 sampai 1
Distribusi	Pola penyebaran data atau probabilitas suatu kejadian dalam ruang sampel
Geografis	Sesuatu yang berhubungan dengan Lokasi, posisi atau karakteristik suatu wilayah
Heterogenitas Spasial	Variasi dalam lokasi atau karakteristik yang berbeda di berbagai lokasi pengamatan
Heterogenitas Temporal	Variasi dalam waktu atau karakteristik yang berbeda di berbagai waktu pengamatan
Korelasi	Hubungan antara dua atau lebih amatan
Lag	Tingkat keterkaitan suatu lokasi terhadap lokasi lainnya
Model	Representasi statistik untuk menggambarkan hubungan antara variabel-variabel dalam suatu data
Parameter	Nilai atau koefisien tetap dalam model atau fungsi
Prediktor	Variabel independen dalam penelitian yang dianggap mempengaruhi variabel respon
Pencilan	Amatan yang memiliki nilai jauh dari amatan lainnya
Regresi	Metode analisis hubungan antara variabel respon dan satu atau lebih variabel prediktor
Residual	Nilai selisih variabel respon dan variabel dugaan
Respon	Variabel dependen dalam penelitian
Variabel	Entitas yang nilainya dapat berubah



## DAFTAR LAMBANG DAN SINGKATAN

Lambang/singkatan	Arti dan Penjelasan
$y_i$	Variabel respon pada lokasi pengamatan ke- $i$
$\varepsilon_i$	Galat ke- $i$
$\mathbf{x}_i^T$	Vektor variabel prediktor pada lokasi pengamatan ke- $i$
$\boldsymbol{\beta}$	Vektor parameter regresi berukuran $(p + 1) \times 1$
$\ln$	Logaritma natural
$db$	Derajat Bebas
$\hat{I}$	Nilai Indeks Moran
$w_{ij}$	Pembobot antara amatan ke- $i$ dan amatan ke- $j$
$(u_i, v_i, t_i)$	Koordinat geografis (lintang, bujur, waktu) pada amatan ke- $i$
$d_{ij}^{ST}$	Jarak spasiotemporal antara amatan $(u_i, v_i, t_i)$ ke amatan $(u_j, v_j, t_j)$
$\mu^S$	Parameter penyeimbang spasial
$\mu^T$	Parameter penyeimbang temporal
$\xi$	Parameter interaksi efek spasial dan temporal
$h$	Parameter penghalus ( <i>bandwidth</i> )
$\rho$	Koefisien korelasi
TPT	Tingkat Pengangguran Terbuka
CV	<i>Cross Validation</i>
BP	<i>Breusch-Pagan</i>
GWR	<i>Geographically Weighted Regression</i>
GTWR	<i>Geographically and Temporally Weighted Regression</i>
GTWAR	<i>Geographically and Temporally Weighted Autoregressive</i>
RGTWAR	<i>Robust Geographically and Temporally Weighted Autoregressive</i>
MLE	<i>Maksimum Likelihood Estimation</i>
MM	<i>Method of Moment</i>

## DAFTAR ISI

	Halaman
HALAMAN JUDUL .....	i
HALAMAN PENGANTAR .....	ii
HALAMAN PENGESAHAN .....	iii
PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA .....	iv
UCAPAN TERIMA KASIH .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
DAFTAR ISTILAH .....	viii
DAFTAR LAMBANG DAN SINGKATAN .....	ix
DAFTAR ISI .....	x
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xiv
DAFTAR LAMPIRAN .....	xv
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. Landasan Teori .....	3
1.5.1. Heterogenitas Spasial .....	3
1.5.2. Heterogenitas Temporal .....	4
1.5.3. <i>Geographically Weighted Regression</i> .....	4
1.5.4. <i>Geographically and Temporally Weighted Regression</i> .....	5
1.5.5. Autokorelasi Spasial dan Temporal .....	7
1.5.6. <i>Lagrange Multiplier Test</i> .....	8

1.5.7. <i>Spatial Autoregressive</i> .....	8
1.5.8. <i>Geographically and Temporally Weighted Autoregressive</i> .....	9
1.5.9. Pencilan .....	9
1.5.10. Regresi <i>Robust</i> .....	10
1.5.11. Uji Signifikansi Parameter .....	12
1.5.12. Uji Kelayakan Model .....	12
1.5.13. Tingkat Pengangguran Terbuka .....	13
BAB II METODOLOGI PENELITIAN.....	14
2.1. Sumber Data.....	14
2.2. Analisis Data.....	14
BAB III HASIL DAN PEMBAHASAN .....	16
3.1. Estimasi Parameter Model.....	16
3.1.1. Estimasi Parameter <i>Geographically and Temporally Weighted Autoregressive</i> .....	16
3.1.2. Estimasi Parameter <i>Robust Geographically and Temporally Weighted Autoregressive</i> .....	19
3.2. Model <i>Robust Geographically and Temporally Weighted Autoregressive</i> .....	21
3.2.1. Eksplorasi Data .....	21
3.2.2. Pengujian Heterogenitas Spasial dan Temporal.....	23
3.2.3. Penentuan Matriks Pembobot <i>Exponential Distance</i> .....	24
3.2.4. Pengujian Autokorelasi Spasial dan Temporal .....	24
3.2.5. Pengujian <i>Lagrange Multiplier</i> .....	25
3.2.6. Penentuan Matriks Pembobot Kernel.....	25
3.2.7. Penentuan Koefisien Korelasi Tiap Observasi .....	25
3.2.8. Pemodelan <i>Geographically and Temporally Weighted Autoregressive</i> .....	26
3.2.9. Pendeteksian Pencilan.....	27
3.2.10. Pemodelan <i>Robust Geographically and Temporally Weighted Autoregressive</i> dengan <i>MM-Estimator</i> .....	28

3.2.11. Pengujian Signifikansi Parameter .....	28
3.2.12. Pengujian Kelayakan Model.....	29
3.2.13. Interpretasi Model RGTWAR.....	29
BAB IV KESIMPULAN .....	32
4.1. Kesimpulan .....	32
4.2. Saran .....	32
DAFTAR PUSTAKA .....	33
LAMPIRAN .....	36

## DAFTAR TABEL

	Halaman
1. Variabel Respon dan Prediktor.....	14
2. Uji <i>Breusch-Pagan</i> .....	23
3. Uji <i>Moran's I</i> dari Variabel Respon .....	24
4. Uji <i>Lagrange Multiplier</i> .....	25
5. Parameter Optimum .....	25
6. Koefisien Korelasi Tiap Observasi.....	25
7. Daerah-Daerah dengan Nilai $\rho = 0$ .....	26
8. Estimasi Parameter Model GTWAR Kota Palopo Tahun 2018 .....	27
9. Estimasi Parameter Model RGTWAR Kota Palopo Tahun 2018 .....	28
10. Uji Signifikansi Parameter Model Kota Palopo Tahun 2018 .....	28
11. Nilai Uji Kelayakan Model.....	29

**DAFTAR GAMBAR**

	Halaman
1. Kuadran Moran Scatterplot .....	10
2. Pola Sebaran TPT pada 24 Kabupaten/Kota di Sulawesi Selatan Tahun 2018-2022 .....	22
3. Boxplot TPT di Sulawesi Selatan Tahun 2018-2022.....	23
4. Pendeteksian Pencilan Spasial .....	27
5. Parameter yang Signifikan dari Setiap Kabupaten/Kota di Sulawesi Selatan Tahun 2018-2022 .....	29

## DAFTAR LAMPIRAN

	Halaman
1. Data Penelitian .....	37
2. Jarak <i>Euclidean</i> .....	38
3. Jarak <i>Exponential Distance Weight</i> .....	39
4. Matriks Pembobot Kernel <i>Fixed Exponential</i> .....	40
5. Koefisien Korelasi .....	41
6. Estimasi Parameter Model <i>Geographically and Temporally Weighted Autoregressive</i> .....	42
7. Estimasi Parameter Model <i>Robust Geographically and Temporally Weighted Autoregressive</i> .....	46
8. Signifikansi Parameter Model <i>Robust Geographically and Temporally Weighted Autoregressive</i> .....	50
9. Model <i>Robust Geographically and Temporally Weighted Autoregressive</i> .....	54
10. Nilai $R_{Adj}^2$ dan AIC lokal setiap Kabupaten/Kota di Sulawesi Selatan Tahun 2018-2022 .....	55





## **BAB I**

### **PENDAHULUAN**

#### **1.1. Latar Belakang**

Regresi spasial merupakan metode statistika yang digunakan dalam memodelkan hubungan antara variabel respon dan variabel prediktor dengan memandang efek spasial. Efek spasial adalah ketergantungan spasial yang terjadi akibat adanya korelasi antarlokasi. Heterogenitas spasial adalah salah satu efek spasial dalam data yang terjadi ketika suatu lokasi memberikan perbedaan respon variabel dengan lokasi lain. Jenis regresi spasial yang digunakan sebagai metode untuk mengatasi data yang mengalami heterogenitas spasial dikenal sebagai *Geographically Weighted Regression* (GWR) (Safitri dan Amaliana, 2021). Metode tersebut merupakan teknik pemodelan regresi dalam analisis data spasial yang menggunakan konsep statistika. Pemodelan GWR menggunakan matriks pembobot yang besarnya tergantung pada kedekatan antarlokasi serta menghasilkan estimasi parameter model yang bersifat lokal untuk setiap lokasi yang berbeda. Model yang dihasilkan oleh metode ini akan menunjukkan interpretasi yang berbeda pada tiap lokasi observasi (Safitri dan Amaliana, 2021).

Pembahasan ruang menjadi fokus utama dalam analisis spasial. Namun selain ruang, waktu menjadi dimensi yang berkaitan dengan proses geografis dan fenomena lingkungan. Suatu data akan mengalami perubahan seiring berjalannya waktu. Hal ini dikenal dengan istilah heterogenitas temporal, yang menunjukkan adanya perubahan pada data karena perbedaan waktu. Dalam beberapa tahun terakhir, terdapat beberapa studi yang membahas efek temporal ke dalam model GWR untuk mengkaji efek spasial dan temporal (Huang dkk., 2010). Mendefinisikan dan mengukur jarak ruang-waktu menjadi masalah yang krusial dalam mengkalkulasi spasial temporal. Disebabkan waktu dan ruang memiliki ukuran efek yang berbeda, Huang (2010) menemukan sebuah cara untuk menggabungkan jarak spasial dan jarak temporal ke dalam jarak spasial temporal dalam model *Geographically and Temporally Weighted Regression* (GTWR).

Jarak ruang-waktu yang dijelaskan oleh GTWR memberikan keakuratan yang baik dalam menganalisis heterogenitas spasial temporal dalam suatu wilayah, dan memberikan hasil yang baik (Erda & Djuraidah, 2018). Namun model GTWR tidak mampu menangkap autokorelasi spasial temporal dalam data. Wu (2014) mengklaim bahwa lebih baik menangani heterogenitas spasial dan temporal serta efek autokorelasi spasial dalam model campuran. Hal ini dikarenakan heterogenitas spasial dan autokorelasi pada umumnya berkaitan dalam konteks pemodelan meskipun secara teoritis kedua permasalahan tersebut dapat dibedakan (Nasir dkk., 2020). Dalam membangun model yang terintegrasi dengan efek autokorelasi spatiotemporal dan heterogenitas, dikembangkan model *Geographically and Temporally Weighted Autoregressive* (GTWAR) untuk mengatasi kedua efek spasial (Wu dkk., 2014).

Terdapat penelitian sebelumnya seperti yang telah dilakukan oleh peneliti sebelumnya. Wu (2014) meneliti harga perumahan di Shenzhen, China dengan menggunakan pemodelan GTWAR. Dibandingkan dengan model OLS, GWR, dan GTWR yang menghasilkan nilai  $R^2$  berturut-turut 0.617, 0.736, dan 0.875, model GTWAR menghasilkan nilai  $R^2$  yang lebih tinggi dengan nilai 0.914. Samsir (2023) juga memodelkan kasus Demam Berdarah di Sulawesi Selatan dengan *Robust Improved* GTWR, menunjukkan hasil *Robust* yang lebih baik dibanding model IGWR. Musdalifah (2023) memodelkan kasus Pneumonia pada Balita di Indonesia dengan RGTWR. Model tersebut menunjukkan MdAD sebesar 6.9661, lebih kecil dibanding nilai MdAD model GTWR sebesar 8.3474.

Pemodelan regresi seperti GTWAR memberikan interpretasi yang akurat dan informatif, namun terkadang ditemukan beberapa permasalahan, salah satunya adanya data pencilan. Amatan yang memiliki nilai jauh dari amatan lainnya mengakibatkan adanya *error* pada *output* yang dihasilkan (Freeman, 1995). Metode yang dapat digunakan untuk mengatasi permasalahan tersebut adalah metode regresi *Robust*. Metode ini dapat mengatasi pencilan dengan mencocokkan model regresi terhadap sebagian besar data tanpa menghilangkan data pencilan (Sugasawa dan Murakami, 2022). Terdapat beberapa metode estimasi yang dikembangkan dalam regresi robust, antara lain *M-Estimator*, *Least Median of Square* (LMS), *Least Trimmed Square* (LTS), *S-Estimator*, dan *MM-Estimator*. *M-Estimator* mempunyai efisiensi yang tinggi, namun memiliki nilai *breakdown point* sebesar 0. LMS, LTS dan *S-Estimator* mempunyai *breakdown point* yang tinggi yaitu sebesar 0.5, namun memiliki efisiensi yang sangat rendah. Sedangkan, *MM-Estimator* memiliki gabungan efisiensi tinggi dari *M-Estimator* dengan *breakdown point* yang tinggi dari *S-Estimator* (Shodiqin dkk., 2018). Regresi robust yang digunakan untuk mengatasi adanya pencilan spasial pada model GTWAR disebut *Robust Geographically and Temporally Weighted Autoregressive* (RGTWAR).

Tingkat Pengangguran Terbuka (TPT) merupakan besaran yang menunjukkan persentase jumlah pengangguran terhadap jumlah angkatan kerja, mengindikasikan besarnya persentase angkatan kerja yang termasuk dalam pengangguran. TPT yang tinggi menunjukkan bahwa terdapat banyak angkatan kerja yang tidak terserap pada pasar kerja. Di berita resmi BPS Sulawesi Selatan, TPT Sulawesi Selatan pada Februari 2023 berada di angka 5,26 persen. Angka ini menunjukkan adanya kenaikan sebesar 0,65 persen dari TPT pada Agustus 2022 yang berada pada angka 4,51 persen. TPT dari kabupaten/kota di Sulawesi Selatan juga menunjukkan adanya heterogenitas dan autokorelasi spasiotemporal. Setiap daerah memiliki karakteristik ekonomi, sosial, dan geografis yang berbeda-beda. Hal ini mengakibatkan variabel yang mempengaruhi tingkat pengangguran terbuka di suatu daerah akan berbeda dengan daerah lainnya. Kondisi angka TPT antarwilayah di Sulawesi Selatan yang beragam tersebut juga dapat menimbulkan adanya nilai pencilan. Berdasarkan penelitian yang telah dipaparkan, maka peneliti ingin melakukan sebuah penelitian dengan **Pemodelan *Robust Geographically and Temporally Weighted Autoregressive* dengan *MM-Estimator* pada Tingkat Pengangguran Terbuka di Sulawesi Selatan.**

## 1.2. Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Data penelitian yang digunakan adalah kasus tingkat pengangguran terbuka di Sulawesi selatan berdasarkan publikasi Badan Pusat Statistik Provinsi Sulawesi Selatan tahun 2018-2022.
2. Pembobot yang digunakan dalam perhitungan matriks pembobot adalah *Fixed Exponential* yang didasarkan pada perhitungan jarak *euclidean*, sedangkan pembobot yang digunakan dalam pembobot *robust* adalah pembobot Tukey.

## 1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mendapatkan bentuk estimasi parameter model *Robust Geographically and Temporally Weighted Autoregressive*.
2. Mendapatkan faktor-faktor yang mempengaruhi Tingkat Pengangguran Terbuka di Sulawesi Selatan tahun 2018-2022.

## 1.4. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Memberikan informasi tentang pemodelan Tingkat Pengangguran Terbuka menggunakan metode *Geographically and Temporally Weighted Autoregressive*.
2. Memberikan informasi tentang analisis regresi RGTWAR dengan *MM-Estimator*.
3. Sebagai inspirasi bagi pembaca untuk mengembangkan analisis regresi spasial dengan metode lain dan/atau kasus lain yang dapat ditemukan dalam kehidupan sehari-hari.
4. Sebagai sumber informasi bagi Dinas Sosial, Dinas Perindustrian dan Tenaga Kerja, atau pihak pemerintah terkait serta masyarakat umum mengenai faktor-faktor yang mempengaruhi Tingkat Pengangguran Terbuka di Sulawesi Selatan.

## 1.5. Landasan Teori

### 1.5.1. Heterogenitas Spasial

Heterogenitas spasial adalah suatu kondisi apabila satu variabel independen yang sama memberikan respon yang tidak sama pada lokasi yang berbeda dalam suatu wilayah pengamatan (Sobari & Jaya, 2023). Heterogenitas spasial dapat diuji dengan menggunakan uji *Breusch-Pagan*. Hipotesis dalam uji ini adalah:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma^2 \text{ (tidak terjadi heterogenitas antar wilayah)}$$

$$H_1: \text{minimal ada } \sigma_i^2 \neq \sigma^2 \text{ (terjadi heterogenitas antar wilayah)}$$

Statistik Uji:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} + \left( \frac{1}{T} \right) \left( \frac{\boldsymbol{\varepsilon}^T \mathbf{W} \boldsymbol{\varepsilon}}{\sigma^2} \right)^2 \sim \chi_{(p)}^2 \quad (1)$$

dengan vektor  $\mathbf{f}$  adalah  $f_i = \frac{\varepsilon_i^2}{\sigma^2} - 1$  dan  $T = \text{tr}[\mathbf{W}^T \mathbf{W} + \mathbf{W}^2]$

Keterangan:

$BP$  : statistic uji *Breusch-Pagan*

$\varepsilon$  : vector residual  $\varepsilon_i$  berukuran  $n \times 1$

$f^T$  :  $(f_1 \ f_2 \ \dots \ f_n)$

$Z$  : matriks  $X$  yang distandarisasi, berukuran  $n \times (p + 1)$

$W$  : matriks pembobot  $W_i$  berukuran  $n \times n$

Pengambilan Keputusan:

Apabila nilai  $BP > \chi^2_{(\alpha,p)}$ , maka  $H_0$  ditolak.

### 1.5.2. Heterogenitas Temporal

Heterogenitas temporal dilakukan untuk melihat keragaman suatu observasi pada periode yang berbeda. Pengujian ini dapat dilakukan melalui diagram *boxplot*. Apabila sebaran nilai variabel yang tergambarkan oleh *boxplot* menunjukkan adanya perbedaan antarvariabel, maka data terindikasi mengalami heterogenitas temporal.

### 1.5.3. Geographically Weighted Regression

Regresi linear berganda merupakan teknik analisis yang tepat dalam menggambarkan hubungan antara variabel respon dan predictor. Namun model yang dihasilkan merupakan model yang bersifat global, berarti model tersebut menggambarkan hubungan antarvariabel secara keseluruhan. Dalam analisis data spasial, diperlukan metode yang mampu menggambarkan hubungan secara lokal atau berdasarkan wilayah, sehingga dihadirkan *Geographically Weighted Regression* (Fotheringham dkk., 2002).

Persamaan regresi digambarkan dengan  $y_i$  sebagai variabel respon dan  $x_{i1}, x_{i2}, \dots, x_{ip}$  sebagai variabel prediktor. Selain itu, model GWR menghasilkan penaksir parameter model yang bersifat lokal untuk setiap titik atau lokasi data tersebut diamati (Sintia dkk., 2022). Sehingga model GWR dapat ditulis sebagai berikut:

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

Keterangan:

$y_i$  : nilai respon untuk lokasi ke-i

$x_{ik}$  : nilai predictor ke-k pada lokasi ke-i,  $k = 1, 2, \dots, p$

$\beta_0(u_i, v_i)$  : konstanta regresi untuk model GWR

$\beta_k(u_i, v_i)$  : koefisien regresi untuk predictor ke-k

$(u_i, v_i)$  : koordinat geografis (lintang, bujur) pada lokasi ke-i

$\varepsilon_i$  : residual pada pengamatan ke-i.  $\varepsilon \sim (0, \sigma^2)$ .

untuk mengestimasi nilai  $\beta$ , digunakan *Ordinary Least Square* (OLS):

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (3)$$

dengan

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (4)$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (5)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad (6)$$

#### 1.5.4. Geographically and Temporally Weighted Regression

*Geographically and Temporally Weighted Regression* (GTWR) merupakan pengembangan dari model GWR yang turut mempertimbangkan elemen spasial dan waktu (Rahmawati dkk., 2023). GTWR memperhitungkan aspek spasial temporal ke dalam estimasi parameter dengan menggunakan matriks pembobot yang dihasilkan dari jarak antar koordinat ( $u, v, t$ ) dari observasi ke- $i$  dengan observasi lainnya (Musdalifah dkk., 2023). Model GTWR dijabarkan sebagai berikut:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (7)$$

direpresentasikan ke dalam bentuk matriks menjadi:

$$\hat{y}_i = x_i^T \hat{\boldsymbol{\beta}}_k(u_i, v_i, t_i) = x_i^T \mathbf{W}(u_i, v_i, t_i) (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (8)$$

$\mathbf{W}(u_i, v_i, t_i)$  adalah matriks pembobot spasial sebagai berikut:

$$\mathbf{W}(u_i, v_i, t_i) = \begin{bmatrix} w_0(u_i, v_i, t_i) & 0 & \dots & 0 \\ 0 & w_1(u_i, v_i, t_i) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_n(u_i, v_i, t_i) \end{bmatrix} \quad (9)$$

Matriks  $W_{ij}$  biasanya diperoleh melalui fungsi kernel. Fungsi ini digunakan untuk menyesuaikan kepadatan dari observasi. Penentuan pembobot spasial yang digunakan sangat penting karena pembobot dapat menjelaskan besarnya pengaruh spasial antar suatu daerah. Pembobot diperoleh dari fungsi kernel (Oktarina dkk., 2024). Terdapat beberapa fungsi kernel yang dapat digunakan, antara lain:

1. Fungsi Kernel *Gaussian*

$$w_j(u_i, v_i, t_i) = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}^{ST}}{h} \right)^2 \right] \quad (10)$$

2. Fungsi Kernel *Exponential*

$$w_j(u_i, v_i, t_i) = \exp \left[ -\left( \frac{d_{ij}^{ST}}{h} \right) \right] \quad (11)$$

3. Fungsi Kernel *Bisquare*

$$w_j(u_i, v_i, t_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}^{ST}}{h}\right)^2\right)^2, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \quad (12)$$

#### 4. Fungsi Kernel *Tricube*

$$w_j(u_i, v_i, t_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}^{ST}}{h}\right)^3\right)^3, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \quad (13)$$

Keterangan:

$w_j(u_i, v_i)$  : Nilai fungsi kernel pada lokasi ke- $i$  dan lokasi ke- $j$

$h$  : Lebar bandwidth

$d_{ij}^{ST}$  : Jarak antara lokasi ke- $i$  dan lokasi ke- $j$

*Bandwidth* ( $h$ ) adalah ukuran jarak dalam fungsi pembobot yang menunjukkan hasil kalibrasi lokal yang dihaluskan. Pemilihan *bandwidth* optimum menjadi hal yang penting karena dapat mempengaruhi hasil pendugaan dari model (Haryanto dkk., 2019). Salah satu metode dalam mendapatkan nilai *bandwidth* optimum adalah metode *Cross Validation* (CV) dengan rumus:

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(h))^2 \quad (14)$$

Keterangan:

$n$  : Banyaknya pengamatan

$y_i$  : Pengamatan ke- $i$

$\hat{y}_{\neq i}(h)$  : Nilai duga pengamatan ke- $i$  yang didapat tanpa pengamatan ke- $i$

Nilai *bandwidth* optimum didapatkan dari nilai  $h$  yang menghasilkan nilai CV minimum.  $d_{ij}$  merepresentasikan jarak antara titik ke- $i$  dan titik ke- $j$ . Perhitungan jarak menggunakan jarak *Euclidean* dengan persamaan:

$$d_{ij}^S = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (15)$$

$$d_{ij}^T = \sqrt{(t_i - t_j)^2} \quad (16)$$

Jarak spasial temporal dapat dikonstruksi sebagai kombinasi linear dari jarak  $d_{ij}^S$  dan  $d_{ij}^T$ :

$$d_{ij}^{ST} = \sqrt{\mu^S [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu^T (t_i - t_j)^2} \quad (17)$$

$u_i$  menunjukkan koordinat *latitude* pada lokasi ke- $i$ , dan  $v_i$  menunjukkan koordinat *longitude* pada lokasi ke- $i$ .  $\mu^S$  merepresentasikan faktor ukuran dari jarak spasial dan  $\mu^T$  merepresentasikan faktor ukuran dari jarak temporal. Sehingga  $W_{ij}$  dapat diekspresikan dalam bentuk:

$$W_{ij} = \exp \left\{ - \left( \frac{\mu^S [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu^T (t_i - t_j)^2}{h_{ST}^2} \right) \right\}$$

$$\begin{aligned}
&= \exp \left\{ - \left( \frac{(u_i - u_j)^2 + (v_i - v_j)^2}{h_S^2} + \frac{(t_i - t_j)^2}{h_T^2} \right) \right\} \\
&= \exp \left\{ - \left( \frac{(d_{ij}^S)^2}{h_S^2} + \frac{(d_{ij}^T)^2}{h_T^2} \right) \right\} \\
&= W_{ij}^S \times W_{ij}^T
\end{aligned}$$

dengan  $h_{ST}^2$ ,  $h_S^2 = h_{ST}^2 / \mu^S$  dan  $h_T^2 = h_{ST}^2 / \mu^T$  adalah parameter dari bandwidth spasial temporal, bandwidth spasial, dan bandwidth temporal. Sehingga apabila nilai bandwidth spasial dan bandwidth temporal telah diperoleh, maka matriks pembobot  $W(u_i, v_i, t_i)$  dan  $\beta(u_i, v_i, t_i)$  juga dapat diperoleh (Ania dkk., 2024).

Wu (2014) mengembangkan Improved GTWR (IGTWR) dengan jarak spasial temporal sebagai berikut:

$$d_{ij}^{ST} = \begin{cases} \mu^S d_{ij}^S + \mu^T d_{ij}^T + 2 \sqrt{\mu^S \mu^T d_{ij}^S d_{ij}^T} \cos(\xi), & t_j < t_i \\ \infty, & t_j \geq t_i \end{cases} \quad (18)$$

dengan  $\xi \in [0, \pi]$ . Parameter  $\mu^S$ ,  $\mu^T$ , dan  $\xi$  perlu disesuaikan untuk menghasilkan output optimal.

### 1.5.5. Autokorelasi Spasial dan Temporal

Autokorelasi spasial menunjukkan adanya korelasi dari suatu wilayah terhadap wilayah lainnya. Untuk melihat nilai autokorelasi spasial secara global, digunakan Indeks Moran dengan hipotesis:

$H_0$ : tidak terjadi autokorelasi

$H_1$ : terjadi autokorelasi

Memperoleh nilai indeks Moran dapat dilakukan dengan menggunakan Persamaan (20):

$$I = \frac{n \sum_{i \neq j} w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{S_0 \sum_{i=1}^n (y_i - \bar{y})^2} \quad (19)$$

dengan:

$$\begin{aligned}
S_0 &= \sum_{i=1}^n \sum_{j=1}^n w_{ij}; S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2; \\
S_2 &= \sum_{i=1}^n \left( \sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji} \right)^2
\end{aligned}$$

Statistik Uji dari indeks moran adalah:

$$|Z_{hitung}| = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1) \quad (20)$$

dengan:

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

Apabila  $Z_{hitung} > Z_{tabel}$ , maka  $H_0$  ditolak. Maka data diasumsikan mengalami autokorelasi spasial.

Nilai indeks moran bernilai kisaran -1 sampai 1. Nilai 0 menandakan data tidak mengalami autokorelasi spasial, berarti daerah tersebut tidak bersifat kelompok dalam wilayahnya. Indeks moran bernilai positif menandakan daerah-daerah pada wilayah tersebut saling berpengaruh secara positif. Suatu daerah yang mengalami kenaikan nilai, akan meningkatkan nilai pada daerah tetangganya. Sedangkan indeks moran yang bernilai negatif, menandakan daerah-daerah pada wilayah tersebut saling berpengaruh negatif, yang berarti meningkatnya nilai suatu wilayah akan menurunkan nilai pada wilayah lainnya.

Perhitungan autokorelasi temporal dapat dilakukan menggunakan indeks Moran dengan pendekatan matriks pembobot  $w_{ij}$  yang didasarkan pada jarak temporal. Adanya autokorelasi temporal menunjukkan adanya ketergantungan suatu data terhadap datanya sendiri dari waktu lain. Apabila kedua efek autokorelasi terpenuhi, dapat disimpulkan bahwa data tersebut mengalami autokorelasi spasiotemporal.

### 1.5.6. Lagrange Multiplier Test

Autokorelasi atau efek ketergantungan dapat terjadi pada variabel respon (lag) dalam data. Untuk menguji efek ketergantungan pada variabel respon, digunakan Uji *Lagrange Multiplier* sebagai dasar untuk memilih model regresi spasial yang sesuai dengan kondisi ketergantungannya (LeSage & Pace, 2009). Pengujian LM pada lag dilakukan dengan hipotesis:

$$H_0: \rho = 0 \text{ (tidak ada autokorelasi dalam lag)}$$

$$H_1: \rho \neq 0 \text{ (terdapat autokorelasi dalam lag)}$$

Statistik Uji:

$$LM_{lag} = \frac{[(e^T \bar{W}_A y) / (e^T e / N)]^2}{\left[ (\bar{W}_A X \hat{\beta})^2 M(\bar{W}_A X \hat{\beta}) / (e^T e / N) \right] + [tr(\bar{W}_A^T \bar{W}_A + \bar{W}_A^2)]} \sim \chi_{(1-\alpha);1}^2 \quad (21)$$

Kriteria uji untuk kedua uji adalah menolak  $H_0$  ketika nilai  $LM_{lag}$  lebih besar dari  $\chi_{(1-\alpha);1}^2$ . Apabila  $LM_{lag}$  signifikan, maka model yang terbentuk adalah model *Spatial Autoregressive* (Azizah, 2023).

### 1.5.7. Spatial Autoregressive

Model *Spatial Autoregressive* (SAR) atau *Spatial Lag Model* (SLM) adalah model yang menggabungkan model regresi linear dengan spasial lag pada variabel respon dengan menggunakan data yang bertipe *cross section* (Anselin, 1988). Model SAR memberikan interpretasi pada data yang mengalami autokorelasi pada lag. Persamaan model SAR dapat dituliskan sebagai berikut:

$$y = \rho \bar{W} y + X \beta + \varepsilon \quad (22)$$

dengan  $\varepsilon \sim N(0, \sigma^2)$

Keterangan:

$\rho$  : Koefisien korelasi spasial lag

$\bar{W}$  adalah matriks pembobot spasial yang digunakan dalam memodelkan regresi spasial seperti model SAR. Berbeda dengan matriks pembobot dalam model GTWR, matriks pembobot dari model SAR tidak menggunakan fungsi kernel.



Nilai yang digunakan dalam pembobotan jarak dalam model ini adalah *exponential distance* dengan rumus:

$$w_{ij}^* = \exp(-d_{ij}) \quad (23)$$

dengan  $d_{ij}$  adalah jarak antar lokasi ke- $i$  terhadap lokasi ke- $j$ . Sedangkan nilai pembobot *exponential distance* yang terstandarisasi dapat diperoleh dengan rumus:

$$w_{ij} = \begin{cases} \frac{w_{ij}^*}{\sum_{j=1}^n w_{ij}^*}, & i \neq j \\ 0, & i = j \end{cases} \quad (24)$$

Matriks pembobot yang terstandarisasi akan memiliki nilai dari rentang 0 hingga 1 sehingga tiap baris dalam pembobotnya menghasilkan jumlah nilai sama dengan 1.

### 1.5.8. Geographically and Temporally Weighted Autoregressive

Penanganan efek heterogenitas spasial temporal maupun efek autokorelasi spasial temporal secara bersamaan, Wu (2014) mengembangkan *Geographically and Temporally Weighted Autoregressive* (GTWAR), yang mengombinasikan IGTWR dengan model regresi berautokorelasi. Autokorelasi spasial diperoleh dengan menghitung spasial lag  $\sum_{j=1}^n W_{ij} y_j$  ke dalam regresi linear:

$$y_i = \rho_i \sum_{j=1}^n \bar{W}_{ij} y_j + \varepsilon_i, i = 1, 2, \dots, n \quad (25)$$

$\rho_i$  merepresentasikan parameter autokorelasi,  $\varepsilon_i$  menunjukkan galat acak dalam model, dan  $\bar{W}_{ij}$  merupakan nilai pembobot dalam matriks pembobot berukuran  $n \times n$ . Mengombinasikan model pada Persamaan (26) ke dalam model IGTWR, Wu (2014) memperkenalkan model GTWAR sebagai berikut:

$$y_i = \beta_0(u_i, v_i, t_i) + \rho(u_i, v_i, t_i) \sum_{j=1}^n \bar{W}_{ij} y_j + \sum_{k=1}^d \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i \quad (26)$$

$$\mathbf{y} = \boldsymbol{\rho}(u_i, v_i, t_i) \bar{\mathbf{W}} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (27)$$

dengan  $\rho(u_i, v_i, t_i)$  merupakan parameter autokorelasi pada observasi ke- $i$ .

### 1.5.9. Pencilan

Pencilan merupakan nilai observasi yang tampak berbeda dengan observasi lainnya pada sekumpulan data. Adanya nilai pencilan dapat memberikan adanya bias dalam mengestimasi parameter model (Erda dkk., 2019). Salah satu metode yang dapat digunakan dalam mendeteksi adanya pencilan dalam data adalah menggunakan *Moran's Scatterplot*. Metode ini akan memberikan sebuah grafik *scatterplot* yang menunjukkan adanya pencilan dalam data melalui empat kuadran:

Kuadran II <i>Low-High (Coldspot)</i>	Kuadran I <i>High-High</i>
Kuadran III <i>Low-Low</i>	Kuadran IV <i>High-Low (Hotspot)</i>

**Gambar 1.** Kuadran Moran Scatterplot

*Moran Scatterplot* dibagi menjadi 4 kuadran. Kuadran I disebut *High-High* (HH), menunjukkan daerah yang mempunyai pengamatan tinggi dikelilingi daerah yang mempunyai nilai pengamatan tinggi. Kuadran II disebut *Low-High* (LH), menunjukkan daerah yang mempunyai pengamatan rendah dikelilingi daerah yang mempunyai nilai pengamatan tinggi. Kuadran III disebut *Low-Low* (LL), menunjukkan daerah yang mempunyai pengamatan rendah dikelilingi daerah yang mempunyai nilai pengamatan yang rendah juga. Kuadran IV disebut *High-Low* (HL), menunjukkan daerah yang mempunyai pengamatan tinggi namun dikelilingi daerah yang mempunyai nilai pengamatan rendah.

Kuadran II dan IV mengindikasikan daerah observasi yang tinggi namun dikelilingi oleh daerah pengamatan yang rendah, begitu pula sebaliknya. Daerah-daerah pada kuadran ini menunjukkan pola observasi yang berbeda dengan daerah tetangganya. Oleh karena itu, observasi yang berada dalam kuadran ini dikategorikan sebagai pencilan spasial dalam data (Shekhar dkk., 2003).

#### 1.5.10. Regresi Robust

Regresi *robust* adalah metode regresi yang digunakan saat residual dari model tidak berdistribusi normal atau terdapat 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). Terdapat beberapa jenis estimasi parameter dari regresi *robust* seperti, *S-Estimator*, *M-Estimator*, dan *MM-Estimator* (Gasul dkk., 2023).

##### a) Regresi Robust S-Estimator

Metode *S-Estimator* merupakan metode regresi *robust* yang memiliki nilai *breakdown* yang tinggi. Metode ini mampu meminimumkan varians dari residual. *S-Estimator* didefinisikan sebagai:

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

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

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

dengan

$$\hat{\sigma} = \begin{cases} \frac{\text{median}|e_i - \text{median}(e_i)|}{0.6745}, & \text{untuk iterasi pertama} \\ \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i e_i^2}, & \text{untuk iterasi kedua dan seterusnya} \end{cases}$$

Estimator  $\hat{\beta}$  pada metode ini diperoleh dengan cara melakukan iterasi hingga diperoleh hasil yang konvergen.

### b) Regresi Robust M-Estimator

Regresi *robust* dengan *M-Estimator* dapat dilakukan dengan meminimumkan residual dengan persamaan:

$$\hat{\beta}_M(u_i, v_i, t_i) = \min_{\beta} \sum_{i=1}^n \rho(\varepsilon_i) = \min_{\beta} \sum_{i=1}^n \rho\left(y_i - \sum_{j=0}^p X_{ij}\beta_j(u_i, v_i, t_i)\right) = 0$$

Untuk mendapatkan skala pada *estimator* ini, biasanya dengan menyelesaikan persamaan:

$$\min_{\beta} \sum_{i=1}^n \rho(u_i) = \min_{\beta} \sum_{i=1}^n \rho\left(\frac{\varepsilon_i}{\hat{\sigma}}\right) = \min_{\beta} \sum_{i=1}^n \rho\left(\frac{y_i - \sum_{j=0}^p X_{ij}\beta_j(u_i, v_i, t_i)}{\hat{\sigma}}\right)$$

dengan  $u_i = \frac{\varepsilon_i}{\hat{\sigma}}$  dan  $\hat{\sigma}$  adalah skala estimasi *robust*. Nilai  $s$  yang sering digunakan adalah:

$$\hat{\sigma} = \frac{\text{median}|e_i - \text{median}(e_i)|}{0.6745}$$

Nilai konstanta 0.6745 membuat  $\hat{\sigma}$  mendekati estimasi tak bias dari  $\sigma$ , dan membuat residual berdistribusi normal (Haryanto dkk., 2019).

### c) Regresi Robust MM-Estimator

Metode *MM-Estimator* merupakan gabungan dari metode estimasi yang memiliki nilai *breakdown* yang tinggi (*S-Estimator*) dan efisiensi yang tinggi (*M-Estimator*). Metode ini memiliki tiga tahap perhitungan. Tahap pertama dilakukan dengan menghitung nilai estimasi parameter awal regresi dengan metode bernilai *breakdown point* tinggi seperti *S-Estimator*. Tahap kedua dilakukan dengan menghitung residual dan skala estimasi *robust* dengan metode *M-Estimator*. Tahap ketiga dilakukan dengan menghitung estimasi parameter akhir dengan metode *MM-Estimator* melalui persamaan:

$$\hat{\beta}_{MM}(u_i, v_i, t_i) = \min_{\beta} \sum_{i=1}^n \rho\left(\frac{\varepsilon_i}{\hat{\sigma}}\right) = \min_{\beta} \sum_{i=1}^n \rho\left(\frac{y_i - \sum_{j=0}^p X_{ij}\beta_j(u_i, v_i, t_i)}{\hat{\sigma}}\right)$$

dengan  $\rho$  adalah fungsi objektif. Fungsi ini digunakan untuk memperoleh fungsi pembobot dalam regresi *robust* (Hujjah Asianingrum & Djuraidah, 2020). Salah satu fungsi pembobot yang dapat digunakan adalah pembobot *Tukey's Bisquare* dengan persamaan:

$$w(u_i) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & \text{untuk } |u_i| \leq c \\ 0, & \text{untuk } |u_i| > c \end{cases} \quad (28)$$

Nilai  $c$  disebut sebagai *tuning constant* dengan nilai  $c = 1.547$  pada metode *S-Estimator* dan  $c = 4.685$  pada metode *M-Estimator* (Fox dan Weisberg, 2010).

### 1.5.11. Uji Signifikansi Parameter

Pengujian signifikansi parameter dilakukan untuk mengetahui variabel independen yang berpengaruh secara signifikan terhadap variabel dependen (L. N. Azizah, 2013). Hipotesis dan statistik uji dalam pengujian ini adalah:

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, \text{ untuk } k = 1, 2, \dots, p$$

$$t = \frac{b_k}{s(b_k)} \quad (29)$$

Keterangan:

$b_k$  : Penduga koefisien variabel ke- $k$

$s(b_k)$  : Standar deviasi untuk penduga koefisien variabel ke- $k$

Menolak  $H_0$  apabila nilai  $t$  lebih besar dibanding nilai tabel  $t_{(\alpha, n-p)}$ .

Pengambilan keputusan dapat dilakukan dengan melihat nilai signifikansi.

### 1.5.12. Uji Kelayakan Model

Metode yang digunakan dalam uji kelayakan model menggunakan nilai *Root Mean Square Error* (RMSE). RMSE adalah besaran untuk melihat *error* yang dihasilkan oleh model. Semakin kecil nilai RMSE, semakin bagus pula model regresi yang terbangun. Persamaan untuk mencari nilai RMSE adalah:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (30)$$

Pengujian kelayakan model dapat dilakukan dengan menghitung besaran Koefisien Determinasi ( $R_{Adj}^2$ ). Besaran koefisien determinasi menunjukkan kelayakan variabel-variabel independent dalam menjelaskan variabel dependen. Nilai koefisien determinasi dapat dicari dengan persamaan:

$$R_{Adj}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (31)$$

Selain itu, nilai *Akaike Information Criterion* (AIC) juga dapat mengukur kebaikan dan kesesuaian model dalam menduga parameter. Suatu model yang baik menunjukkan nilai AIC yang lebih kecil. Nilai AIC dapat dihitung melalui:

$$AIC = 2k - 2 \ln L(\beta) \quad (32)$$

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

### 1.5.13. Tingkat Pengangguran Terbuka

Tingkat Pengangguran Terbuka (TPT) adalah nilai yang menunjukkan persentase jumlah pengangguran terbuka terhadap jumlah angkatan kerja. Angkatan kerja merupakan penduduk yang berada dalam usia kerja, yakni berusia 15-65 tahun. Sedangkan pengangguran terbuka adalah penduduk yang tidak mempunyai pekerjaan dan mencari pekerjaan, penduduk yang tidak mempunyai pekerjaan dan mempersiapkan usaha, penduduk yang tidak mempunyai pekerjaan dan tidak sedang mencari pekerjaan karena merasa tidak mungkin mendapatkan pekerjaan, serta meraka yang sudah mempunyai pekerjaan namun belum mulai bekerja.

Berdasarkan data dari BPS, "mencari pekerjaan" adalah kegiatan seseorang yang sedang mencari pekerjaan pada saat dilakukan survei, yakni orang yang belum pernah bekerja dan sedang berusaha mendapatkan pekerjaan, orang yang sudah pernah bekerja namun karena suatu hal, orang tersebut berhenti atau diberhentikan dan sedang berusaha untuk mendapatkan pekerjaan. Kegiatan "mempersiapkan usaha" adalah suatu kegiatan yang dilakukan oleh seseorang dalam rangka mempersiapkan suatu pekerjaan baru, yang bertujuan untuk memperoleh keuntungan atas resiko sendiri, baik dengan dan/atau tanpa mempekerjakan karyawan dibayar maupun tak dibayar (BPS, 2023).

Definisi "tidak mencari pekerjaan karena merasa tidak mungkin mendapatkan pekerjaan" mencakup individu yang telah menyerah dalam mencari pekerjaan karena merasa tidak ada kesempatan yang sesuai atau mereka yang merasa tidak memiliki kualifikasi yang dibutuhkan. Kelompok ini termasuk dalam kategori pengangguran terselubung, yang juga mencerminkan tantangan yang lebih dalam dalam pasar tenaga kerja. Pengelompokan ini penting karena memberikan gambaran lebih lengkap tentang tingkat kesulitan yang dihadapi oleh angkatan kerja dalam mendapatkan pekerjaan yang layak.

Penting untuk dicatat bahwa TPT tidak hanya digunakan sebagai indikator ekonomi, tetapi juga sebagai dasar bagi pengambilan kebijakan publik. Pemerintah dan pembuat kebijakan dapat menggunakan data TPT untuk merumuskan dan mengevaluasi kebijakan ketenagakerjaan, pendidikan, serta program pelatihan dan peningkatan keterampilan. Jika TPT tinggi, pemerintah perlu mengembangkan program pelatihan kerja, insentif bagi perusahaan untuk menciptakan lapangan kerja, atau investasi dalam sektor-sektor yang dapat menyerap banyak tenaga kerja. Dengan demikian, analisis TPT dapat membantu mengarahkan upaya untuk mengurangi pengangguran dan meningkatkan kesejahteraan ekonomi secara keseluruhan (BPS, 2023).

## BAB II METODOLOGI PENELITIAN

### 2.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari Badan Pusat Statistik Provinsi Sulawesi Selatan tahun 2018-2022. Variabel yang digunakan dalam penelitian ini ditampilkan dalam Tabel 1.

**Tabel 1.** Variabel Respon dan Prediktor

Variabel	Keterangan	Deskripsi
$Y$	Tingkat Pengangguran Terbuka	Persentase jumlah pengangguran terhadap jumlah angkatan kerja. Angkatan kerja adalah penduduk usia 15-65 tahun.
$X_1$	PDRB	Jumlah nilai tambah atas barang dan jasa yang dihasilkan oleh berbagai unit produksi pada suatu wilayah dalam jangka waktu tertentu (biasanya satu tahun)
$X_2$	Tingkat Partisipasi Angkatan Kerja	Persentase banyaknya angkatan kerja terhadap banyaknya penduduk yang berumur sepuluh tahun ke atas.
$X_3$	Kepadatan Penduduk	Perbandingan jumlah penduduk yang tinggal dalam satu kilometer persegi wilayah.
$X_4$	Angka Melek Huruf	Perbandingan jumlah penduduk berumur $\geq 15$ tahun yang dapat membaca dan menulis kalimat sederhana dengan huruf latin dan atau huruf lainnya, terhadap jumlah penduduk umur $\geq 15$ tahun.

### 2.2. Analisis Data

Metode analisis yang digunakan dalam penelitian ini adalah regresi spatiotemporal dengan *Robust Geographically and Temporally Weighted Autoregressive (RGTWAR)* menggunakan *MM-Estimator*. Pengolahan data dibantu dengan perangkat lunak *R-Studio*. Tahapan analisis data dalam penelitian ini adalah sebagai berikut:

1. Melakukan eksplorasi awal pada data untuk melihat sebaran data melalui peta tematik dari sudut kewilayahannya.
2. Mengidentifikasi efek heterogenitas dari data dengan menggunakan uji *Breusch-Pagan* sesuai pada Persamaan (1).
3. Menentukan matriks jarak antarwilayah. Penentuan jarak pada penelitian ini menggunakan jarak *Euclidean* pada Persamaan (17).
4. Menentukan matriks pembobot *exponential distance* untuk mengidentifikasi efek autokorelasi dari data dengan menggunakan uji *Moran's I* pada Persamaan (20). Apabila hasil menunjukkan terdapat autokorelasi pada

data, analisis dilanjutkan dengan uji *Lagrange Multiplier* untuk melihat letak autokorelasi pada data sesuai Persamaan (21).

5. Menentukan matriks pembobot kernel untuk melihat kedekatan antar wilayah. Matriks pembobot yang digunakan pada penelitian ini menggunakan kernel *Fixed-Exponential* dari matriks jarak yang telah ditentukan sebelumnya.
6. Menentukan parameter  $\mu^S$ ,  $\mu^T$ ,  $\xi$ , dan  $h$  optimal dengan nilai *cross-validation*.
7. Menentukan nilai koefisien autokorelasi tiap observasi  $\rho(u_i, v_i, t_i)$  dengan memaksimalkan fungsi likelihood.
8. Melakukan pemodelan GTWAR dengan bobot dan nilai koefisien autokorelasi masing-masing observasi.
9. Melakukan pendeteksian pencilan pada residual model GTWAR menggunakan *moran's scatterplot*.
10. Mengestimasi parameter RGTWAR dengan *MM-Estimator*
11. Melakukan signifikansi parameter model regresi.
12. Mengevaluasi model yang dihasilkan dengan RMSE,  $R_{Adj}^2$ , dan AIC. Model yang lebih baik menunjukkan nilai RMSE dan AIC yang lebih kecil dan  $R_{Adj}^2$  yang lebih besar.
13. Menarik kesimpulan.