

MODEL REGRESI *ROBUST IMPROVED GEOGRAPHICALLY AND TEMPORALLY WEIGHTED* UNTUK ANGKA KEJADIAN DEMAM BERDARAH DENGUE DI SULAWESI SELATAN TAHUN 2016-2021

ROBUST IMPROVED GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION MODEL FOR DENGUE HEMORRHAGIC FEVER INCIDENCE IN SOUTH SULAWESI IN 2016-2021

SAMSIR ADITYA ANIA



**PROGRAM STUDI STATISTIKA
SEKOLAH PASCASARJANA
UNIVERSITAS HASANUDDIN
MAKASSAR**

2024



**MODEL REGRESI *ROBUST IMPROVED GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED* UNTUK ANGKA KEJADIAN DEMAM
BERDARAH DENGUE DI SULAWESI SELATAN TAHUN 2016-2021**

Tesis

sebagai salah satu syarat untuk mencapai gelar Magister

Program Studi Statistika

Disusun dan diajukan oleh :

SAMSIR ADITYA ANIA

H062212011

kepada

PROGRAM MAGISTER STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2024



Optimized using
trial version
www.balesio.com

TESIS

MODEL REGRESI *ROBUST IMPROVED GEOGRAPHICALLY AND TEMPORALLY WEIGHTED* UNTUK ANGKA KEJADIAN DEMAM BERDARAH DENGUE DI SULAWESI SELATAN TAHUN 2016-2021

SAMSIR ADITYA ANIA

H062212011

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin pada tanggal 24 Januari 2024 dan dinyatakan telah memenuhi syarat kelulusan.

Menyetujui,

Pembimbing Utama *uj>*

Dr. Nirwan, M.Si.

NIP. 19630306 198702 1 002

Pembimbing Pendamping

Dr. Erna Tri Herdiani, S.Si., M.Si.

NIP. 19750429 200003 2 001

Ketua Program Studi
Magister Statistika

Tri Herdiani, S.Si., M.Si.

50429 200003 2 001

Dekan Fakultas MIPA
Universitas Hasanuddin



Dr. Eng. Amiruddin, M.Si.

NIP. 19720515 199702 1 002



PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul "**Model Regresi *Robust Improved Geographically and Temporally Weighted* untuk Angka Kejadian Demam Berdarah Dengue di Sulawesi Selatan Tahun 2016-2021**" adalah benar karya saya dengan arahan dari komisi pembimbing Dr. Nirwan, M.Si. sebagai Pembimbing Utama dan Dr. Erna Tri Herdiani, S.Si, M.Si. sebagai Pembimbing Pendamping. 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 tesis ini. Sebagian dari isi tesis ini telah *accepted* di Jurnal ***Communications in Mathematical Biology and Neuroscience*** sebagai artikel dengan judul "***Robust Spatial Temporal Analysis with Improved Geographically and Temporally Weighted Regression Model of Dengue Incidence Rate***".

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 24 Januari 2024
Yang menyatakan,



Samsir Aditya Ania
NIM H062212011



UCAPAN TERIMA KASIH

Segala Puji Syukur kepada **Allah SWT**, Tuhan Yang Maha Esa, atas berkat, bimbingan dan kasih karunia-Nya yang dilimpahkan kepada penulis, serta shalawat dan salam semoga selalu tercurahkan kepada baginda tercinta, Nabi yang paling dimuliakan, pemimpin orang-orang bertakwa, **Muhammad SAW** yang dinanti-nantikan syafaatnya di akhirat kelak. Limpahan doa kepada keluarga serta sahabat Rasulullah SAW. *Alhamdulillah*, berkat rahmat dan karunia serta mukzizat-Nya sehingga penulis dapat menyelesaikan tesis ini sebagai salah satu syarat akademik untuk meraih gelar Magister Sains pada Program Studi Magister Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulisan tugas akhir ini tentunya tidak lepas dari bantuan berbagai pihak baik moril maupun materil. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang tak terhingga untuk Almarhum Ayah **La Ode Saninu Ania** dan Ibunda tercinta **Dra. Mimi Sumarni Muhidin** yang tak kenal lelah mendoakan, memberikan dukungan, dan selalu melimpahkan cinta dan kasih sayangnya kepada penulis sehingga mereka menjadi motivasi terbesar penulis dalam menyelesaikan tugas akhir ini. Untuk saudara-saudaraku tercinta **Sandy Agusta Ania, Sabrina Afrilia Ania, Hj. Shanaz Ariesta Ania, Samudra Airlangga Ania,** dan **L.M. Safaat Adicandra Ania** terima kasih telah memberikan semangat, motivasi, dukungan, dan doa yang diberikan kepada penulis. Penulis paling berterimakasih kepada **H. Herman Tamsil** telah membantu memfasilitasi penulis sehingga bisa melanjutkan pendidikan hingga jenjang magister.

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

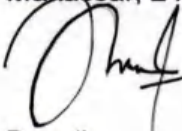
1. Bapak **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.** selaku Rektor Universitas Hasanuddin.
2. Bapak **Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. Ibu **Dr. Anna Islamiyati, S.Si., M.Si.** selaku Ketua Departemen Statistika rta seluruh jajarannya, serta segenap dosen pengajar dan staff artemen Statistika. Selaku Tim Penguji terimakasih atas saran dan kritik membangun dalam penyempurnaan penyusunan tugas akhir ini.



4. Ibu **Dr. Erna Tri Herdiani, S.Si, M.Si** selaku Ketua Program Magister Statistika sekaligus pembimbing pendamping yang telah bersedia meluangkan waktu dan pikirannya untuk membimbing dan memberi arahan dengan penuh kesabaran selama penulisan tugas akhir ini.
5. Ibu **Dr. Nirwan, M.Si.** selaku dosen pembimbing utama yang telah bersedia meluangkan waktu dan pikirannya untuk membimbing dan memberi arahan dengan penuh kesabaran selama penulisan tugas akhir ini.
6. Ibu **Prof. Dr. Nurtiti Sunusi, S.Si., M.Si** dan **Sri Astuti Thamrin, S.Si., M.Stat., Ph.D.** selaku Tim Penguji. Terima kasih telah meluangkan waktu dan telah memberikan kritik dan saran yang membangun dalam penyempurnaan penyusunan tugas akhir ini.
7. Teman-teman Pondok Hisyam, **Muhammad Ashar, Andi Arkan Alam Putra, Muh. Naim, Hasbianto,** dan **Abu Hurairah.** *Jazaakumullahu Khairan* masih selalu mengingatkan dalam kebaikan dan mengajarkan berbagai hal hingga sekarang. Semoga Allah *Subhanahu Wa Ta'ala* menjaga kita semua. Amin...
8. Bestie Lab yang teristimewa, Kak **Siswanto,** Kak **Hedi Kuswanto, Agung Muh. Takdir, Muhammad Idman, Ratmila, Ainun Utari, Andi Isna Yunita, Alimatun Najiha,** dan **Asnidar.** Dengan penuh rasa syukur untuk setiap cerita suka duka yang penuh warna dan segala bentuk dukungan kalian yang sangat berarti penulis ucapkan terimakasih.
9. Teman-teman Mahasiswa Program Magister Statistika angkatan I sampai angkatan VIII, terkhusus Kak **Masjidil Aqsha** dan Kak **Muhammad Fadil** terima kasih atas bantuan yang luar biasa kepada penulis.
10. Seluruh pihak yang sedikit banyaknya telah berpartisipasi dalam penulisan tugas akhir ini yang tidak dapat penulis sebutkan satu persatu.

Semoga Allah SWT memberikan balasan yang berlipat ganda kepada semuanya yang telah membantu dalam penulisan tugas akhir ini. Penulis berharap semoga tulisan ini dapat memberikan manfaat bagi semua pihak yang memerlukannya. Dengan segala kerendahan hati, penulis meminta maaf atas segala kekurangan dalam penulisan tugas akhir ini.

Makassar, 24 Januari 2024


Penulis



ABSTRAK

SAMSIR ADITYA ANIA. **Model Regresi *Robust Improved Geographically and Temporally Weighted* untuk Angka Kejadian Demam Berdarah Dengue di Sulawesi Selatan Tahun 2016-2021** (dibimbing oleh Nirwan dan Erna Tri Herdiani).

Demam Berdarah Dengue (DBD) masih menjadi masalah kesehatan masyarakat di dunia, terutama negara-negara tropis seperti Indonesia. Provinsi Sulawesi Selatan merupakan salah satu wilayah di Indonesia yang memiliki angka kejadian DBD tertinggi di tahun 2016 dan mengalami fluktuasi hingga terjadi peningkatan pada tahun 2021. Penyebaran angka kejadian DBD di Sulawesi Selatan ini mengalami keragaman setiap Kabupaten/Kota selama tahun 2016 hingga 2021. Analisis statistika yang dapat digunakan ketika data terjadi pengaruh keragaman secara spasial yaitu pemodelan Improved Geographically and Temporally Weighted Regression (IGTWR). Metode IGTWR mempertimbangkan aspek waktu (temporal) dan interaksi antara spasial-temporal yang menyebabkan jarak yang diukur dalam dimensi ruang (spasial) dapat berpengaruh pada jarak temporal. Pemodelan IGTWR tidak kekar terhadap data yang memiliki pencilan. Oleh karena itu, estimasi parameter model IGTWR menggunakan metode estimasi MM untuk mengatasi pencilan membentuk model Robust Improved Geographically and Temporally Weighted Regression (RIGTWR). Metode RIGTWR dapat digunakan untuk mengetahui faktor-faktor lingkungan sosial ekonomi yang dapat mempengaruhi angka kejadian DBD, yaitu kepadatan penduduk, persentase penduduk miskin, jumlah sarana kesehatan, dan persentase akses sanitasi layak. Hasil menunjukkan bahwa model RIGTWR memiliki nilai RMSE dan MAD terkecil dibandingkan model IGTWR yaitu masing-masing sebesar 37.2886 dan 15.3776 serta memiliki nilai Adjusted R-Square sebesar 75.65%. Pengujian parsial model RIGTWR menunjukkan faktor terbesar mempengaruhi angka kejadian DBD adalah persentase penduduk miskin dengan arah pengaruh positif dan jumlah sarana kesehatan dengan arah pengaruh negatif. Faktor yang dominan signifikan di setiap wilayah dan waktu dipengaruhi oleh persentase akses sanitasi layak.

Kata Kunci: analisis spasial, pencilan spasial, estimasi MM, RIGTWR, angka kejadian DBD



ABSTRACT

SAMSIR ADITYA ANIA. **Robust Improved Geographically and Temporally Weighted Regression Model for Dengue Hemorrhagic Fever Incidence in South Sulawesi in 2016-2021** (supervised by Nirwan and Erna Tri Herdiani).

Dengue Hemorrhagic Fever (DHF) is still a public health problem in the world, especially in tropical countries like Indonesia. South Sulawesi Province is one of the regions in Indonesia that has the highest DHF incidence rate in 2016 and experienced fluctuations until an increase in 2021. The spread of the DHF incidence rate in South Sulawesi experienced diversity in each district / city from 2016 to 2021. Statistical analysis that can be used when data is affected by spatial diversity is Improved Geographically and Temporally Weighted Regression (IGTWR) modeling. The IGTWR method considers the time aspect (temporal) and the interaction between spatial-temporal which causes the distance measured in the space dimension (spatial) can affect the temporal distance. IGTWR modeling is not robust to data that has outliers. Therefore, the IGTWR model parameter estimation uses the MM estimation method to overcome outliers to form the Robust Improved Geographically and Temporally Weighted Regression (RIGTWR) model. The RIGTWR method was used to determine the socioeconomic environmental factors that can affect the dengue incidence rate, namely population density, percentage of poor people, number of health facilities, and percentage of access to proper sanitation. The results show that the RIGTWR model has the smallest RMSE and MAD values compared to the IGTWR model which is 37.2886 and 15.3776 respectively and has an Adjusted R-Square value of 75.65%. Partial testing of the RIGTWR model shows that the biggest factors affecting the dengue incidence rate are the percentage of poor people with a positive direction of influence and the number of health facilities with a negative direction of influence. The dominant significant factor in each region and time is influenced by the percentage of access to proper sanitation.

Keyword: spatial analysis, spatial outliers, MM estimation, RIGTWR, dengue incidence rate



DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
PERNYATAAN PENGAJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN TESIS.....	iv
UCAPAN TERIMA KASIH.....	v
ABSTRAK.....	vii
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Analisis Regresi.....	6
2.2 Pencilan.....	7
2.3 Regresi <i>Robust</i>	9
2.3.1 Estimasi S (<i>Scale</i>).....	9
2.3.2 Estimasi MM (<i>Method of Moment</i>).....	10
2.3.3 Analisis Regresi Spasial.....	11



2.4.1 Pembobotan Spasial	11
2.4.2 Ketergantungan Spasial	12
2.4.3 Keragaman Spasial	13
2.4.3.1 <i>Geographically Weighted Regression</i>	14
2.4.3.2 <i>Geographically and Temporally Weighted Regression</i>	15
2.4.3.3 <i>Improved Geographically and Temporally Weighted Regression</i>	17
2.4.3.4 <i>Robust Improved Geographically and Temporal Weighted Regression</i>	18
2.5 Pengujian Parameter Model.....	18
2.6 Demam Berdarah Dengue	20
2.7 Kerangka Konseptual	22
BAB III METODE PENELITIAN.....	23
3.1 Sumber Data.....	23
3.2 Identifikasi Variabel	23
3.3 Analisis Data.....	24
BAB IV HASIL DAN PEMBAHASAN.....	27
4.1 Estimasi Parameter	27
4.2 Eksplorasi Angka Kejadian DBD di Sulawesi Selatan.....	31
4.3 Hubungan Linear Variabel.....	33
4.4 Uji Keragaman Spasial dan Temporal	34
4.5 Pendeteksian Pencilan	36
4.6 Pembentukan Bobot Spasial	37
4.7 Analisis Model RIGTWR dengan Estimasi MM	40
4.8 Pengujian Keباikan Model.....	45
4.9 Perbandingan Pencilan	47
4.10 Hasil Pengujian Parsial Model RIGTWR.....	49
PENUTUP	51
Kesimpulan	51



5.2	Saran	52
	DAFTAR PUSTAKA.....	53
	LAMPIRAN.....	57



Optimized using
trial version
www.balesio.com

DAFTAR TABEL

Nomor urut	Halaman
Tabel 3.1 Identifikasi Data.....	23
Tabel 4.1 Analisis Deskripsi Data DBD di Sulawesi Selatan.....	31
Tabel 4.2 Hasil Uji Multikolinearitas	34
Tabel 4.3 Hasil Statistik Uji <i>Breusch-Pagan</i>	34
Tabel 4.4 Hasil Pendeteksian Pencilan Metode RWBP	36
Tabel 4.5 Nilai Parameter Model RIGTWR.....	38
Tabel 4.6 Hasil Pemodelan RIGTWR	41
Tabel 4.7 Hasil Analisis Kebaikan Model.....	46



DAFTAR GAMBAR

Nomor urut	Halaman
Gambar 2.1	Ilustrasi dari Pengembangan Jarak Spasial-Temporal..... 17
Gambar 2.2	Kerangka Konseptual 22
Gambar 4.1	Sebaran Data Angka Kejadian DBD di Sulawesi Selatan 32
Gambar 4.2	Plot Hubungan Linier antar Variabel Bebas dan Terikat 33
Gambar 4.3	Plot Angka Kejadian DBD Tahun 2016 - 2021 35
Gambar 4.4	Pendeteksian Pencilan <i>Deleted Studentized Residual</i> 36
Gambar 4.5	Peta sebaran koefisien kepadatan penduduk (X_1)..... 42
Gambar 4.6	Peta sebaran koefisien persentase penduduk miskin (X_2) 43
Gambar 4.7	Peta sebaran koefisien jumlah sarana kesehatan (X_3)..... 44
Gambar 4.8	Peta sebaran koefisien persentase akses sanitasi layak (X_4) 45
Gambar 4.9	Nilai Peluang Pencilan Model IGTWR dan RIGTWR 47
Gambar 4.10	Nilai Peluang Pencilan Model IGTWR dan RIGTWR 48
Gambar 4.11	Sebaran Pengaruh Variabel Bebas Pada Model RIGTWR 49



DAFTAR LAMPIRAN

Nomor urut	Halaman
Lampiran 1. Hubungan Linear Variabel Bebas Terhadap Variabel Terikat Tahun 2016 – 2021	58
Lampiran 2. Hasil Analisis RWBP Model IGTWR dan RIGTWR Tahun 2016 .	61
Lampiran 3. Hasil Analisis RWBP Model IGTWR dan RIGTWR Tahun 2017 .	62
Lampiran 4. Hasil Analisis RWBP Model IGTWR dan RIGTWR Tahun 2018 .	63
Lampiran 5. Hasil Analisis RWBP Model IGTWR dan RIGTWR Tahun 2019 .	64
Lampiran 6. Hasil Analisis RWBP Model IGTWR dan RIGTWR Tahun 2020 .	65
Lampiran 7. Hasil Analisis RWBP Model IGTWR dan RIGTWR Tahun 2021 .	66
Lampiran 8. Matrix Bobot Spasial.....	67
Lampiran 9. Matrix Bobot <i>Robust</i> Konvergen	68
Lampiran 10. Hasil Estimasi Parameter Beta RIGTWR Tahun 2016.....	69
Lampiran 11. Hasil Estimasi Parameter Beta RIGTWR Tahun 2017.....	70
Lampiran 12. Hasil Estimasi Parameter Beta RIGTWR Tahun 2018.....	71
Lampiran 13. Hasil Estimasi Parameter Beta RIGTWR Tahun 2019.....	72
Lampiran 14. Hasil Estimasi Parameter Beta RIGTWR Tahun 2020.....	73
Lampiran 15. Hasil Estimasi Parameter Beta RIGTWR Tahun 2021.....	74
Lampiran 16. Perbandingan Nilai DBD Aktual dengan Hasil Prediksi Model IGTWR dan RIGTWR Tahun 2016 – 2021	75



BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi spasial merupakan salah satu metode statistika yang digunakan untuk mengetahui hubungan antara variabel terikat dan satu atau lebih variabel bebas yang menggunakan data spasial. Data spasial adalah data yang berisi informasi atribut dan informasi wilayah geografis yang umumnya berbentuk peta dan memiliki koordinat geografis yaitu lintang dan bujur (Djuraidah, 2021). Regresi data spasial memiliki pengaruh spasial, seperti ketergantungan spasial dan keragaman spasial (Anselin, 1988). Ketergantungan spasial menunjukkan pada pengamatan di suatu wilayah bergantung pada pengamatan di wilayah yang bertetangga. Keragaman spasial menunjukkan pada nonstasioneritas dari hubungan antara variabel bebas dengan variabel terikat dalam ruang sehingga terjadi keragaman pada setiap wilayah pengamatan (Djuraidah, 2021). Pemodelan keragaman spasial antara lain menggunakan metode ekspansi spasial, *Geographically Weighted Regression* (GWR), dan spasial *regime* (A. Fotheringham, 2002).

Model keragaman spasial yang sederhana telah dikembangkan oleh banyak peneliti adalah GWR. GWR merupakan metode yang digunakan untuk menangani pengaruh keragaman spasial dengan membangun model regresi lokal pada setiap wilayah amatan yang hanya melibatkan satu waktu (Sholihin, 2018). Menurut A. S. Fotheringham et al. (2015) regresi data spasial dengan melibatkan pengamatan beberapa waktu merupakan hal yang penting dalam analisis spasial sehingga dimungkinkan untuk melakukan estimasi parameter yang lebih akurat. Penambahan unsur waktu ini dapat dilakukan dengan model *Geographically and Temporally Weighted Regression* (GTWR). Model GTWR memiliki keuntungan yaitu menghasilkan estimasi parameter yang bersifat lokal untuk setiap wilayah dan waktu, sehingga model lebih representatif (Huang et al., 2010). Model GTWR tidak memperhatikan interaksi jarak spasial-temporal, sehingga Wu et al. (2014)

mendasarkan model *Improved Geographically and Temporally Weighted on* (IGTWR).



Wu et al. (2014) menggabungkan model IGTWR dan model regresi autokorelasi untuk melakukan pengembangan model GWR agar mengatasi pengaruh non stasioner dan autokorelasi secara bersamaan yang diterapkan pada data harga perumahan di Kota Shenzhen, Cina pada tahun 2001 hingga 2008. Hasil penelitian memperoleh model IGTWR lebih baik dalam mengatasi keragaman spasial tanpa permasalahan autokorelasi dibandingkan dengan model MKT, GWR, dan GTWR. Sholihin (2018) meneliti tentang model IGTWR yang merupakan pengembangan model GTWR dengan menambahkan interaksi jarak spasial-temporal dan membangun program analisis untuk mengestimasi parameter model IGTWR menggunakan R-Studio yang diimplementasikan pada kasus pertumbuhan ekonomi di Jawa Tengah tahun 2011 hingga 2015. Hasil penelitiannya menunjukkan bahwa model IGTWR menghasilkan nilai RMSE dan AIC yang lebih kecil dibandingkan dengan model MKT dan GTWR. Sifriyani et al. (2022) menggunakan model IGTWR pada bidang kesehatan khususnya pada kasus Covid-19 di Kalimantan Timur. Hasil menunjukkan model IGTWR lebih efektif dalam memodelkan kasus Covid-19 di Kalimantan Timur dibandingkan dengan model MKT dan GTWR.

Penelitian sebelumnya Sholihin (2018) dan Sifriyani et al. (2022) melakukan estimasi parameter pada model IGTWR menggunakan metode *Weighted Least Squares* (WLS). Zhang & Mei (2011) berpendapat bahwa metode WLS tidak kekar (*robust*) terhadap pencilan data, sehingga apabila terdapat pencilan maka akan mengakibatkan terciptanya model regresi yang bias dan kekeliruan dalam menyimpulkan hubungan antar variabel. Yu & Yao (2017) mengusulkan beberapa metode yang *robust* untuk mengatasi adanya pencilan pada data yaitu estimasi M (*Maximum Likelihood Type*), estimasi LMS (*Least Median of Square*), estimasi LTS (*Least Trimmed Square*), estimasi S (*Scale*), dan estimasi MM (*Method of Moment*). Menurut Rahman & Widodo (2018) estimasi LMS, LTS, S, dan MM memiliki *breakdown point* yang tinggi dibandingkan dengan estimasi lainnya yaitu sebesar 50%. Sedangkan estimasi M memiliki *breakdown point* yang rendah tetapi efisiensinya tinggi dalam mengestimasi parameter yaitu sebesar 95%. Beberapa estimasi regresi *robust* yang telah disebutkan hanya estimasi MM yang memiliki *breakdown point* dan efisiensi tinggi yang masing-masing sebesar 50% dan 95% (Sifriyani et al., 2018). Selain itu, estimasi MM mampu mendeteksi pencilan pada data bebas dan terikat (Khotimah et al., 2020).



Beberapa peneliti sebelumnya menunjukkan bahwa data kasus DBD teridentifikasi adanya pencilan sehingga membutuhkan metode analisis yang khusus mengatasi kondisi pencilan (Idris et al., 2018; Saputro et al., 2017). Selain itu, Gani et al. (2022) melakukan analisis kecenderungan dan visualisasi data spasial dalam kasus DBD di Indonesia selama 2007 – 2022. Hasil analisisnya menunjukkan bahwa kasus DBD cenderung mengalami penurunan di Pulau Jawa dan cenderung mengalami kenaikan di Provinsi luar Pulau Jawa yang salah satunya adalah Provinsi Sulawesi Selatan. Menurut data Dinas Kesehatan Provinsi Sulawesi Selatan angka kejadian DBD tertinggi terjadi pada tahun 2016 sebesar 89 kasus per 100.000 penduduk dan seiring bertambahnya waktu angka kejadian DBD mengalami peningkatan dan penurunan hingga pada tahun 2021 angka kejadian DBD mengalami peningkatan dari tahun sebelumnya sebanyak 29 atau 30 kasus DBD per 100.000 penduduk meningkat menjadi 39 kasus per 100.000 penduduk.

Hal ini harus menjadi perhatian penting pemerintah dalam menekan penyebaran kasus DBD dengan mengetahui berbagai faktor yang mempengaruhi kejadian DBD. Menurut Arsin (2013) salah satu faktor penyebaran kasus DBD adalah faktor lingkungan khususnya lingkungan sosial ekonomi yaitu kepadatan penduduk, fasilitas kesehatan, kemiskinan, dan fasilitas sanitasi yang tidak memadai. Carmo et al. (2020) berpendapat bahwa faktor-faktor yang mempengaruhi angka kejadian DBD berkaitan tentang populasi, pendidikan, pendapatan, perumahan dan kerentanan sosial. Nabila & Yotenka (2021) menyebutkan faktor yang mempengaruhi kejadian DBD adalah tatanan kawasan sehat, persentase penduduk miskin, sarana kesehatan, dan kebijakan PHBS.

Berdasarkan uraian diatas, peneliti tertarik untuk membentuk model *Robust Improved Geographically and Temporally Weighted Regression* (RIGTWR) untuk menguji faktor-faktor yang berpengaruh terhadap angka kejadian DBD di Sulawesi Selatan. Angka kejadian DBD di Sulawesi Selatan menunjukkan terjadinya keragaman temporal dan penyebarannya diamati di setiap Kabupaten/Kota yang beragam sehingga dapat menyebabkan beberapa wilayah memiliki nilai yang berbeda dari wilayah lainnya dan juga angka kejadian DBD sering teridentifikasi

Oleh karena itu, peneliti melakukan penelitian dengan judul "**Model Robust Improved Geographically and Temporally Weighted untuk kejadian Demam Berdarah Dengue di Sulawesi Selatan Tahun 2016-**



1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka diperoleh permasalahan yang akan dikaji dalam penelitian ini sebagai berikut:

1. Bagaimana mengestimasi parameter model RIGTWR menggunakan metode estimasi MM?
2. Bagaimana membangun model RIGTWR dengan estimasi MM pada angka kejadian DBD untuk setiap Kabupaten/Kota Provinsi Sulawesi Selatan dalam kurun waktu 2016 - 2021?
3. Bagaimana menentukan faktor-faktor yang berpengaruh terhadap angka kejadian DBD untuk setiap Kabupaten/Kota Provinsi Sulawesi Selatan dalam kurun waktu 2016 - 2021?

1.3 Batasan Masalah

Batasan-batasan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah angka kejadian DBD di Sulawesi Selatan dalam kurun waktu 2016-2021.
2. Metode yang digunakan untuk mengatasi pencilan pada data adalah regresi *robust* dengan menggunakan estimasi MM yang merupakan gabungan dari metode dengan estimasi S yang memiliki nilai *breakdown* tinggi dan estimasi M yang memiliki efisiensi tinggi.
3. Fungsi pembobot spasial yang digunakan adalah *Tukey Bisquare* dengan jarak yang digunakan adalah jarak *euclidean* dan pembentukan fungsi pembobot *robust* menggunakan fungsi kernel *exponential*.

1.4 Tujuan Penelitian

Tujuan dalam penelitian ini adalah sebagai berikut:

1. Melakukan estimasi parameter model RIGTWR menggunakan metode estimasi MM.
2. Memodelkan RIGTWR dengan estimasi MM pada angka kejadian DBD untuk setiap Kabupaten/Kota Provinsi Sulawesi Selatan dalam kurun waktu 2016 –



entukan faktor-faktor yang berpengaruh signifikan terhadap angka kejadian DBD untuk setiap Kabupaten/Kota Provinsi Sulawesi Selatan dalam kurun waktu 2016 – 2021.

1.5 Manfaat Penelitian

Manfaat dalam penelitian ini adalah sebagai berikut:

1. Sebagai bentuk pengembangan ilmu pengetahuan di bidang statistika dan kesehatan tentang analisis regresi *robust* IGTWR dengan estimasi MM untuk mengatasi data yang mengandung pencilan dan menghasilkan informasi yang lebih akurat.
2. Sebagai bantuan kepada Pemerintah terkhusus pada Dinas Kesehatan Provinsi Sulawesi Selatan dan sektor terkait lainnya untuk dapat menggunakan informasi hasil penelitian sebagai bahan pertimbangan dalam perencanaan program pengendalian DBD pada setiap Kabupaten/Kota di Provinsi Sulawesi Selatan.



Optimized using
trial version
www.balesio.com

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan analisis yang mempelajari tentang hubungan antara satu atau lebih variabel bebas dengan satu variabel terikat (Fox, 2015). Analisis hubungan antara satu variabel bebas dan satu variabel terikat disebut analisis regresi linier sederhana, sedangkan analisis hubungan antara beberapa variabel bebas dengan satu variabel terikat disebut analisis regresi linier berganda (Rawlings et al., 1998). Secara umum model regresi linier berganda ditulis sebagai berikut (Fox, 2015):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i = \beta_0 + \sum_{k=1}^p \beta_k X_{ik} + \varepsilon_i \quad (2.1)$$

untuk $i = 1, 2, \dots, n$ dan $k = 1, 2, \dots, p$

dalam notasi matriks, dapat ditulis sebagai berikut:

$$Y = X\beta + \varepsilon \quad (2.2)$$

dengan

$$Y = [y_1 \ y_2 \ \dots \ y_n]'$$

$$\beta = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_p]'$$

$$\varepsilon = [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n]'$$

$$X = [1 \ x_{i1} \ x_{i2} \ \dots \ x_{ip}] = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

dengan Y adalah vektor variabel terikat yang berukuran $n \times 1$ dengan n merupakan banyaknya pengamatan, β merupakan vektor parameter regresi atau koefisien variabel bebas yang berukuran $(p + 1) \times 1$ dengan p merupakan banyaknya variabel bebas, ε merupakan vektor galat yang berukuran $n \times 1$, dan X merupakan matriks variabel bebas yang berukuran $n \times (p + 1)$.



2.2 Pencilan

Pencilan merupakan suatu pengamatan pada kumpulan data yang memiliki pola berbeda dengan kumpulan data. Keberadaan pencilan dapat mengganggu model regresi yang mengakibatkan kekeliruan dalam pengambilan kesimpulan. Salah satu metode yang baik dalam mendeteksi pencilan dalam regresi linear adalah *studentized deleted residuals* (Arimie et al., 2020) dengan rumus berikut:

$$t_i = r_i \sqrt{\frac{n - p - 1}{n - p - r_i^2}}$$

dengan r_i merupakan *standardized residuals* ke- i seperti berikut:

$$r_i = \frac{\hat{\varepsilon}_i}{\sqrt{MSE(1 - h_{ii})}}$$

dengan $MSE = \sqrt{SSE/n - p - 1}$ adalah rata-rata kuadrat galat dan $h_i = \mathbf{x}'_i(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_i$ merupakan nilai *leverage* yang dapat dikatakan sebagai nilai pencilan yang terjadi pada variabel bebas. Nilai *leverage* yang dideteksi sebagai pencilan ketika nilai $h_i > 2p/n$ (Yan & Su, 2009). Sedangkan data dideteksi sebagai pencilan menggunakan metode *studentized deleted residuals* ketika nilai $|t_i| > t_{\frac{\alpha}{2}, n-p-1}$ (Arimie et al., 2020).

Sedangkan dalam regresi spasial metode yang baik dalam mendeteksi pencilan menggunakan pendekatan *Random Walk* (RW). Pendekatan ini ada dua metode yaitu RWBP (*Random Walk on Bipartite Graph*) dan RVEC (*Random Walk on Exhaustive Combination*). Metode RWBP lebih baik daripada metode RVEC karena RWBP disusun berdasarkan atribut spasial dan non spasial pada data dan dideteksi dengan menghitung jarak antara pengamatan spasial dengan tetangga terdekat melalui dua grafik yang terboboti. Metode RWBP dilakukan dengan langkah-langkah berikut (Liu et al., 2010):

1. Pembentukan Dua Grafik Terboboti

Pembentukan grafik terdiri dari tiga tahapan dasar. Pertama, mengelompokkan variabel non spasial berdasarkan metode penggerombolan K-
kedua, membentuk dua grafik yaitu bagian kiri berisi data spasial dan bagian kanan berisi data non spasial yang telah digerombolkan sebelumnya. Ketiga, menghitung nilai kemiripan pengamatan spasial dan kelompok (nilai *edge*)



berdasarkan variabel non spasial dari pengamatan spasial dan nilai pusat dari kelompok yang terbentuk sebagai berikut:

$$E(P_i, C_j) = \frac{1}{e^{|Atr(P_i) - Ctr(C_j)|^\alpha}, 0 < \alpha \leq 2$$

Keterangan:

$E(P_i, C_j)$: Nilai *edge* atau nilai kemiripan pengamatan spasial dan kelompok

$Atr(P_i)$: Variabel non spasial ke- i

$Ctr(C_j)$: Nilai pusat dari kelompok ke- j

α : Nilai konstanta yang diperoleh berdasarkan jangkauan dari nilai variabel non spasial

2. Menghitung Kemiripan antara Pengamatan Spasial

Perhitungan kemiripan antara pengamatan spasial dilakukan dengan menerapkan metode RW pada dua grafik terboboti. Metode RW dimulai dari simpul ke- i dan secara iterasi menuju tetangga terdekat dengan peluang tertentu dengan peluang sebesar c untuk kembali ke simpul awal.

$$s_p = (1 - c)(I - cW_p)^{-1} e_p$$

Keterangan:

s_p : Vektor peluang yang menjelaskan skor kemiripan antara titik ke- p dan titik lainnya

c : Konstanta yang bernilai 0.1

W_p : Matriks ketetanggaan dari titik ke- p yang telah dinormalisasi

e_p : Nilai *edge* untuk titik ke- p

3. Menghitung skor kemiripan untuk setiap pasangan pengamatan spasial dengan rumus:

$$Sim(p_i, p_j) = \frac{(s_{p_i}, s_{p_j})}{\sqrt{s_{p_i}, s_{p_i}} \times \sqrt{s_{p_j}, s_{p_j}}}$$

Keterangan:

s_{p_i} : Vektor peluang yang menjelaskan skor kemiripan antara titik ke- p dan titik lainnya pada wilayah ke- i

s_{p_j} : Vektor peluang yang menjelaskan skor kemiripan antara titik ke- p dan titik lainnya pada wilayah ke- j



- Mengidentifikasi pencilan spasial berdasarkan nilai kemiripan dimana nilai kemiripan terkecil akan dideteksi sebagai pencilan. Selain itu, *software R* menyediakan *package* yang dapat menghitung nilai peluang berdasarkan nilai kemiripan.

2.3 Regresi *Robust*

Regresi *robust* merupakan metode yang mempertimbangkan data pencilan dalam pembentukan model. Metode ini digunakan untuk menganalisis model regresi dengan distribusi galat yang tidak normal atau pada beberapa kondisi data mengandung pencilan yang mempengaruhi model regresi yang dihasilkan (Zhang & Mei, 2011). Penggunaan metode regresi *robust* menghasilkan model yang bersifat *robust* atau resisten terhadap pencilan (Shodiqin et al., 2018). Estimasi yang bersifat resisten artinya estimasi tersebut relatif tidak terpengaruh ketika terjadi perubahan besar pada bagian kecil data atau sebaliknya, ketika terdapat perubahan kecil pada bagian besar data (Setiarini & Listyani, 2017).

Yu & Yao (2017) mengusulkan beberapa metode regresi *robust* yang resisten terhadap adanya pencilan yaitu estimasi M, estimasi LMS, estimasi LTS, estimasi S, dan estimasi MM. Penelitian ini menggunakan estimasi MM yang memiliki *breakdown point* dan efisiensi tinggi.

2.3.1 Estimasi S (*Scale*)

Estimasi S pertama kali diperkenalkan oleh Rousseeuw & Yohai (1984) dengan *breakdown point* yang dapat mencapai hingga 50%. Estimasi S menggunakan standar deviasi galat untuk mengatasi kelemahan median. Menurut Yu & Yao (2017) estimasi S didefinisikan sebagai berikut:

$$\hat{\beta}_S = \arg \min \sum_{i=1}^n \rho \left(\frac{y_i - X_i \hat{\beta}(u_i, v_i, t_i)}{\hat{\sigma}_{WLS}} \right) \quad (2.3)$$

dengan ρ adalah fungsi objektif *Tukey Bisquare* sesuai Persamaan 2.4.

$$\rho(z_i) = \begin{cases} \frac{z_i^2}{2} - \frac{z_i^4}{2c^2} + \frac{z_i^6}{6c^4} & , |z_i| \leq c \\ \frac{c^2}{6} & , |z_i| > c \end{cases} \quad (2.4)$$



Equation 2.3 can be done by finding its derivative $\hat{\beta}_{WLS}$ so that it is obtained:

$$\sum_{i=1}^n x_{ij} \psi \left(\frac{y_i - X_i \hat{\beta}(u_i, v_i, t_i)}{\hat{\sigma}_{WLS}} \right) = 0$$

dengan ψ disebut fungsi pengaruh yang merupakan turunan dari fungsi objektif Persamaan 2.4. Sehingga diperoleh fungsi pembobot pada Persamaan 2.5.

$$w_i^* = \frac{\psi(z_i)}{z_i} = \begin{cases} \left(1 - \left(\frac{z_i}{c}\right)^2\right)^2 & , |z_i| \leq c \\ 0 & , |z_i| > c \end{cases} \quad (2.5)$$

dengan nilai c disebut *tuning constant* sebesar 1.547, $z_i = \varepsilon_i / \hat{\sigma}_{WLS}$, dan $\hat{\sigma}_{WLS}$ merupakan variansi galat dengan rumus berikut:

$$\hat{\sigma}_{WLS} = \begin{cases} \frac{\text{median}|\varepsilon_i - \text{median}(\varepsilon)|}{0.6745} & , \text{iterasi} = 1 \\ \sqrt{\frac{1}{nK} \sum_{i=1}^n w_i^* \varepsilon_i^2} & , \text{iterasi} > 1 \end{cases}$$

2.3.1 Estimasi MM (*Method of Moment*)

Estimasi MM diperkenalkan oleh Yohai (1987) dengan menggabungkan estimasi S yang memiliki *high breakdown point* sebesar 50% dan estimasi M yang memiliki efisiensi tinggi mencapai 95%. Estimasi MM diawali dengan pemodelan menggunakan estimasi S. Kemudian hasil estimasi S digunakan sebagai nilai awal dalam pemodelan estimasi M sehingga membentuk model estimasi MM. Pada umumnya digunakan fungsi objektif *Tukey Bisquare* baik pada estimasi S maupun estimasi M (Khotimah et al., 2020). Estimasi MM merupakan estimasi yang meminimumkan fungsi objektif dari galat sesuai Persamaan 2.6 (Susanti et al., 2014):

$$\hat{\beta}_{MM} = \arg \min \sum_{i=1}^n \rho \left(\frac{y_i - X_i \hat{\beta}(u_i, v_i, t_i)}{\hat{\sigma}_S} \right) \quad (2.6)$$

dengan

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

dan menggunakan fungsi pembobot seperti pada Persamaan 2.5 dengan nilai *constant* $c = 4.685$.



2.4 Analisis Regresi Spasial

Analisis regresi spasial adalah analisis regresi yang digunakan untuk tipe data spasial atau data yang memiliki efek spasial (*spatial effect*). Regresi spasial didasari dari berkembangnya metode regresi linier berganda. Pengembangan itu berdasarkan adanya pengaruh wilayah atau spasial pada data yang dianalisis (Anselin, 1988). Pada saat data memiliki komponen wilayah, terdapat dua masalah yang muncul yaitu ketergantungan spasial antar amatan dan keragaman spasial dalam proses pemodelan (LeSage, 1999). Komponen lain yang diperlukan dalam pemodelan spasial adalah pembobotan spasial.

2.4.1 Pembobotan Spasial

Pemodelan pada GWR sangat bergantung pada pembobot spasial yang digunakan. Menurut A. Fotheringham (2002), pembobot terbentuk dari fungsi kernel yang digunakan untuk mengestimasi parameter pada model GWR. Fungsi kernel umumnya mendefinisikan bobot sedemikian rupa sehingga data amatan yang letaknya dekat ke titik- i dalam ruang memiliki pengaruh yang lebih besar dari titik- i . Pembobot tersebut dilambangkan sebagai W yang merupakan diagonal matriks dari w_j (bobot tiap titik data dengan wilayah ke- j). Ada beberapa fungsi kernel yang dapat digunakan untuk pembobotan spasial, yaitu:

1. Fungsi Kernel *Gaussian*

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right)$$

2. Fungsi Kernel *Bisquare*

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^2\right)^2 & , \text{ untuk } d_{ij} < h \\ 0 & , \text{ untuk } d_{ij} \geq h \end{cases}$$

3. Fungsi Kernel *Eksponensial*

$$w_j(u_i, v_i) = \exp\left(-\left(\frac{d_{ij}}{h}\right)\right) \quad (2.7)$$

4. Fungsi Kernel *Tricube*

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^3\right)^3 & , \text{ untuk } d_{ij} < h \\ 0 & , \text{ untuk } d_{ij} \geq h \end{cases}$$



dengan d_{ij} merupakan jarak antara wilayah ke- i dengan wilayah ke- j yang dirumuskan sebagai berikut:

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

Menurut A. Fotheringham (2002), lebar jendela (*bandwidth*) yang dilambangkan dengan h merupakan ukuran jarak fungsi pembobot yang dapat mengukur sejauh mana pengaruh suatu wilayah dengan wilayah lainnya. Semakin besar h menunjukkan bobot yang terbentuk relatif lebih tinggi untuk daerah yang lebih terpencil, sedangkan semakin kecil h menunjukkan bobot yang terbentuk relatif lebih rendah untuk daerah yang lebih terpencil (Djuraidah, 2021). Pemilihan *bandwidth* yang optimal dapat dilakukan dengan melihat nilai *Cross Validation* (CV) minimum. Perhitungan CV menggunakan rumus berikut

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

dengan $\hat{y}_{\neq i}$ adalah nilai estimasi dari model GWR tanpa melibatkan wilayah ke- i .

2.4.2 Ketergantungan Spasial

Menurut Grekousis (2020) ketergantungan spasial adalah korelasi antar nilai amatan yang berkaitan dengan wilayah pada satu variabel yang sama. Jika terdapat pola sistematis dalam penyebaran sebuah variabel, maka terdapat ketergantungan spasial. Salah satu uji yang digunakan dalam pemeriksaan keberadaan ketergantungan spasial antar wilayah adalah statistik *Moran's I*. Hipotesis yang digunakan dalam uji ini adalah sebagai berikut:

$H_0: I = 0$ (Tidak terdapat ketergantungan spasial antar wilayah pengamatan)

$H_1: I \neq 0$ (Terdapat ketergantungan spasial antar wilayah pengamatan)

Menurut Djuraidah (2021) Indeks Moran memiliki rumus sesuai Persamaan 2.8.

$$I = \frac{n \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.8)$$

Statistik uji yang digunakan yaitu:

$$) = \frac{I - E(I)}{\sqrt{Var(I)}} \approx N(0,1) \quad (2.9)$$



dengan

$$E(I) = -\frac{1}{n-1}$$

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

$$S_0 = \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)$$

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

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

Keterangan:

- I = Indeks Moran
- n = banyaknya wilayah kejadian
- x_i = nilai pengamatan pada wilayah ke- i
- x_j = nilai pengamatan pada wilayah ke- j
- \bar{x} = rata-rata nilai pengamatan
- w_{ij} = elemen pada pembobot terstandarisasi antara daerah i dan j
- $E(I)$ = nilai ekspektasi dari I
- $Var(I)$ = nilai variansi dari I

Pengambilan keputusan pada uji ini adalah tolak H_0 pada taraf signifikansi α jika nilai $|Z(I)| > Z_{\frac{\alpha}{2}}$ atau $p\text{-value} < \alpha$ maka terdapat ketergantungan spasial antar wilayah pengamatan.

2.4.3 Keragaman Spasial

Keragaman spasial adalah efek spasial yang menggunakan pendekatan titik karena adanya keragaman di setiap wilayah yang disebabkan perbedaan karakteristik satu wilayah dengan wilayah lainnya. Pengujian keragaman spasial

akukan dengan uji *Breusch Pagan* (BP) (Anselin, 1988). Pengujian ini hipotesis sebagai berikut:



$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

(Tidak terdapat keragaman spasial antar wilayah pengamatan)

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ untuk } i, j = 1, 2, \dots, p$$

(Terdapat keragaman spasial antar wilayah pengamatan)

Statistik uji BP adalah

$$BP = \frac{1}{2} \mathbf{f}' \mathbf{Z} (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{Z}' \mathbf{f} \sim \chi_{(p)}^2 \quad (2.10)$$

dengan vektor \mathbf{f} adalah $f_i = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1 \right)$, \mathbf{Z} adalah matriks variabel terikat berukuran $n \times (p + 1)$ dan sudah dibakukan untuk setiap pengamatan dengan $i = 1, 2, \dots, n$, ε_i^2 adalah kuadrat tengah galat untuk pengamatan ke- i dan σ^2 merupakan variansi galat. Statistik uji BP menyebar χ_p^2 dengan p adalah banyaknya parameter regresi. Keputusan menolak H_0 , jika nilai BP lebih besar dari χ_p^2 maka terdapat keragaman spasial (Putra, 2019).

2.4.3.1 Geographically Weighted Regression

Model GWR merupakan pengembangan dari metode kuadrat terkecil dalam regresi klasik menjadi regresi terboboti spasial. Djuraidah (2021) menjelaskan bahwa keragaman spasial terjadi apabila nonstasionerotas dari hubungan variabel bebas dengan variabel terikat dalam wilayah. Hal ini mengakibatkan setiap wilayah pengamatan memiliki nilai parameter regresi yang berbeda-beda. Model GWR umumnya dapat ditulis pada Persamaan 2.11 (Nugroho & Sumarminingsih, 2021).

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (2.11)$$

Keterangan:

y_i : Nilai amatan variabel terikat untuk wilayah ke- i

u_i : Koordinat lintang (*latitude*) dari wilayah amatan ke- i

v_i : Koordinat bujur (*longitude*) dari wilayah amatan ke- i

$\beta_0(u_i, v_i)$: Intersep pada wilayah amatan ke- i

$\beta_k(u_i, v_i)$: Koefisien regresi variabel bebas ke- k pada wilayah amatan ke- i

: Nilai amatan variabel bebas ke- k pada wilayah amatan ke- i

: Galat wilayah pengamatan ke- i



2.4.3.2 Geographically and Temporally Weighted Regression

Model GTWR merupakan pendekatan yang efektif untuk menangani adanya keragaman spasial dan temporal (Huang et al., 2010). Model GTWR merupakan pengembangan dari model GWR dengan menambahkan unsur waktu (temporal). Berbeda dengan model GWR, GTWR menggabungkan informasi spasial dan temporal pada matriks pembobot dalam mengidentifikasi adanya keragaman spasial dan temporal. Model GTWR sebanyak p variabel bebas dengan variabel terikat y_i pada wilayah (u_i, v_i, t_i) untuk setiap amatan dituliskan pada Persamaan 2.12 (Sholihin, 2018):

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (2.12)$$

Keterangan:

$\beta_0(u_i, v_i, t_i)$: Intersep pada wilayah amatan (u_i, v_i) dan waktu amatan t_i

$\beta_k(u_i, v_i, t_i)$: Koefisien regresi variabel bebas ke- k pada wilayah amatan ke- i dan waktu amatan t_i

Estimasi parameter model GTWR $\hat{\beta}_k(u_i, v_i, t_i)$ untuk variabel bebas sebanyak k pada titik ke- i diperoleh menggunakan metode WLS dengan memasukkan bobot yang berbeda untuk setiap wilayah dan waktu. Estimasi parameter dapat dituliskan pada Persamaan 2.13.

$$\hat{\beta}(u_i, v_i, t_i) = [X^T W(u_i, v_i, t_i) X]^{-1} X^T W(u_i, v_i, t_i) Y \quad (2.13)$$

dengan $W(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in})$ dan $W(u_i, v_i, t_i)$ adalah matriks pembobot pada pengamatan (u_i, v_i) waktu ke- t_i dengan n adalah jumlah pengamatan. Fungsi jarak spasial-temporal (d_{ij}^{ST}) terdiri atas gabungan fungsi jarak spasial (d_{ij}^S) dan fungsi jarak temporal (d_{ij}^T), yang dituliskan pada Persamaan 2.14 (Huang et al., 2010).

$$\begin{cases} (d_{ij}^S)^2 = (u_i - u_j)^2 + (v_i - v_j)^2 \\ (d_{ij}^T)^2 = (t_i - t_j)^2 \\ (d_{ij}^{ST})^2 = \lambda [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu [(t_i - t_j)^2] \end{cases} \quad (2.14)$$



dengan λ dan μ digunakan sebagai parameter penyeimbang terhadap pengaruh antara wilayah dan waktu pada pengukuran jarak spasial-temporal. Persamaan 2.14 diperoleh:

$$\begin{aligned}
w_{ij} &= \exp \left\{ - \left(\frac{\lambda [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu (t_i - t_j)^2}{h_{ST}^2} \right) \right\} \\
&= \exp \left\{ - \left(\frac{[(u_i - u_j)^2 + (v_i - v_j)^2]}{h_S^2} + \frac{(t_i - t_j)^2}{h_T^2} \right) \right\}
\end{aligned}$$

Misalkan $h_S^2 = \frac{h_{ST}^2}{\lambda}$ dan $h_T^2 = \frac{h_{ST}^2}{\mu}$, maka diperoleh hasil:

$$\begin{aligned}
w_{ij} &= \exp \left\{ - \left(\frac{(d_{ij}^S)^2}{h_S^2} + \frac{(d_{ij}^T)^2}{h_T^2} \right) \right\} \\
&= \exp \left\{ - \left(\frac{(d_{ij}^S)^2}{h_S^2} \right) \right\} \times \exp \left\{ - \left(\frac{(d_{ij}^T)^2}{h_T^2} \right) \right\} \\
&= w_{ij}^S \times w_{ij}^T
\end{aligned}$$

dengan $w_{ij}^S = \exp \left\{ - \left(\frac{(d_{ij}^S)^2}{h_S^2} \right) \right\}$ dan $w_{ij}^T = \exp \left\{ - \left(\frac{(d_{ij}^T)^2}{h_T^2} \right) \right\}$

Keterangan:

h_S : *Bandwidth* spasial

h_T : *Bandwidth* temporal

h_{ST} : *Bandwidth* spasial-temporal

Misalkan τ merupakan parameter rasio dari $\tau = \frac{\mu}{\lambda}$ dengan $\lambda \neq 0$ maka diperoleh Persamaan 2.15 (Huang et al., 2010):

$$\frac{(d_{ij}^{ST})^2}{\lambda} = [(u_i - u_j)^2 + (v_i - v_j)^2] + \tau (t_i - t_j)^2 \quad (2.15)$$

Parameter τ berfungsi untuk memperbesar atau memperkecil rasio jarak temporal terhadap jarak spasial. Parameter ini didapatkan dari kriteria CV minimum melalui inialisasi nilai τ awal yang dituliskan pada Persamaan 2.16.

$$CV(\tau) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(\tau))^2 \quad (2.16)$$

Selanjutnya estimasi parameter λ dan μ bisa diperoleh dengan metode *iterative*

kan hasil estimasi τ yang menghasilkan CV minimum.



2.4.3.3 Improved Geographically and Temporally Weighted Regression

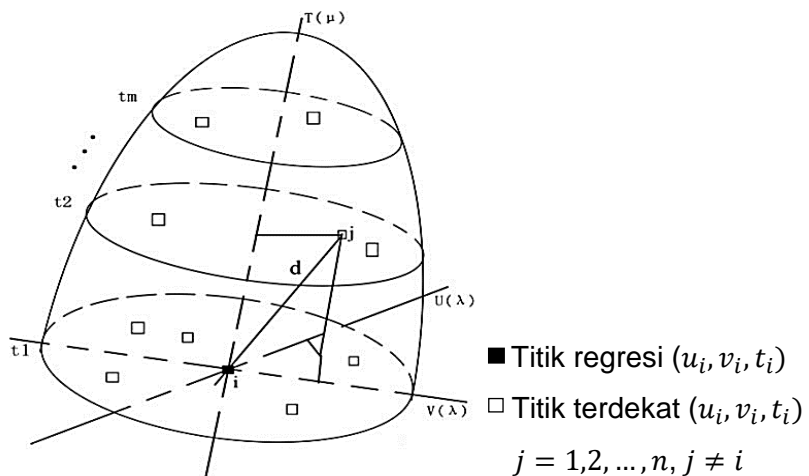
Matriks pembobot yang digunakan pada model GTWR menggunakan operator sederhana yaitu penjumlahan, operator tersebut digunakan untuk mengukur jarak spasial-temporal dengan kombinasi linier antara jarak spasial dan jarak temporal yang ditunjukkan pada Persamaan 2.17 (Huang et al., 2010).

$$d_{ST}^2 = d_S^2 \otimes d_T^2 = \lambda d_S^2 + \mu d_T^2 \quad (2.17)$$

dengan λ dan μ adalah parameter yang disesuaikan untuk menyeimbangkan pengaruh skala yang digunakan untuk mengukur jarak spasial dan temporal dalam masing-masing sistem koordinat. Spesifikasi tersebut mengasumsikan bahwa sistem koordinat spasial-temporal bersifat tegak lurus. Dengan demikian, jarak yang diukur dalam dimensi ruang tidak berpengaruh pada jarak temporal sehingga tidak sesuai untuk pemodelan interaksi spasial-temporal. Wu et al. (2014) mendefinisikan operator ' \otimes ' yang lebih kompleks sebagai penyusun model IGTRW sesuai Persamaan 2.18.

$$\begin{cases} d_{ij}^{ST} = d_{ij}^S \otimes d_{ij}^T = \lambda d_{ij}^S + \mu d_{ij}^T + 2\sqrt{\lambda\mu d_{ij}^S d_{ij}^T} \cos(\xi) & , t_j \leq t_i \\ d_{ij}^{ST} = \infty & , t_j > t_i \end{cases} \quad (2.18)$$

dengan t_i dan t_j adalah waktu pengamatan pada wilayah ke- i dan ke- j . Parameter λ , μ , dan $\xi \in [0, \pi]$ adalah parameter penyeimbang yang diperoleh dengan mengoptimalkan koefisien determinasi melalui prosedur CV. Parameter ξ digunakan untuk mengukur interaksi pengaruh wilayah dan waktu. Gambar 2.1 menunjukkan contoh pengembangan jarak spasial-temporal.



Gambar 2.1 Ilustrasi dari Pengembangan Jarak Spasial-Temporal



2.4.3.4 Robust Improved Geographically and Temporal Weighted Regression

Regres *robust* pada model IGTWR merupakan pemodelan yang digunakan pada data yang terdapat pencilan. Model RIGTWR untuk wilayah dan waktu ke- i yang mengandung pencilan sesuai Persamaan 2.19.

$$\rho(y_i) = \rho \left[\beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \right] \quad (2.19)$$

dengan: $i = 1, 2, 3, \dots, n$ dan $x_{ik} = x_{i1}, x_{i2}, \dots, x_{ip}$

Regresi *robust* dengan estimasi MM dilakukan dengan meminimumkan fungsi objektif dari galat sesuai Persamaan 2.20.

$$\hat{\beta}_{MM} = \arg \min \sum_{i=1}^n \rho \left(\frac{y_i - X_i \hat{\beta}(u_i, v_i, t_i)}{\hat{\sigma}_s} \right) \quad (2.20)$$

Estimasi MM dilakukan dengan menggunakan metode *Iteratively Reweighted Least Square* (IRLS). Pada iterasi ini, nilai bobot w^* akan berubah nilainya di setiap iterasi, sehingga diperoleh Persamaan 2.21.

$$\hat{\beta}(u_i, v_i, t_i)^m = [X^T W(u_i, v_i, t_i) w^{*m} X]^{-1} X^T W(u_i, v_i, t_i) w^{*m} Y \quad (2.21)$$

dengan m merupakan banyaknya iterasi

Pada w^{*m} pembobot yang diberikan, diperoleh estimasi sesuai Persamaan 2.22.

$$\hat{\beta}(u_i, v_i, t_i)^{m+1} = [X^T W(u_i, v_i, t_i) w^{*m} X]^{-1} X^T W(u_i, v_i, t_i) w^{*m} Y \quad (2.22)$$

Perhitungan Persamaan 2.21 dan 2.22 akan terus berulang hingga diperoleh estimasi yang konvergen, yakni ketika selisih nilai $\hat{\beta}(u_i, v_i, t_i)^{m+1}$ dan $\hat{\beta}(u_i, v_i, t_i)^m$ lebih kecil dari 0.000 (Erda & Djuraidah, 2019).

2.5 Pengujian Parameter Model

Pengujian parameter model bertujuan untuk mengetahui parameter-parameter yang signifikan dalam mempengaruhi variabel terikat. Model GWR memiliki model di setiap wilayah, sehingga pengujian parameter dilakukan pada wilayah secara parsial. Hipotesis pengujian parameter adalah sebagai berikut (Liang et al., 2000):



$$H_0 : \beta_k(u_i, v_i) = 0$$

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

Estimasi parameter $\hat{\beta}(u_i, v_i)$ mengikuti sebaran normal multivariat dengan rata-rata $\beta(u_i, v_i)$ dan matriks varian-kovarian $C_i C_i^T \sigma^2$. Matriks C_i diperoleh dengan rumus sebagai berikut:

$$C_i = (X^T W(u_i, v_i))^{-1} X^T W(u_i, v_i)$$

Sehingga didapatkan bentuk sebaran normal baku, yaitu:

$$\frac{\hat{\beta}_k(u_i, v_i) - \beta_k(u_i, v_i)}{\sigma \sqrt{c_{kk}}} \sim N(0,1)$$

dengan c_{kk} merupakan elemen diagonal ke- k dari matriks $C_i C_i^T$.

Pada kondisi hipotesis alternatif (H_1), koefisien regresi yang beragam pada setiap wilayah yang ditentukan dengan model GWR. Jumlah Kuadrat Galat (JKG) yang diperoleh dari model GWR sesuai Persamaan 2.23.

$$\begin{aligned} JKG_{GWR} &= \hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}) \\ JKG_{GWR} &= \mathbf{y}^T (\mathbf{I} - \mathbf{L})^T (\mathbf{I} - \mathbf{L}) \mathbf{y} \end{aligned} \quad (2.23)$$

dengan matriks $\mathbf{L} = \mathbf{x}_i^T ([X^T W(u_i, v_i, t_i) X]^{-1} X^T W(u_i, v_i, t_i) \mathbf{y})$ sehingga statistik uji yang digunakan pada pengujian parameter secara parsial sesuai Persamaan 2.24 (Nakaya et al., 2005).

$$t_k(u_i, v_i) = \frac{\hat{\beta}(u_i, v_i)}{\hat{\sigma} \sqrt{c_{kk}}} \quad (2.24)$$

dengan $\hat{\sigma} = \sqrt{\frac{JKG_{GWR}}{\delta_1}}$ dan mengikuti sebaran t dengan derajat bebas $\left(\frac{\delta_1^2}{\delta_2^2}\right)$. Nilai JKG_{GWR} diperoleh seperti pada Persamaan 2.23, sedangkan nilai δ dihitung sesuai Persamaan 2.25.

$$\delta_i = \text{tr}([(I - L)^T (I - L)]^i), \text{ untuk } i = 1 \text{ dan } 2 \quad (2.25)$$

Pengambilan keputusan pada uji ini adalah tolak H_0 jika nilai $|t_k(u_i, v_i)| > t_{\alpha}(n - p - 1)$ atau jika $p\text{-value} < \alpha$ dengan p adalah banyaknya variabel bebas et al., 2005).



2.6 Demam Berdarah Dengue

Demam Berdarah Dengue (DBD) adalah penyakit infeksi yang disebabkan oleh virus dengue yang mengakibatkan demam akut. Penyakit DBD merupakan penyakit menular yang disebabkan oleh virus dengue dan ditularkan oleh nyamuk *Aedes Aegypti* kepada manusia. Faktor-faktor yang berperan dalam penyebaran DBD merupakan interaksi dinamis antara *host* (manusia), *agent* (virus), dan *environmet* (lingkungan). Faktor lingkungan adalah kondisi berpengaruh yang bukan bagian dari virus maupun manusia, tetapi mampu menginteraksikan antara virus dengan manusia. Terdapat beberapa faktor lingkungan yaitu lingkungan, fisik, kimia, biologi, dan sosial ekonomi (Arsin, 2013).

Alvarez et al. (2023) melakukan penelitian tentang faktor sosial ekonomi terhadap angka kejadian DBD. Penelitiannya menghasilkan kesimpulan bahwa daerah dengan sosial ekonomi rendah akan meningkatkan penyebaran demam berdarah dibandingkan dengan daerah dengan sosial ekonomi yang tinggi. Medronho & Queiroz (2021) berpendapat bahwa faktor-faktor yang mempengaruhi angka kejadian DBD, yaitu kepadatan penduduk, persentase wilayah perkotaan, pendapatan rata-rata kepala rumah tangga, dan persentase rumah tangga yang terhubung dengan pasoka air, pembuangan limbah yang memadai, serta mempunyai tempat pengumpulan sampah umum. Carmo et al. (2020) mengatakan faktor yang mempengaruhi angka kejadian DBD diantaranya berkaitan tentang populasi, pendidikan, pendapatan, perumahan dan kerentanan sosial.

Berdasarkan peneliti-peneliti sebelumnya, tesis ini membahas tentang faktor sosial ekonomi yang mempengaruhi angka kejadian DBD yaitu kepadatan penduduk, persentase penduduk miskin, jumlah sarana kesehatan, dan persentase akses sanitasi layak. Penjelasannya sebagai berikut:

1. Kepadatan penduduk merupakan indikator dari tekanan penduduk. Kepadatan penduduk di suatu wilayah dibandingkan dengan luas tanah yang ditempati dan dinyatakan sebagai banyaknya penduduk perkilometer persegi (km^2). Masalah kependudukan akan berdampak pada bidang sosial-ekonomi, politik, pertahanan, keamanan serta masalah kesehatan. Perubahan jumlah

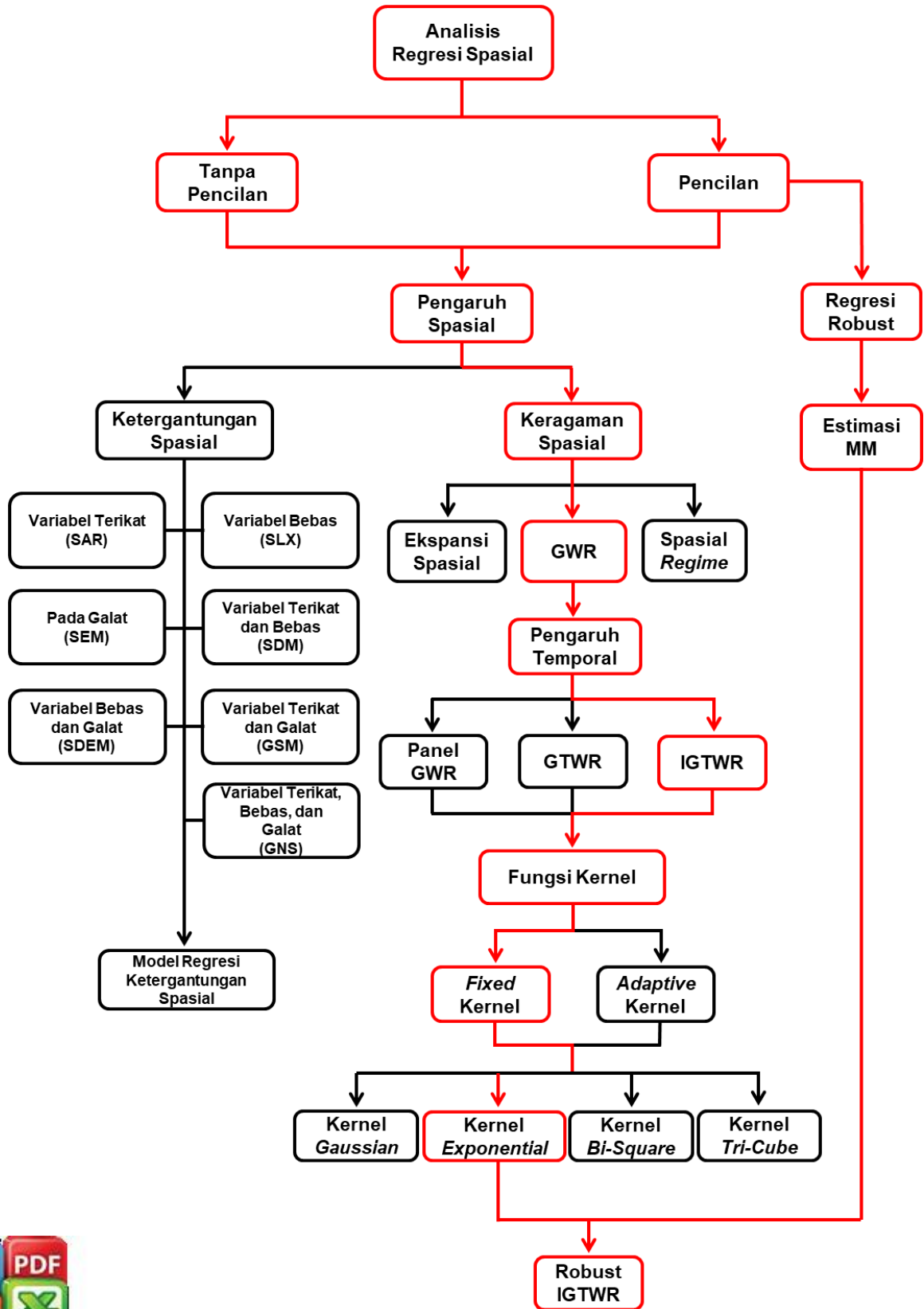


ilasi pada periode waktu tertentu merupakan salah satu ciri terjadinya perubahan poulasi di wilayah tersebut. Peningkatan populasi akan jakibatkan perubahan kepadatan penduduk.

2. Kemiskinan adalah suatu keadaan dimana ketidakmampuan suatu sisi ekonomi untuk memenuhi kebutuhan dasar pokok seperti makanan, pakaian dan tempat tinggal yang diukur dari sisi pengeluaran. Berdasarkan hal tersebut, penduduk miskin merupakan penduduk yang memiliki rata-rata pengeluaran perkapita per bulan lebih rendah dari garis kemiskinan (GK). Garis kemiskinan sendiri dapat diartikan sebagai suatu keadaan dimana nilai kebutuhan makanan yang disetarakan dengan 2100 kalori per kapita per hari dan kebutuhan minimum untuk pakaian maupun tempat tinggal kurang dari rata-rata pengeluaran per kapita per bulan.
3. Menurut Undang-Undang Republik Indonesia No. 23 Tahun 1992 tentang kesehatan bahwa sarana kesehatan merupakan tempat yang digunakan untuk menyelenggarakan upaya kesehatan seperti rumah sakit, klinik, puskesmas serta yang lainnya.
4. Akses sanitasi layak merupakan suatu fasilitas sanitasi yang memenuhi syarat kesehatan seperti tersedianya kloset menggunakan leher angsa dan tempat pembuangan akhir tinja dengan menggunakan tangki septik maupun Sistem Pengolahan Air Limbah (SPAL). Fasilitas sanitasi yang layak juga bisa digunakan oleh rumah tangga sendiri atau bersama dengan rumah tangga yang lain.



2.7 Kerangka Konseptual



Gambar 2.2 Kerangka Konseptual

