

**PEMODELAN *ROBUST GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED REGRESSION* DENGAN
M-ESTIMATOR PADA PENDERITA PNEUMONIA
USIA BALITA DI INDONESIA**

SKRIPSI



MUSDALIFAH

H051181024

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
OKTOBER 2022**

**PEMODELAN *ROBUST GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED REGRESSION* DENGAN
M-ESTIMATOR PADA PENDERITA PNEUMONIA
USIA BALITA DI INDONESIA**

SKRIPSI

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

MUSDALIFAH

H051181024

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

**MAKASSAR
OKTOBER 2022**

LEMBAR PERNYATAAN KEOTENTIKAN

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

Pemodelan *Robust Geographically and Temporally Weighted Regression* dengan *M-Estimator* pada Penderita Pneumonia Usia Balita di Indonesia

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

Makassar, 4 Oktober 2022



Musdantah

NIM H051181024

**PEMODELAN ROBUST GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED REGRESSION DENGAN M-
ESTIMATOR PADA PENDERITA PNEUMONIA USIA BALITA
DI INDONESIA**

Disetujui Oleh:

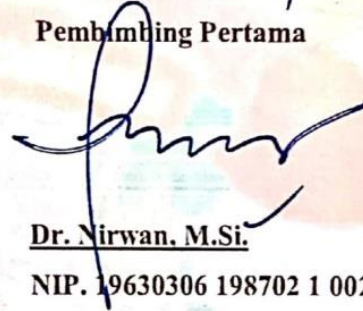
Pembimbing Utama



Siswanto, S.Si., M.Si.

NIP. 19920107 201903 1 012

Pembimbing Pertama



Dr. Nirwan, M.Si.

NIP. 19630306 198702 1 002

Ketua Program Studi



Dr. Nurhid Sunusi, S.Si., M.Si.

NIP. 19720117 199703 2 002

Pada 4 Oktober 2022

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Musdalifah

NIM : H051181024

Program Studi : Statistika

Judul Skripsi : *Pemodelan Robust Geographically and Temporally Weighted Regression dengan M-Estimator pada Penderita Pneumonia Usia Balita di Indonesia*

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

DEWAN PENGUJI

1. Ketua : Siswanto, S.Si., M.Si. (.....)
2. Sekretaris : Dr. Nirwan, M.Si. (.....)
3. Anggota : Andi Kresna Jaya, S.Si., M.Si. (.....)
4. Anggota : Dra. Nasrah Sirajang, M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 4 Oktober 2022

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji hanya milik Allah *Subhanallahu Wa Ta'ala* atas segala limpahan rahmat, nikmat dan hidayah-Nya yang telah diberikan kepada penulis sampai saat ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa sallam*, kepada para keluarga, tabi'in, tabi'ut tabi'in, serta orang-orang sholeh yang haq hingga kadar Allah berlaku atas diri-diri mereka. *Alhamdulillahirobbil'alamiin*, berkat rahmat dan kekuatan yang diberikan oleh Allah *Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi yang berjudul "***Pemodelan Robust Geographically and Temporally Weighted Regression dengan M-Estimator pada Penderita Pneumonia Usia Balita di Indonesia***" sebagai salah satu syarat memperoleh gelar sarjana pada Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis mampu sampai di titik ini berkat dukungan dan bantuan dari pihak-pihak yang senantiasa berada di sisi penulis. Oleh karenanya, penulis haturkan rasa terima kasih yang tulus dari dalam hati, terutama untuk orang tua penulis, Ayahanda **Haeruddin** dan Almarhumah Ibunda tersayang, **Hasbiah** yang telah menjadi pendukung nomor satu, sahabat, orang tua, dan motivator penulis selama ini. Atas cinta, pengorbanan, dan kesabaran hati serta dengan ikhlas mengiringi setiap langkah penulis dengan doa yang selalu terucap dan restu yang selalu tercurah. Untuk Saudara tercinta, **Fadly, Fadlan, Tita** dan **Tuti** serta keluarga besar penulis atas semangat yang selalu diberikan kepada penulis selama ini.

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 beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurtiti Sunusi S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf yang telah membekali ilmu dan kemudahan

kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.

4. **Bapak Siswanto, S.Si., M.Si.**, selaku Pembimbing Utama dan **Bapak Dr. Nirwan M.Si.**, selaku Pembimbing Pendamping dan Penasehat Akademik yang dengan penuh kesabaran telah meluangkan waktu, tenaga dan pemikirannya di tengah berbagai kesibukan dan prioritasnya untuk senantiasa memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal hingga selesainya penulisan tugas akhir ini.
5. **Bapak Andi Kresna Jaya S.Si., M.Si.** dan **Ibu Dra. Nasrah Sirajang, M.Si.**, selaku Tim Penguji yang telah memberikan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini serta waktu yang telah diberikan kepada penulis.
6. **Nurul Qarimah, Evaletrina Gracelita Marisda, Rappe, Yuniarti, Eka Safitri, Ismawati S,** dan **Fitriyani**, sahabat sekaligus keluarga yang sangat berharga bagi penulis, yang memberikan dukungan yang begitu besar, setia mendampingi, menyemangati, dan selalu ada berbagi suka dan duka bersama penulis selama bertahun-tahun, mulai dari bangku SMA hingga saat ini.
7. Sahabat tercinta, **Nurhidaya L., Andi Sri Yulianti,** dan **Sri Indriani Amil** atas kebersamaannya mengiringi perjalanan penulis dalam menjalani pendidikan, serta atas dukungan, bantuan, semangat dan kenangan berharga dari awal perkuliahan hingga saat ini.
8. **Yustika, Nurul Rezki, Nurul Ikhsani, Isra Rizka Utami, Hajratul Ashwad, Nur Anugrah Yusuf, Ainun, Siti Aisyah, Riskayani, Andi Umami Melin Aicha** dan **Teman-teman Statistika 2018** yang tidak bisa penulis sebutkan satu-persatu namanya. Terimakasih atas bantuan dan dukungan serta masa-masa perkuliahan yang telah dibagi bersama.
9. **Teman-teman KKN Tematik Universitas Hasanuddin Gelombang 106 Wilayah Maros 6 Kecamatan Simbang, Tanralili dan Moncongloe**, yang telah berbagi kenangan berharga selama masa-masa KKN dan memberikan pengalaman yang begitu luar biasa selama menjalani masa pengabdian.

10. Teman-teman tim PKM-RSH, **Nurul Rezki**, **Evaletrina Gracelita Marisda** dan **Ismail**, serta dosen pendamping **Bapak Siswanto, S.Si., M.Si.** atas pengalaman berharga yang tidak akan penulis dapatkan ditempat lain.
11. Kepada semua pihak yang tidak dapat penulis sebutkan satu-persatu, semoga segala dukungan dan partisipasi yang diberikan kepada penulis bernilai ibadah disisi Allah *Subhanallahu Wa Ta'ala*.

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

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Makassar, 4 Oktober 2022



Musdalifah

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIK**

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

Nama : Musdalifah
NIM : H051181024
Program Studi : Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Jenis Karya : Skripsi

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

“Pemodelan Robust Geographically and Temporally Weighted Regression dengan M-Estimator pada Penderita Pneumonia Usia Balita di Indonesia”

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

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal, 4 Oktober 2022

Yang menyatakan



(Musdalifah)

ABSTRAK

Pneumonia merupakan penyakit yang menyebabkan peradangan pada paru-paru dan menjadi salah satu penyakit yang paling banyak menginfeksi balita. Sebagai penyakit menular langsung, ada kemungkinan pengaruh keragaman lokasi terhadap jumlah penderita pneumonia. *Robust Geographically and Temporally Weighted Regression* (RGTWR) merupakan metode yang digunakan untuk memodelkan suatu data dengan mempertimbangkan heterogenitas lokasi dan waktu, serta mengatasi pencilan pada data. Data yang digunakan adalah jumlah penderita pneumonia usia balita beserta faktor-faktor yang diduga memengaruhinya yaitu jumlah puskesmas, kepadatan penduduk, persentase balita dengan imunisasi dasar lengkap, persentase balita diberi ASI eksklusif 0-6 bulan, persentase penduduk miskin, persentase balita 6-59 bulan mendapat vitamin A dan persentase rumah tangga dengan akses sanitasi layak. Penelitian ini dilakukan untuk memodelkan penderita pneumonia usia balita serta mendapatkan faktor-faktor yang signifikan memengaruhi jumlah penderita di masing-masing amatan. RGTWR menghasilkan model yang optimal dengan nilai R^2 sebesar 99.7037%, *Mean Absolute Deviation* sebesar 906.2049 dan *Median Absolute Deviation* sebesar 497.0159 dibanding model *Geographically and Temporally Weighted Regression*. Kepadatan Penduduk, Persentasi bayi diberi imunisasi dasar lengkap, persentase penduduk miskin dan persentase balita 6-59 bulan mendapat vitamin A menjadi faktor yang memengaruhi jumlah penderita pneumonia usia balita disebagian besar lokasi di 34 provinsi dan 5 tahun amatan.

Kata Kunci: *Robust Geographically and Temporally Weighted Regression, M-Estimator, Pneumonia, Pencilan*

ABSTRACT

Pneumonia is a disease that causes inflammation of the lungs and is one of the most common diseases infecting toddlers. As a directly infectious disease, there is a possibility of the influence of location diversity on the number of pneumonia sufferers. Robust Geographically and Temporally Weighted Regression (RGTWR) is a method used to model data by considering the heterogeneity of location and time and to overcome outliers in the data. The data used is the number of pneumonia sufferers aged under five and the factors that are thought to influence it, namely the number of health centers, population density, percentage of children under five with complete basic immunizations, percentage of children under five who are exclusively breastfed 0-6 months, percentage of poor people, percentage of toddlers 6-59 months. received vitamin A and the percentage of households with access to proper sanitation. This study was conducted to model pneumonia sufferers under five and to find out the factors that significantly affect the number of sufferers in each observation. RGTWR produces an optimal model with an R^2 value of 99.7037%, a Mean Absolute Deviation of 906.2049, and a Median Absolute Deviation of 497.0159 compared to the Geographically and Temporally Weighted Regression model. Population Density, the percentage of infants given complete basic immunization, the percentage of poor people and the percentage of children aged 6-59 months receiving vitamin A are factors that influence the number of pneumonia sufferers under five in most locations in 34 provinces and 5 years of observation.

Keywords: Robust Geographically and Temporally Weighted Regression, M-Estimator, Pneumonia, Outliers

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
LEMBAR PERNYATAAN KEONTETIKAN	iii
HALAMAN PERSETUJUAN PEMBIMBING	iv
HALAMAN PENGESAHAN	v
KATA PENGANTAR	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	ix
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Regresi Linier Berganda.....	5
2.2 Pengujian Asumsi Residual.....	6
2.3 Uji Heterogenitas Spasial	8
2.4 <i>Geographically Weighted Regression</i>	8
2.5 <i>Geographically and Temporally Weighted Regression</i>	10
2.6 <i>Spatial Outlier</i>	12
2.7 <i>Robust Regression</i> dengan <i>M-Estimator</i>	14
2.8 <i>Robust Geographically and Temporally Weighted Regression</i>	15
2.9 Pneumonia	16
BAB III METODE PENELITIAN	18
3.1 Sumber Data	18
3.2 Metode Penelitian.....	18

BAB IV HASIL DAN PEMBAHASAN.....	22
4.1 Eksplorasi Data Penderita Pneumonia Usia Balita di Indonesia.....	22
4.2 Pemodelan Regresi Linier Berganda	24
4.3 Uji Heterogenitas Spasial dan Temporal.....	28
4.4 Analisis model <i>Geographically and Temporally Weighted Regression</i>	29
4.5 Deteksi Pencilan Model GTWR.....	34
4.6 Analisis Model RGTWR dengan <i>M-Estimator</i>	36
BAB V KESIMPULAN DAN SARAN	50
5.1 Kesimpulan.....	50
5.2 Saran.....	50
DAFTAR PUSTAKA	51

DAFTAR TABEL

Tabel 3.1 Variabel Penelitian.....	18
Tabel 4.1 Deskripsi Data Jumlah Penderita Pneumonia Usia Balita di Indonesia.....	22
Tabel 4.2 ANOVA Model Regresi Linier Berganda	24
Tabel 4.3 Uji Parsial Model Regresi Linear Berganda	25
Tabel 4.4 Nilai VIF Variabel Bebas.....	27
Tabel 4.5 <i>Bandwidth</i> Model GWR dan GTWR.....	29
Tabel 4.6 Nilai Parameter Model GTWR	30
Tabel 4.7 Parameter Model GTWR	31
Tabel 4.8 Model GTWR	32
Tabel 4.9 Uji Parsial Parameter Model GTWR	33
Tabel 4.10 Peluang Pencilan Residual Model GTWR.....	35
Tabel 4.11 Model RGTWR Hasil Iterasi	37
Tabel 4.12 Uji Parsial Parameter Model RGTWR.....	45
Tabel 4.13 Jumlah Amatan Masuk Model	46
Tabel 4.14 Peluang Pencilan Residual Model RGTWR	47
Tabel 4.15 Pemilihan Model Terbaik.....	49

DAFTAR GAMBAR

Gambar 4.1 Plot Hubungan Linier antar Variabel Bebas dan Terikat	23
Gambar 4.2 Q-Q Plot Residual Model Regresi Linier Berganda.....	26
Gambar 4.3 <i>Boxplot</i> Sebaran Jumlah Penderita Pneumonia Usia Balita	28
Gambar 4.4 <i>Boxplot</i> Residual Model GTWR	34
Gambar 4.5 Peta Sebaran Parameter β_1	38
Gambar 4.6 Peta Sebaran Parameter β_2	39
Gambar 4.7 Peta Sebaran Parameter β_3	40
Gambar 4.8 Peta Sebaran Parameter β_4	41
Gambar 4.9 Peta Sebaran Parameter β_5	42
Gambar 4.10 Peta Sebaran Parameter β_6	43
Gambar 4.11 Peta Sebaran Parameter β_7	44
Gambar 4.12 <i>Boxplot</i> Residual Model RGTWR.....	47

DAFTAR LAMPIRAN

Lampiran 1 Deteksi Peluang Pencilan Dengan Metode RWBP Tahun 2016	56
Lampiran 2 Deteksi Peluang Pencilan Dengan Metode RWBP Tahun 2017	58
Lampiran 3 Deteksi Peluang Pencilan Dengan Metode RWBP Tahun 2018	60
Lampiran 4 Deteksi Peluang Pencilan Dengan Metode RWBP Tahun 2019	62
Lampiran 5 Deteksi Peluang Pencilan Dengan Metode RWBP Tahun 2020	64
Lampiran 6 Parameter Beta Model RGTWR Tahun 2016.....	66
Lampiran 7 Parameter Beta Model RGTWR Tahun 2017.....	69
Lampiran 8 Parameter Beta Model RGTWR Tahun 2018.....	72
Lampiran 9 Parameter Beta Model RGTWR Tahun 2019.....	75
Lampiran 10 Parameter Beta Model RGTWR Tahun 2020.....	78
Lampiran 11 Uji Parsial Parameter Beta Model RGTWR Tahun 2016.....	81
Lampiran 12 Uji Parsial Parameter Beta Model RGTWR Tahun 2017.....	84
Lampiran 13 Uji Parsial Parameter Beta Model RGTWR Tahun 2018.....	87
Lampiran 14 Uji Parsial Parameter Beta Model RGTWR Tahun 2019.....	90
Lampiran 15 Uji Parsial Parameter Beta Model RGTWR Tahun 2020.....	93

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pneumonia merupakan salah satu jenis penyakit menular langsung, sama halnya dengan diare, *typhus*, HIV/AIDS, tuberkulosis paru dan kusta. Kementerian Kesehatan Republik Indonesia (Kemenkes RI) melalui publikasi Profil Kesehatan Indonesia menyebutkan bahwa penanganan pneumonia sampai saat ini lebih diprioritaskan pada pengendalian pneumonia balita (Kemenkes RI, 2021). Usia Balita (Bawah Lima Tahun) menurut Departemen Kesehatan RI (2009) adalah anak dengan rentang usia 0-5 tahun (Amin, 2017). Pneumonia pada balita ditandai dengan batuk dan atau tanda kesulitan bernapas yaitu adanya nafas cepat, kadang disertai tarikan dinding dada bagian bawah kedalam.

Dilansir dari laman *website* resmi *World Health Organization* (WHO), pada tahun 2018 pneumonia menjadi penyebab kematian infeksi tunggal terbesar pada anak-anak di seluruh dunia. Lebih dari 800.000 kematian balita diseluruh dunia, atau sekitar satu balita per 39 detik disebabkan oleh pneumonia. Pneumonia membunuh 740.180 jiwa anak di bawah usia 5 tahun pada tahun 2019, menyumbang 14% dari semua kematian pada anak berusia dibawah 5 tahun (WHO, 2021).

Jumlah anak usia balita 0-4 tahun di Indonesia pada tahun 2018 diperkirakan sekitar 23.729.583 jiwa (Kemenkes, 2019). Estimasi global menunjukkan bahwa pada tahun tersebut, setiap satu jam ada 71 anak di Indonesia tertular pneumonia dan lebih dari 19.000 balita meninggal akibat pneumonia, atau lebih dari 2 balita setiap jamnya. *United Nations International Children's Emergency Fund* (UNICEF) mencatat 25.481 kematian balita Indonesia karena infeksi pernapasan akut, menempatkan Indonesia di peringkat ketujuh dunia dengan beban pneumonia tertinggi pada tahun 2017. Angka kematian tersebut menunjukkan bahwa 17% kematian balita di dunia akibat pneumonia berasal dari Indonesia (KOMPAS, 2021). Besarnya jumlah penderita pneumonia tersebut menyebabkan pada tahun 2019, pneumonia menjadi penyebab kematian balita (usia 12-59 bulan) kedua terbanyak setelah diare (WHO, 2021).

Penelitian yang dilakukan oleh Ristiani (2021) menyebutkan faktor yang memengaruhi balita terkena pneumonia di Jawa Barat diantaranya adalah jumlah penduduk miskin, kepadatan penduduk, gizi kurang dan air layak konsumsi. Renika dan Amin (2021) menyebutkan faktor yang memengaruhi pneumonia pada balita di Pulau Jawa tahun 2018 adalah kepadatan penduduk, jumlah puskesmas, dan persentase penduduk dengan akses sanitasi layak. Faktor yang memengaruhi kasus pneumonia balita di Jawa Barat menurut penelitian Nadya dkk. (2017) adalah jumlah balita dengan gizi buruk, persentase bayi diberi imunisasi dasar lengkap dan jumlah puskesmas. Selanjutnya, penelitian yang dilakukan oleh Islamiyah (2020) menyimpulkan bahwa kasus pneumonia balita di Sulawesi Selatan dipengaruhi oleh persentase bayi mendapat vitamin A, bayi diberi ASI eksklusif dan persentase bayi diberi imunisasi dasar lengkap.

Faktor-faktor diatas merupakan beberapa faktor yang memengaruhi jumlah kasus pneumonia pada balita di beberapa wilayah. Pneumonia sebagai penyakit menular langsung menyebabkan adanya kemungkinan pengaruh lokasi atau spasial yang mempengaruhi jumlah kasus pneumonia pada balita, seperti yang dibuktikan oleh penelitian di atas. Menurut Fotheringham dkk (2002), salah satu metode yang dapat digunakan untuk menganalisis faktor yang berpengaruh terhadap suatu variabel dengan mempertimbangkan lokasi yang berbeda atau adanya keragaman spasial adalah dengan model *Geographically Weighted Regression* (GWR). Analisis GWR mengakomodasi unsur lokasi tanpa mengikut sertakan unsur waktu pengamatan dalam pemodelan (Huang dkk., 2010). Menurut Fotheringham dkk (2015), penambahan unsur waktu perlu dilakukan agar didapat pendugaan parameter yang lebih akurat. Penambahan unsur waktu pada GWR dapat dilakukan dengan metode *Geographically and Temporally Weighted Regression* (GTWR) yang mengakomodasi adanya heterogenitas secara spasial (lokasi) dan secara temporal (Huang dkk., 2010).

Pendugaan koefisien regresi pada analisis GTWR dilakukan dengan menggunakan metode kuadrat terkecil terboboti. Menurut Zhang dan Mei (2011), metode tersebut tidak kekar (*robust*) terhadap pencilan (*outlier*), sehingga apabila terdapat pencilan pada data, maka akan mengakibatkan terciptanya koefisien penduga yang bias dan mengakibatkan kekeliruan dalam menyimpulkan

hubungan regresi. Salah satu metode yang mengakomodir adanya pencilan adalah *Robust Geographically and Temporally Weighted Regression* (RGTWR) dengan *M-Estimator*. *M-Estimator* merupakan pengembangan dari *Maximum Likelihood Estimation* (MLE) yang dilakukan dengan meminimumkan fungsi objektif dari galat agar model yang dihasilkan kekar terhadap pencilan. Menurut Fox (2002), fungsi objektif merupakan fungsi yang digunakan untuk mencari fungsi pembobot pada regresi kekar.

Penelitian tentang RGTWR dengan *M-Estimator* telah diterapkan sebelumnya oleh Erda (2018), dengan melakukan pendugaan model RGTWR dengan *M-Estimator* pada kasus jumlah peserta KB Aktif di Jawa Timur Tahun 2009-2016. Penelitian tersebut menghasilkan model RGTWR dengan *M-Estimator* yang mampu menurunkan nilai *Mean Absolute Deviation* dan *Median Absolute Deviation* hasil pemodelan GTWR. Putra dan Oktora (2021) melakukan pemodelan RGTWR dengan *MM-Estimator* pada deforestasi di Pulau Sumatera tahun 2011-2019. Putra (2019) melakukan pemodelan RGTWR dengan *GM-Estimator* dan *S-Estimator* untuk tingkat kriminalitas di Provinsi Jawa Tengah dan Jawa Timur tahun 2011-2015.

Berdasarkan uraian tersebut diatas, peneliti tertarik untuk melakukan pemodelan RGTWR dengan *M-Estimator* pada penderita pneumonia usia balita di Indonesia. Metode tersebut dianggap mampu mengatasi keragaman spasial dan temporal serta pencilan pada data, sehingga hasil yang didapatkan akan mampu menjelaskan kondisi sebenarnya di lapangan secara lebih akurat. Data yang digunakan pada penelitian ini adalah jumlah penderita pneumonia pada balita di Indonesia tahun 2016-2020 beserta faktor-faktor yang mungkin memengaruhinya.

1.2 Rumusan Masalah

1. Bagaimana pemodelan RGTWR dengan *M-Estimator* pada penderita pneumonia usia balita di Indonesia tahun 2016-2020?
2. Faktor-faktor apa sajakah yang secara signifikan memengaruhi jumlah penderita pneumonia usia balita di Indonesia pada tahun 2016-2020?

1.3 Batasan Masalah

Penelitian ini dibatasi pada 7 variabel yang diduga mempengaruhi jumlah penderita pneumonia usia balita di Indonesia tahun 2016-2020. Metode yang digunakan adalah RGTWR dengan menggunakan *M-Estimator*, dan fungsi kernel yang digunakan adalah fungsi Kernel *Exponential*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Memperoleh model RGTWR dengan *M-Estimator* pada kasus jumlah penderita pneumonia usia balita di Indonesia tahun 2016-2020.
2. Mendapatkan faktor-faktor yang secara signifikan memengaruhi jumlah penderita pneumonia usia balita di Indonesia tahun 2016-2020.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Memberikan informasi kepada pihak berkepentingan mengenai faktor yang mempengaruhi jumlah penderita pneumonia pada balita di Indonesia.
2. Menambah dan mengembangkan wawasan terkait metode RGTWR dengan *M-Estimator*. Penelitian ini dapat menjadi rujukan untuk penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Regresi Linier Berganda

Analisis regresi linier merupakan suatu metode untuk menjelaskan hubungan antara variabel respon dan variabel prediktor secara linier (Draper & Smith, 1988). Model regresi linier hubungan antara variabel respon (y) dan variabel prediktor ($x_1, x_2, x_3, \dots, x_p$) untuk p variabel prediktor secara umum ditulis sebagai berikut.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i$$

Dalam bentuk matriks, persamaan umum untuk model regresi linier adalah sebagai berikut.

$$Y = X\beta + \varepsilon$$

dengan;

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}; X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1p} \\ 1 & X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix}; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix};$$

$$i = 1, 2, \dots, n; k = 1, 2, 3, \dots, p.$$

Keterangan:

n : Jumlah pengamatan

p : Jumlah prediktor

Y_i : Nilai observasi variabel respon pada pengamatan ke- i

X_{ik} : Nilai observasi variabel prediktor ke- k pada pengamatan ke- i

β_0 : Nilai intersep model regresi

β_p : Koefisien regresi variabel prediktor ke- p

ε_i : Residual pada pengamatan ke- i dengan asumsi independen, identik, dan berdistribusi normal, dengan mean nol dan varians konstan σ^2 (Fotheringham dkk., 2002).

2.2 Pengujian Asumsi Residual

1. Asumsi Independen

Asumsi independen merupakan asumsi dari model regresi yang mengharuskan tidak terdapat korelasi antar residual. Uji yang digunakan untuk mendeteksi kasus autokorelasi adalah uji *Durbin Watson* (Drapper & Smith, 1991). Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \rho = 0 \text{ (tidak terjadi autokorelasi)}$$

$$H_1 : \rho \neq 0 \text{ (terjadi autokorelasi)}$$

Statistik uji yang digunakan:

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

Keterangan:

d : Nilai Durbin Watson

ε_i : Residual pada pengamatan ke- i

ε_{i-1} : Residual pada pengamatan ke- $i - 1$

Setelah nilai statistik uji Durbin Watson didapatkan, selanjutnya adalah membandingkan hasil tersebut dengan tabel *Durbin Watson* (DW). Tabel DW terdiri atas batas bawah (d_L) dan batas atas (d_U). Berikut beberapa keputusan setelah membandingkan dengan tabel DW.

$$d < d_L : \text{Tolak } H_0$$

$$d > 4 - d_L : \text{Tolak } H_0$$

$$d_U < d < 4 - d_L : \text{Terima } H_0$$

$$d_L \leq d \leq d_U \text{ dan } 4 - d_U \leq d \leq 4 - d_L : \text{Tidak ada keputusan}$$

2. Uji Normalitas *Kolmogorov-Smirnov*

Uji *Kolmogorov-Smirnov* bertujuan untuk mengetahui distribusi suatu data telah berdistribusi normal atau tidak (Drapper & Smith, 1991). Hipotesis yang digunakan yaitu:

$$H_0 : F_n(\varepsilon) = F_0(\varepsilon) \text{ (Residual berdistribusi normal)}$$

$$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon) \text{ (Residual tidak berdistribusi normal)}$$

Statistik uji:

$$D = \text{Max } |F_n(\varepsilon) - F_0(\varepsilon)|$$

Tolak H_0 apabila $D > D_\alpha$.

Keterangan:

D_α : Nilai kritis untuk uji Kolmogorov Smirnov satu sampel, diperoleh dari tabel Kolmogorov Smirnov satu sampel

$F_n(\varepsilon)$: Nilai peluang kumulatif (fungsi distribusi kumulatif) sampel

$F_0(\varepsilon)$: Nilai peluang kumulatif (fungsi distribusi kumulatif) teoritis

3. Uji Multikolinearitas

Multikolinearitas adalah kondisi ketika terdapat korelasi antar variabel prediktor yang dapat menyebabkan kesalahan estimasi parameter pada model regresi linier. Adanya multikolinearitas dapat diketahui melalui nilai *Variance Inflated Factor* (VIF). Nilai VIF didapatkan melalui perhitungan berikut ini.

$$VIF = \frac{1}{1 - R_j^2}$$

R_j^2 adalah koefisien determinasi antara variabel prediktor x_j dengan variabel prediktor lainnya. Jika nilai VIF > 10, maka dapat dikatakan terdapat multikolinearitas (Laksana, 2018).

4. Asumsi Identik (Homoskedastisitas)

Asumsi identik (homoskedastisitas) berarti bahwa varians pada residual sama atau identik. Kebalikanannya adalah kasus heteroskedastisitas, yaitu jika kondisi varians residual tidak identik (Gujarati, 2003).

$$var(y_i) = var(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n$$

Uji identik dapat dilakukan dengan menggunakan uji *Glejser*. Hipotesis yang digunakan adalah sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (residual identik)

$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$ (residual tidak identik)

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|^2)}{p}}{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)}{n - p - 1}}$$

Jika $F_{hitung} > F_{\alpha, p, (n-p-1)}$ atau $p - value < \alpha$, maka tolak H_0 yang berarti asumsi identik tidak terpenuhi (Laksana, 2018).

2.3 Uji Heterogenitas Spasial

Pengujian heterogenitas spasial dilakukan untuk mengetahui adanya perbedaan karakteristik satu wilayah dengan wilayah lainnya. Adanya heterogenitas spasial dapat menghasilkan parameter regresi berbeda di tiap lokasi pengamatan, dan dapat diuji dengan menggunakan statistik uji *Breuch-Pagan* dengan hipotesis sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (tidak ada heterogenitas spasial)

$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, n$ (ada heterogenitas spasial)

Statistik uji yang digunakan adalah seperti berikut.

$$BP = \left(\frac{1}{2}\right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (2.1)$$

dengan elemen vector \mathbf{f} adalah

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

Keterangan:

e_i : residual pada pengamatan ke- i

\mathbf{Z} : matriks berukuran $n \times (p + 1)$ yang berisi vektor yang sudah dinormalstandarkan untuk tiap pengamatan. H_0 akan ditolak jika nilai $BP > \chi_{\alpha; p}^2$ (Laksana, 2018).

2.4 Geographically Weighted Regression

Model *Geographically Weighted Regression* (GWR) merupakan metode statistika yang digunakan untuk menganalisis pengaruh antara peubah prediktor terhadap respon berdasarkan adanya keragaman spasial (Fotheringham, 2002). Pada metode GWR, setiap parameter regresi dihitung pada setiap lokasi sehingga setiap titik lokasi mempunyai nilai parameter regresi yang berbeda-beda. Hubungan antara peubah respon Y dan peubah penjelas x_1, x_2, \dots, x_k pada model GWR untuk lokasi ke- i adalah (Huang dkk., 2010):

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

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

Keterangan :

y_i : Nilai pengamatan peubah respon untuk lokasi pengamatan ke- i

- $\beta_0(u_i, v_i)$: Intersep pada lokasi pengamatan ke- i
- $\beta_k(u_i, v_i)$: Koefisien peubah penjelas ke- k pada lokasi pengamatan ke- i
- x_{ik} : Pengamatan peubah penjelas ke- k pada lokasi pengamatan ke- i
- ε_i : Residual pengamatan pada lokasi ke- i

Pemodelan pada GWR sangat bergantung pada pembobot spasial yang digunakan. Pembobot \mathbf{W} merupakan diagonal matriks dari w_j (bobot tiap titik data dengan lokasi ke- j). Menurut Fotheringham dkk (2002), pemilihan pembobot spasial yang digunakan sangat penting karena pembobot dapat menjelaskan besarnya pengaruh ketetanggaan suatu daerah. Pembobot yang terbentuk diperoleh dari fungsi Kernel. Terdapat beberapa fungsi Kernel yang dapat digunakan untuk pembobotan, antara lain:

1. Fungsi Kernel *Gaussian*

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

2. Fungsi Kernel *Exponential*

$$w_j(u_i, v_j) = \exp \left[-\left(\frac{d_{ij}}{h} \right) \right] \tag{2.2}$$

3. Fungsi Kernel *Bisquare*

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

4. Fungsi Kernel *Tricube*

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

Keterangan:

- $w_j(u_i, v_j)$: Nilai fungsi Kernel pada lokasi ke- i dan lokasi ke- j
- h : Lebar jendela (*bandwidth*)
- d_{ij} : Jarak antara lokasi ke- i dan lokasi ke- j berdasarkan titik koordinat $(u_i - v_i)$ dan $(u_j - v_j)$

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

Menurut Fotheringham dkk (2002), lebar jendela (*bandwidth*) merupakan ukuran jarak fungsi pembobot yang dapat mengukur sejauh mana pengaruh suatu lokasi dengan lokasi lainnya. Pemilihan lebar jendela yang optimal dapat dilakukan melalui metode *Cross Validation* (CV). Lebar jendela optimal dipilih berdasarkan CV minimum. Perhitungan CV menggunakan rumus sebagai berikut:

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(h)]^2 \quad (2.3)$$

Keterangan :

CV : Nilai *Cross Validation*

y_i : Nilai observasi variabel respon pada lokasi ke- i

$\hat{y}_{\neq i}$: Nilai dugaan dari model GWR tanpa melibatkan lokasi ke- i dalam proses estimasi

h : *Bandwidth*

2.5 Geographically and Temporally Weighted Regression

Geographically-Temporally Weighted Regression (GTWR) merupakan metode pengembangan dari GWR yang mengakomodasi adanya heterogenitas secara spasial dan secara temporal (Huang dkk., 2010). Persamaan umum untuk model GTWR untuk p peubah respon dengan peubah penjelas y_i pada lokasi (u_i, v_i, t_i) adalah sebagai berikut:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (2.4)$$

Dengan:

y_i : Nilai pengamatan peubah respon untuk lokasi pengamatan ke- i

$\beta_0(u_i, v_i, t_i)$: Intersep pada lokasi pengamatan ke- (u_i, v_i) dan waktu ke- t_i

$\beta_k(u_i, v_i, t_i)$: Koefisien peubah penjelas ke- k pada lokasi pengamatan ke- (u_i, v_i) dan waktu ke- t_i

x_{ik} : Nilai pengamatan peubah penjelas ke- k di lokasi pengamatan ke- i dengan $k = 1, 2, \dots, p$

ε_i : Residual pengamatan pada lokasi ke- i

Parameter model GTWR $\hat{\beta}_k(u_i, v_i, t_i)$ untuk setiap k peubah pada lokasi ke- i diperoleh dengan menggunakan metode kuadrat terkecil terboboti (*Weighted*

Least Square) dengan memberi pembobot berbeda untuk setiap lokasi dan waktu. Pendugaan parameter dituliskan sebagai berikut.

$$\hat{\beta}(u_i, v_i, t_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{Y} \quad (2.5)$$

dengan

$$\mathbf{W}(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in}) \quad (2.6)$$

Keterangan :

n : Jumlah data observasi

$\mathbf{W}(u_i, v_i, t_i)$: Matriks pembobot pada pengamatan (u_i, v_i) dan waktu ke t_i .

Penyusunan model GTWR mengasumsikan pengamatan yang berdekatan dengan lokasi ke- i memiliki pengaruh yang lebih besar terhadap $\hat{\beta}(u_i, v_i, t_i)$ dibandingkan dengan lokasi yang terletak lebih jauh dari lokasi ke- i . Kedekatan lokasi observasi ini memiliki dua unsur, yaitu kedekatan spasial dan kedekatan temporal (Winarso dan Yasin, 2016).

Fungsi jarak spasial-temporal (d_{ij}^{ST}) terdiri dari gabungan fungsi jarak spasial dan fungsi jarak temporal, dituliskan sebagai berikut:

$$\begin{aligned} (d_{ij}^S)^2 &= (u_i - u_j)^2 + (v_i - v_j)^2 \\ (d_{ij}^T)^2 &= (t_i - t_j)^2 \\ (d_{ij}^{ST})^2 &= \lambda [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu [(t_i - t_j)^2] \end{aligned} \quad (2.7)$$

dengan λ dan μ merupakan parameter penyeimbang terhadap pengaruh perbedaan satuan antara lokasi dan waktu pada pengukuran jarak spasial temporal (Huang dkk., 2010).

Berdasarkan persamaan tersebut kemudian diperoleh:

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

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

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

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

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

Keterangan:

h_S : *Bandwidth* spasial

h_T : *Bandwidth* temporal

h_{ST} : *Bandwidth* spasial-temporal

Menurut Liu dkk. (2017), misalkan τ merupakan parameter rasio dari $\tau = \frac{\mu}{\lambda}$ dengan $\lambda \neq 0$ maka diperoleh persamaan:

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

Parameter τ berfungsi untuk memperbesar atau memperkecil rasio dari jarak temporal terhadap jarak spasial. Parameter ini didapatkan dari kriteria CV minimum melalui inisialisasi nilai τ awal sebagai berikut:

$$CV(\tau) = \sum_{l=1}^n [y_{=1} - \hat{y}_{\neq i}(\tau)]^2 \quad (2.8)$$

Selanjutnya penduga parameter λ dan μ bisa diperoleh dengan metode iteratif berdasarkan hasil penduga τ yang menghasilkan CV minimum.

2.6 Spatial Outlier

Outlier (pencilan) merupakan pengamatan yang tidak mengikuti sebagian besar pola data dan nilai pengamatannya terletak jauh dari pusat data. Pendeteksian pencilan pada data spasial dapat menggunakan pendekatan *Random Walk* (RW). Pendekatan ini disusun berdasarkan atribut spasial dan nonspasial pada data dan dideteksi dengan menghitung jarak antara pengamatan spasial dengan tetangga terdekat melalui dua grafik (*Random Walk Bipartite*, RWBP) yang terboboti. RWBP dilakukan dengan langkah-langkah seperti berikut.

1. Pembentukan Dua Grafik Terboboti

Pembentukan grafik terdiri dari tiga tahapan dasar. Pertama, mengelompokkan peubah nonspasial berdasarkan metode penggerombolan K -

Means. Kemudian, membentuk dua grafik, bagian kiri berisi data spasial dan bagian kanan berisi data nonspasial yang telah digerombolkan sebelumnya. Selanjutnya, menghitung nilai *edge* (nilai kemiripan pengamatan spasial dan kelompok) berdasarkan peubah nonspasial dari pengamatan spasial dan nilai pusat dari kelompok yang terbentuk yang dihitung dengan cara sebagai berikut:

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

Keterangan :

$Atr(P_i)$: Peubah nonspasial ke- i

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

α : Nilai konstanta yang diperoleh berdasarkan jangkauan dari nilai peubah nonspasial.

2. Menghitung Kemiripan antara Pengamatan Spasial

Perhitungan kemiripan antara pengamatan spasial dilakukan dengan menerapkan metode jalan acak pada dua grafik terboboti. Jalan acak berarti langkah 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 tiap 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 lokasi ke- i

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

- Mengidentifikasi pencilan spasial berdasarkan nilai kemiripan dimana nilai kemiripan terkecil akan dideteksi sebagai pencilan (Liu dkk., 2010).

2.7 Robust Regression dengan M -Estimator

Robust regression (regresi kekar) merupakan metode regresi yang digunakan ketika sebaran dari residual tidak normal dan atau adanya berapa pencilan yang berpengaruh terhadap model (Olive, 2005). Regresi kekar digunakan untuk memberikan hasil pendugaan yang *robust* terhadap adanya pencilan (Chen 2002). Menurut Fox (2002), regresi kekar dengan M -Estimator dilakukan dengan meminimumkan fungsi objektif dari galat dengan persamaan sebagai berikut:

$$M = \sum_{i=1}^n \rho(\varepsilon_i)$$

Keterangan:

M : Nilai Penaksiran M

ρ : Fungsi objektif

ε_i : Residual pada pengamatan ke- i

Menurut Fox (2002), fungsi objektif merupakan fungsi yang digunakan untuk mencari fungsi pembobot pada regresi kekar. Terdapat beberapa fungsi objektif yang dapat digunakan pada regresi kekar, yaitu:

- Fungsi Objektif *Huber*

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

dengan

$$\rho'(\varepsilon_i) = \begin{cases} \varepsilon_i^*, & |\varepsilon_i^*| \leq c \\ c, & \varepsilon_i^* > c \\ -c, & \varepsilon_i^* < -c \end{cases}$$

dan fungsi pembobot :

$$w_i = \begin{cases} 1, & |\varepsilon_i^*| \leq c \\ \frac{c}{|\varepsilon_i^*|}, & |\varepsilon_i^*| > c \end{cases}$$

2. Fungsi Objektif *Tukey*

$$\rho(\varepsilon_i) = \begin{cases} \frac{c^2}{6} \left\{ 1 - \left[1 - \left(\frac{\varepsilon_i^*}{c} \right)^2 \right]^2 \right\}, & |\varepsilon_i^*| \leq c \\ \frac{c^2}{6}, & |\varepsilon_i^*| > c \end{cases}$$

dengan

$$\rho'(\varepsilon_i) = \begin{cases} \varepsilon_i^* \left(1 - \left(\frac{\varepsilon_i^*}{c} \right)^2 \right)^2, & |\varepsilon_i^*| \leq c \\ 0, & |\varepsilon_i^*| > -c \end{cases}$$

dan fungsi pembobot :

$$w_i = \begin{cases} \left(1 - \left(\frac{\varepsilon_i^*}{c} \right)^2 \right)^2, & |\varepsilon_i^*| \leq c \\ 0, & |\varepsilon_i^*| > c \end{cases} \quad (2.9)$$

$\varepsilon_i^* = \frac{\varepsilon_i}{\hat{\sigma}}$, dan konstanta c untuk fungsi pembobot *Huber* bernilai 1.345 dan untuk pembobot *Tukey* bernilai 4.685. Nilai tersebut merupakan konstanta yang menghasilkan efisiensi tinggi dengan galat menyebar normal dan dapat kekar terhadap pencilan (Fox, 2002).

2.8 *Robust Geographically and Temporally Weighted Regression*

Regresi yang kekar pada GTWR dilakukan dengan menambahkan pembobot pada persamaan (2.4). Model *Robust Geographically and Temporally Weighted Regression* (RGTWR) untuk lokasi dan waktu ke- i yang mengandung pencilan adalah sebagai berikut:

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

dengan : $i = 1, 2, \dots, n$; $x_i = x_{i1}, x_{i2}, \dots, x_{ip}$

RGTWR dengan *M-Estimator* dilakukan dengan meminimumkan galat dengan persamaan sebagai berikut:

$$\hat{\beta}_m = \min_{\beta} \sum_{i=1}^n \rho(\varepsilon_i) = \min_{\beta} \sum_{i=1}^n \rho(y_i - X_i^T \beta) = \sum_{i=1}^n \left(\rho(y_i - X_i^T \beta(u_i, v_i, t_i)) \right)$$

M-Estimator dilakukan dengan menggunakan metode *Iteratively Reweighted Least Square* (IRLS) atau iterasi kuadrat terkecil terboboti. Pada iterasi ini, nilai w_i akan berubah nilainya disetiap iterasinya, sehingga:

$$\hat{\beta}(u_i, v_i, t_i)^m = (X_i' W^{m-1} X_i)^{-1} X_i' W^{m-1} y_i \quad (2.10)$$

dengan m merupakan banyaknya iterasi.

Pada W^m pembobot yang diberikan, diperoleh penduga:

$$\hat{\beta}(u_i, v_i, t_i)^{m+1} = (X_i' W^m X_i)^{-1} X_i' W^m y_i$$

Perhitungan persamaan diatas akan terus berulang hingga diperoleh penduga yang konvergen, yakni ketika selisih nilai $\hat{\beta}(u_i, v_i, t_i)^{m+1}$ dan $\hat{\beta}(u_i, v_i, t_i)^m$ mendekati nilai nol (Erda, 2018).

2.9 Pneumonia

Pneumonia merupakan peradangan yang mengenai parenkim paru, distal dari bronkiolus respiratorius, dan alveoli, serta menimbulkan konsolidasi jaringan paru dan gangguan pertukaran gas setempat (Dahlan, 2014). Sampai saat ini program dalam pengendalian pneumonia lebih diprioritaskan pada pengendalian pneumonia balita. Gejala pneumonia pada balita ditandai dengan batuk atau tanda kesulitan bernapas yaitu adanya napas cepat, kadang disertai tarikan dinding dada bagian bawah ke dalam (Kementian Kesehatan Republik Indonesia, 2021).

Pada umumnya, pneumonia dikategorikan dalam penyakit menular yang ditularkan melalui udara, dengan sumber penularan penyakit adalah penderita pneumonia yang menyebarkan kuman dalam bentuk droplet ke udara pada saat batuk atau bersin. Untuk selanjutnya, kuman penyebab pneumonia masuk ke saluran pernapasan melalui proses inhalasi (udara yang dihirup), atau dengan cara penularan langsung, yaitu percikan droplet yang dikeluarkan oleh penderita saat batuk, bersin, dan berbicara langsung terhirup oleh orang disekitar penderita, atau memegang dan menggunakan benda yang telah terkena sekresi saluran pernapasan penderita. Berdasarkan penelitian Mardjanis Said (Said, 2010), faktor yang menyebabkan penyebaran pneumonia balita antara lain:

1. Faktor internal meliputi status gizi, pemberian ASI eksklusif, imunisasi dasar lengkap, dan Bayi dengan Berat Badan Lahir Rendah (BBLR).
2. Faktor eksternal meliputi tersedianya puskesmas di setiap daerah, tersedianya air minum layak, kemiskinan dan kondisi rumah.