

**PEMODELAN *NONPARAMETRIC GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED REGRESSION* DENGAN
PENDEKATAN *SPLINE TRUNCATED* PADA TINGKAT
PENGANGGURAN TERBUKA DI PROVINSI
SULAWESI SELATAN**

SKRIPSI



MUH. FAHMI ALKAF

H051201021

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

UNIVERSITAS HASANUDDIN

MAKASSAR

2024



**PEMODELAN *NONPARAMETRIC GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED REGRESSION* DENGAN
PENDEKATAN *SPLINE TRUNCATED* PADA TINGKAT
PENGANGGURAN TERBUKA DI PROVINSI
SULAWESI SELATAN**

SKRIPSI

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

MUH. FAHMI ALKAF

H051201021

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

UNIVERSITAS HASANUDDIN

MAKASSAR

MEI 2024



LEMBAR PERNYATAAN KEOTENTIKAN

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

Pemodelan *Nonparametric Geographically and Temporally Weighted Regression* dengan Pendekatan *Spline Truncated* pada Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan

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

Makassar, 14 Mei 2024

A 10000 Indonesian postage stamp featuring a Garuda emblem and the text 'METERAL TEMPEL' and 'B9EALX194995852'. A handwritten signature is written over the stamp.

Muh. Fahmi Alkaf

NIM H051201021



**PEMODELAN *NONPARAMETRIC GEOGRAPHICALLY AND
TEMPORALLY WEIGHTED REGRESSION* DENGAN
PENDEKATAN *SPLINE TRUNCATED* PADA TINGKAT
PENGANGGURAN TERBUKA DI PROVINSI
SULAWESI SELATAN**

Disetujui Oleh:

Pembimbing Utama



Dr. Anna Islamiyati, S.Si., M.Si.

NIP. 19770808 200501 2 002

Ketua Program Studi



Dr. Anna Islamiyati, S.Si., M.Si.

NIP. 19770808 200501 2 002

Pada 14 Mei 2024



HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Muh. Fahmi Alkaf

NIM : H051201021

Program Studi : Statistika

Judul Skripsi : *Pemodelan Nonparametric Geographically and Temporally Weighted Regression* dengan Pendekatan *Spline Truncated* pada Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan

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

DEWAN PENGUJI

1. Ketua : Dr. Anna Islamiyati, S.Si., M.Si.
2. Anggota : Prof. Dr. Dr. Georgina Maria Tinungki, M.Si.
3. Anggota : Dr. Erna Tri Herdiani, S.Si., M.Si.



n di : Makassar

: 14 Mei 2024

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur kehadirat Allah *Subhanahu Wa Ta'ala* atas limpahan rahmat dan karunianya sehingga penulis dapat menyelesaikan penyusunan tugas akhir ini. Shalawat serta salam senantiasa tercurah kepada baginda *Rasulullah* Muhammad *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan para sahabatnya. *Alhamdulillah rabbil'alamin*, berkat nikmat kemudahan yang diberikan oleh Allah *Subhanahu Wa Ta'ala* sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “**Pemodelan *Nonparametric Geographically and Temporally Weighted Regression* dengan Pendekatan *Spline Truncated* pada Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan**” sebagai suatu syarat untuk memperoleh gelar sarjana pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis menyadari bahwa penyelesaian tugas akhir ini tidak terlepas dari bantuan dan dorongan dari berbagai pihak baik dalam bentuk materil maupun moril, sehingga dengan segala keterbatasan pengetahuan dan kemampuan, penulis dapat menyelesaikan tugas akhir ini. Oleh karena itu, ucapan terima kasih serta penghargaan yang setinggi-tingginya senantiasa terucap kepada Ayahanda **Muhammad Side** dan ibunda **Almh. Nurhaedah** yang telah membesarkan dan memberikan didikan yang sebaik-baiknya kepada penulis, memberikan dukungan penuh, pengorbanan yang luar biasa, kesabaran hati, serta keikhlasan dalam menemani langkah penulis dengan doa dan restu mulianya. Ucapan terima kasih juga penulis haturkan kepada saudara-saudara tersayang penulis **Muh. Irham Maulana**, **Muhammad Fikri Haikal**, dan **Nurfadillah Alhumaerah** yang selalu menghibur dan memberi semangat kepada penulis dalam menyelesaikan tugas akhir ini, serta keluarga besar penulis, terima kasih atas doa mulia dan dukungannya selama ini.

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

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. Anna Islamiyati S.Si., M.Si.**, selaku Ketua Departemen Statistika sekaligus Pembimbing Utama dan Penasehat akademik penulis yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya dalam memberikan arahan, dorongan semangat dan motivasi kepada penulis dari awal hingga selesainya penulisan tugas akhir ini serta selama menjadi mahasiswa di Departemen Statistika.
4. **Ibu Prof. Dr. Dr. Georgina Maria Tinungki, M.Si.** dan **Ibu Dr. Erna Tri Herdiani, S.Si., M. Si** selaku Tim Penguji yang telah meluangkan waktu dalam memberikan motivasi serta kritikan yang membangun kepada penulis dalam penyempurnaan tugas akhir ini.
5. Segenap Dosen Pengajar dan Staf yang telah memberikan ilmu dan kemudahan kepada penulis dalam berbagai hal selama menempuh pendidikan sarjana di Departemen Statistika.
6. Teman diskusi terbaik bagi penulis, **Ahmad Mukhlis Mursidin, Fadlan Amin, dan Ahmad Ryan Al-Aqsha** terima kasih telah menjadi sahabat dan teman diskusi dalam proses penyelesaian tugas akhir ini.
7. Sahabat-sahabat “**RANDOM**” penulis selama melewati masa perkuliahan, **Ahmad Mukhlis Mursidin, Fadlan Amin, Eunike Laurine Randa, Krisna Dwi Kayana, Siti Aisyah Surianda, Andi Muhammad Alif Reza Bahri, Baharuddin Aldi Jayanto, Ahmad Dzaki Fauzan, dan Muhammad Rais Rifat**. Terimakasih atas persaudaraannya dan bantuan yang diberikan kepada penulis sehingga bisa melalui proses perkuliahan dengan mudah dan sangat berkesan. Penulis merasa bersyukur dipertemukan dengan mereka dalam proses perkuliahan.
8. Teman seperjuangan di Statistika 2020, terkhusus **Yoel Sarapang, Nur Aliah Ramadani, Divia Ratri, Andi Muhammad Hakam**. Terima kasih atas kebersamaan, kebahagiaan, kesedihan, serta kebaikannya menjadi sosok guru penulis. Terima kasih telah mengukir kenangan indah bersama penulis na masa perkuliahan.



LEMBAR PERNYATAAN KEOTENTIKAN

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

Pemodelan *Nonparametric Geographically and Temporally Weighted Regression* dengan Pendekatan *Spline Truncated* pada Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan

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

Makassar, 14 Mei 2024

A 10000 Indonesian postage stamp (Meterai Tempel) with a signature over it. The stamp features the Garuda Pancasila emblem and the text 'REPUBLIK INDONESIA', '10000', 'METERAI TEMPEL', and the serial number 'C. B9EALX194995852'.

Muh. Fahmi Alkaf

NIM H051201021



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

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

Nama : Muh. Fahmi Alkaf
NIM : H051201021
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 *Nonparametric Geographically and Temporally Weighted Regression* dengan Pendekatan *Spline Truncated* pada Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan”

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

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar tanggal 14 Mei 2024.

nyatakan,

hmi Alkaf)



ABSTRAK

Tingkat Pengangguran Terbuka (TPT) merupakan suatu indikator yang mengukur persentase jumlah pengangguran terhadap jumlah angkatan kerja. TPT di Provinsi Sulawesi Selatan menunjukkan adanya heterogenitas spasial dan temporal di saat yang bersamaan. Selain itu, pola hubungan antarvariabel menyebar secara acak sehingga sulit untuk menaksir kurva regresinya. Untuk mengatasi masalah tersebut digunakan model pengembangan regresi spasial dan regresi nonparametrik yaitu metode *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated*. Penelitian ini bertujuan untuk memodelkan TPT di Provinsi Sulawesi Selatan menggunakan metode *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated*. Berdasarkan hasil analisis, model terbaik yang diperoleh menggunakan fungsi pembobot *fixed kernel exponential* dengan titik knot dengan nilai GCV paling minimum yaitu pada kombinasi titik knot. Model lokal yang diperoleh bervariasi pada masing-masing observasi, dan diperoleh 4 kelompok wilayah berdasarkan faktor-faktor yang berpengaruh secara signifikan terhadap TPT di Provinsi Sulawesi Selatan.

Kata Kunci: Regresi Nonparameterik, Spasial, *Spline Truncated*, Temporal, Tingkat Pengangguran Terbuka



ABSTRACT

The Open Unemployment Rate is an indicator that measures the percentage of the unemployed in the total workforce. TPT in South Sulawesi Province shows spatial and temporal heterogeneity at the same time. In addition, the pattern of relationships between variables spreads randomly, so it is difficult to estimate the regression curve. To overcome this problem, a spatial regression and nonparametric regression development model was used, namely the Nonparametric Geographically and Temporally Weighted Regression method with a truncated spline approach. This research aims to model TPT in South Sulawesi Province using the Nonparametric Geographically and Temporally Weighted Regression method with a truncated spline approach. Based on the analysis results, the best model obtained uses a fixed kernel exponential weighting function with knot points with the minimum GCV value, namely the combination of knot points. The local models obtained varied for each observation, and 4 regional groups were obtained based on factors that significantly influenced TPT in South Sulawesi Province.

Keywords: Nonparametric Regression, Open Unemployment Rate, Spatial, Spline Truncated, Temporal.



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
LEMBAR PERNYATAAN KEOTENTIKAN	iii
HALAMAN PENGESAHAN	v
KATA PENGANTAR	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	ix
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1 Tingkat Pengangguran Terbuka.....	5
2.2 Regresi Linear Berganda.....	6
2.3 Regresi <i>Spline Truncated</i>	7
2.4 Analisis Data Spasial.....	8
2.5 Uji Heterogenitas Spasial dan Temporal	9
2.6 <i>Geographically and Temporally Weighted Regression</i>	10
2.7 Jarak Spasio-Temporal	11
2.8 Matriks Pembobot	13
2.8 <i>Nonparametric Geographically Weighted Regression</i>	14
2.9 Pengujian Kesesuaian Model	15
Pengujian Signifikansi Parameter	15
2.1 Uji Signifikansi Parameter Simultan.....	15
2.2 Uji Signifikansi Parameter Parsial	16



BAB III METODOLOGI PENELITIAN	17
3.1 Sumber Data.....	17
3.2 Variabel Penelitian	17
3.3 Metode Analisis.....	17
BAB IV HASIL DAN PEMBAHASAN.....	19
4.1 Estimasi Parameter Model NGTWR dengan Pendekatan <i>Spline Truncated</i>	19
4.2 Penerapan Model NGTWR pada TPT di Provinsi Sulawesi Selatan.....	22
4.2.1 Eksplorasi Data	22
4.2.2 Pola Hubungan Antarvariabel	23
4.2.3 Pengujian Heterogenitas Spasial dan Temporal	25
4.2.4 Pemilihan Titik Knot Optimum.....	26
4.2.5 Pemilihan Model Terbaik	27
4.2.7 Pemodelan NGTWR dengan Pendekatan <i>Spline Truncated</i>	27
4.2.8 Pengujian Hipotesis.....	28
4.2.9 Interpretasi Model	31
BAB V KESIMPULAN DAN SARAN	34
5.1 Kesimpulan	34
5.2 Saran.....	34
DAFTAR PUSTAKA.....	35
LAMPIRAN.....	38



DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Jarak Spasio-Temporal (Huang dkk., 2010)	12
Gambar 4. 1 Peta Sebaran TPT.....	22
Gambar 4. 2 Pola Hubungan Antarvariabel.....	24
Gambar 4. 3 Uji Heterogenitas Temporal.....	25
Gambar 4. 4 Estimasi Parameter τ	27
Gambar 4. 5 Estimasi Parameter μ dan λ	28
Gambar 4. 6 Pemetaan Hasil Uji Signifikansi Parameter secara Parsial	31



DAFTAR TABEL

Tabel 3. 1 Variabel Penelitian	17
Tabel 4. 1 Pendeteksian Multikolinearitas.....	24
Tabel 4. 2 Uji Heterogenitas Spasial	25
Tabel 4. 3 Pemilihan Titik Knot Optimal	26
Tabel 4. 4 Letak Titik Knot.....	26
Tabel 4. 5 Pemilihan Model Terbaik.....	27
Tabel 4. 6 Hasil Uji Signifikansi Parameter secara Parsial	30



DAFTAR LAMPIRAN

Lampiran 1. Data Variabel Respon dan Variabel-variabel Prediktor	39
Lampiran 2. Hasil Iterasi Parameter τ	40
Lampiran 3. Hasil Iterasi Parameter μ dan λ	41
Lampiran 4. Matriks Pembobot Spasio-temporal	43
Lampiran 5. Pemilihan Titik Knot Optimum (1 Titik Knot)	44
Lampiran 6. Pemilihan Titik Knot Optimum (2 Titik Knot)	45
Lampiran 7. Pemilihan Titik Knot Optimum (3 Titik Knot)	46
Lampiran 8. Hasil Estimasi Parameter	47
Lampiran 9. Nilai thitung Setiap Observasi	50



BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengangguran merupakan salah satu masalah yang dialami di Indonesia. Pengangguran merupakan suatu masalah yang sangat rumit karena dipengaruhi oleh beberapa faktor dan sangat mempengaruhi faktor-faktor ekonomi yang lain (Muslim, 2014). Salah satu indikator yang digunakan untuk mengukur tingkat pengangguran di Indonesia adalah Tingkat Pengangguran Terbuka (TPT). TPT merupakan persentase jumlah pengangguran terhadap jumlah angkatan kerja. Menurut Badan Pusat Statistik (BPS), berdasarkan Survei Angkatan Kerja Nasional (Sakernas) pada Agustus 2022, TPT di Indonesia mencapai angka 5.97%. Sulawesi Selatan menduduki peringkat ke-22 dari 35 provinsi berdasarkan TPT di Indonesia. Menurut BPS Provinsi Sulawesi Selatan, pada Agustus 2022 TPT di Sulawesi Selatan mencapai 4.51%. Meskipun tergolong rendah dibanding provinsi-provinsi lain di Indonesia, namun masih terdapat masalah-masalah pengangguran jika diuraikan menurut kabupaten/kota. Oleh karena itu, untuk menganalisis faktor-faktor yang mempengaruhi TPT di Sulawesi Selatan, dapat dilakukan menggunakan metode analisis regresi.

Analisis regresi merupakan metode analisis yang digunakan untuk melihat pengaruh variabel prediktor terhadap variabel respon yang dipilih. Metode analisis regresi yang terdiri dari dua atau lebih variabel prediktor disebut analisis regresi linear berganda. Dalam metode regresi linear berganda, salah satu yang paling penting adalah estimasi parameter regresi. Salah satu metode dalam estimasi parameter regresi adalah metode kuadrat terkecil (*Ordinary Least Square*) dengan meminimalkan jumlah *error* (Sopianti, 2013). Dalam metode kuadrat terkecil terdapat asumsi klasik yang harus dipenuhi. Salah satu asumsi yang termasuk dalam asumsi klasik adalah asumsi kehomogenan ragam yang berarti bahwa setiap residual memiliki varians yang sama (Gujarati, 2003). Kehomogenan ragam akan diperoleh pada pemodelan data spasial karena adanya pengaruh lokasi pada analisis. Sehingga metode yang tepat digunakan adalah metode analisis spasial (Lutfiani dkk., 2019).



Geographically Weighted Regression (GWR) merupakan salah satu pemodelan regresi yang digunakan data yang menggunakan pendekatan titik (Lutfiani dkk., 2019). GWR merupakan metode yang tepat untuk membentuk suatu analisis regresi yang bersifat lokal pada setiap lokasi, karena melakukan pembobotan berdasarkan posisi dari suatu lokasi terhadap lokasi lainnya (Hutabarat & Saefuddin, 2011). Lutfiani dkk. (2019) memodelkan faktor-faktor yang mempengaruhi jumlah penduduk miskin di Provinsi Jawa Timur dengan menggunakan metode GWR, dan Dewi dkk. (2016) melakukan pemodelan faktor penyebab kecelakaan lalu lintas di Jawa Timur berdasarkan metode *Geographically Weighted Regression*. Penelitian-penelitian terdahulu banyak melakukan pemodelan dalam bentuk regresi parametrik, karena model membentuk pola data yang diketahui. Akan tetapi, sering kali diperoleh kasus pola data yang tidak beraturan. Sehingga penggunaan model regresi nonparametrik lebih akurat untuk digunakan (Jao dkk., 2022).

Regresi nonparametrik merupakan pemodelan regresi apabila tidak diketahui kurva regresinya. Berbeda dengan regresi parametrik yang sejak awal diketahui bentuk kurva regresinya seperti linear, kuadratik, kubik, eksponensial, dan lain sebagainya. Regresi nonparametrik bersifat fleksibel karena model yang akan menentukan sendiri kurva regresinya. Oleh karena itu, diperlukan pendekatan untuk mengestimasi kurva regresi nonparametrik (Islamiyati, 2014). Salah satu metode regresi nonparametrik yang dapat digunakan adalah metode *spline truncated*. Metode *spline truncated* memiliki beberapa keunggulan berupa bentuk kurva yang fleksibel dan lebih sederhana secara matematis (Pratiwi & Wijaya, 2022). Nurcahyani dkk. (2019) memodelkan rata-rata lama sekolah di Pulau Jawa menggunakan regresi *spline truncated*, Ramli dkk. (2020) memodelkan nilai *Return on Assets (ROA)* menggunakan regresi nonparametrik *spline truncated* pada data longitudinal, Jao dkk. (2022) melakukan pemodelan regresi *spline Poisson* pada tingkat kematian bayi di Sulawesi Selatan, dan Wongkar & Fahmuddin (2023) menganalisis faktor-faktor yang mempengaruhi TPT di Provinsi Sulawesi Selatan menggunakan metode regresi nonparametrik *Spline Truncated*.

...gunaan metode regresi nonparametrik telah banyak dikembangkan pada ...ng tidak memperhatikan adanya perbedaan karakteristik lingkungan dan ...ografis yang dapat menyebabkan adanya variasi pada data (Sifriyani,



Kartiko, dkk., 2018). Oleh sebab itu, untuk mengatasi permasalahan tersebut dilakukan pengembangan metode pada bidang regresi nonparametrik dan regresi spasial. Rifaldi (2023) melakukan pemodelan *Geographically Weighted Regression* dengan pendekatan *spline truncated* pada kasus balita gizi kurang di Provinsi Sulawesi Selatan, Serena dkk. (2021) melakukan penelitian mengenai aplikasi pendekatan *spline truncated* dalam model GWR pada pencemaran daerah aliran Sungai Mahakam, dan Saputri dkk. (2023) melakukan penelitian mengenai aplikasi metode *Nonparametric Geographically Weighted Regression* pada Tingkat Pengangguran Terbuka di Indonesia.

Selain efek spasial, efek temporal juga dapat digunakan pada proses analisis untuk memperoleh model yang lebih informatif karena akan membentuk model pada setiap lokasi dan waktu pengamatan (Djuraidah dkk., 2021). Namun, model GWR hanya dapat mengakomodasi adanya efek keragaman spasial, sehingga untuk menganalisis data yang mengandung efek spasial sekaligus temporal digunakan metode *Geographically and Temporally Weighted Regression* (GTWR) (Debataraja dkk., 2021). Berbeda dengan model GWR, pada model GTWR informasi spasial dan temporal digabungkan pada matriks pembobot untuk mengakomodasi adanya efek spasial dan temporal. Berdasarkan uraian tersebut, penelitian ini akan difokuskan pada pengkajian model GTWR dengan pendekatan *spline truncated* yang akan digunakan untuk memodelkan faktor-faktor yang mempengaruhi TPT di Provinsi Sulawesi Selatan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah yang akan dibahas pada penelitian ini adalah sebagai berikut:

1. Bagaimana hasil estimasi model *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated*?
2. Bagaimana pemodelan TPT di Provinsi Sulawesi Selatan menggunakan metode *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated*?



Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Orde polinomial yang diterapkan pada model regresi nonparametrik adalah linear ($l = 1$)
2. Penentuan titik knot optimum menggunakan metode *Generalized Cross Validation* (GCV).
3. Pemilihan model terbaik dengan membandingkan nilai *Cross Validation* (CV) dengan menggunakan matriks pembobot *fixed Kernel*.
4. Menggunakan metode *Maximum Likelihood Weighting Estimation* (MLWE) dalam melakukan estimasi parameter.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah sebagai berikut:

1. Memperoleh hasil estimasi model *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated*.
2. Memperoleh model TPT di Provinsi Sulawesi Selatan menggunakan metode *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated*.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Bagi peneliti, menambah wawasan terkait pemodelan *Nonparametric Geographically and Temporally Weighted Regression* dengan pendekatan *spline truncated* serta memperoleh informasi terkait penerapan metode pada kasus TPT di Provinsi Sulawesi Selatan.
2. Bagi pemerintah, menjadi suatu acuan dalam pengambilan kebijakan untuk mengatasi masalah pengangguran di Provinsi Sulawesi Selatan.



BAB II

TINJAUAN PUSTAKA

2.1 Tingkat Pengangguran Terbuka

Pengangguran (*unemployment*) didefinisikan sebagai suatu keadaan seseorang yang tergolong dalam kategori angkatan kerja (*labor force*) tidak memiliki pekerjaan dan secara aktif sedang mencari pekerjaan (Nanga, 2001). Menurut Badan Pusat Statistik (BPS), indikator untuk mengukur masalah pengangguran adalah TPT. TPT merupakan persentase sejumlah pengangguran terhadap jumlah angkatan kerja. TPT mengindikasikan besarnya persentase angkatan kerja yang termasuk dalam pengangguran. TPT mampu menunjukkan kemampuan ekonomi untuk menciptakan lapangan kerja yang mampu menyerap persediaan (*supply*) tenaga kerja yang ada. Semakin tinggi nilai indikator ini, semakin banyak persediaan tenaga kerja yang tidak termanfaatkan. Indikator ini dapat memberikan sinyal tentang kinerja pasar kerja dan berlangsungnya kondisi ekonomi tertentu, seperti resesi, perubahan siklus bisnis dan teknologi, dan lain-lain. Tingkat pengangguran terbuka diperoleh dengan Persamaan (2.1):

$$TPT = \frac{\text{Jumlah pengangguran}}{\text{Jumlah angkatan kerja}} \times 100\% \quad (2.1)$$

Berdasarkan penelitian-penelitian terdahulu, beberapa faktor yang diduga mempengaruhi TPT adalah sebagai berikut:

1. Indeks Pembangunan Manusia (IPM) merupakan indikator yang terdiri dari 3 komposit antara lain pendidikan, kesehatan dan penghidupan yang layak. Nilai IPM yang tinggi menandakan kualitas sumber daya manusia yang tinggi pula (Hasibuan, 2023).
2. Angka Partisipasi Murni (APM) merupakan indikator yang digunakan untuk memperoleh proporsi dari penduduk berdasarkan kelompok usia tertentu bersekolah pada jenjang yang seharusnya (Sari dkk., 2022).
3. Jumlah Penduduk Miskin adalah indikator jumlah penduduk yang tidak memiliki akses pada sarana dan prasarana yang memadai, serta kualitas hidup jauh di bawah standar sehingga sulit untuk mencapai kualifikasi SDM memadai (Sa'adah & Ardyan, 2016).



4. Produk Domestik Regional Bruto (PDRB) yang meningkat merupakan indikator pertumbuhan ekonomi. Pertumbuhan ekonomi yang meningkat dapat memberikan kesempatan bagi industri untuk berkembang dan menciptakan lapangan pekerjaan. Peningkatan PDRB diharapkan dapat menyerap lebih banyak tenaga kerja, karena dengan peningkatan PDRB dapat meningkatkan kapasitas produksi (Muslim, 2014).

2.2 Regresi Linear Berganda

Metode regresi adalah metode yang digunakan untuk menyatakan pola hubungan antara satu variabel respon dan satu atau lebih variabel prediktor. Regresi linear merupakan metode yang memodelkan hubungan antara variabel respon (y) dan variabel prediktor ($x_1, x_2, x_3, \dots, x_m$) secara linear. Secara umum, model regresi linear berganda dengan m variabel prediktor adalah sebagai berikut (Gujarati, 2003):

$$y_i = \beta_0 + \sum_{j=1}^m \beta_j x_{ji} + \varepsilon_i \quad (2.2)$$

$$i = 1, 2, 3, \dots, n ; j = 1, 2, 3, \dots, m$$

dengan

- y_i : Variabel respon pada pengamatan ke- i
- x_{ji} : Variabel prediktor ke- j pada pengamatan ke- i
- β_0 : Koefisien intersep dari model
- β_j : Koefisien regresi
- ε_i : galat pengamatan ke- i

Dari Persamaan (2.2), dapat diuraikan menjadi persamaan-persamaan berikut ini:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \dots + \beta_m x_{m1} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \dots + \beta_m x_{m2} + \varepsilon_2 \\ y_3 &= \beta_0 + \beta_1 x_{13} + \beta_2 x_{23} + \dots + \beta_m x_{m3} + \varepsilon_3 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \dots + \beta_m x_{mn} + \varepsilon_n \end{aligned} \quad (2.3)$$

in bentuk matriks, Persamaan (2.3) dapat dituliskan sebagai berikut:



$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{m1} \\ 1 & x_{12} & x_{22} & \dots & x_{m2} \\ 1 & x_{13} & x_{23} & \dots & x_{m3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{mn} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dan dapat dinyatakan sebagai berikut:

$$Y = X\beta + \varepsilon \quad (2.4)$$

Metode estimasi parameter pada model regresi linear berganda, dapat dilakukan menggunakan metode *Ordinary Least Square* (OLS). Estimasi parameter regresi menggunakan metode OLS, dapat dilakukan dengan cara meminimumkan jumlah kuadrat error berdasarkan Persamaan (2.4) (Gujarati, 2003).

$$\min(\varepsilon' \varepsilon = (Y - X\beta)'(Y - X\beta)) = 0$$

$$\frac{\partial \varepsilon' \varepsilon}{\partial \beta} = 0$$

$$Y'Y - 2Y'X\beta + \beta X'X\beta = 0$$

Sehingga diperoleh hasil estimasi parameter sebagai berikut:

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.5)$$

Hasil estimasi menggunakan dapat dipercaya jika memenuhi asumsi pada OLS atau biasa disebut dengan asumsi klasik. Uji Asumsi klasik adalah analisis yang dilakukan untuk mengetahui masalah-masalah asumsi klasik di dalam sebuah model regresi linear OLS (Mardiatmoko, 2020).

2.3 Regresi *Spline Truncated*

Spline merupakan penggalan (*piecewise*) polinomial tersegmen (*truncated*) yang kontinu. Fungsi *spline* mampu menyesuaikan diri lebih efektif terhadap pola data yang naik atau turun secara tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus (Eubank, 1999). Dalam analisis regresi nonparametrik *spline*, jika terdapat satu respon dan satu prediktor maka regresi tersebut disebut regresi nonparametrik *spline* univariabel. Sedangkan apabila terdapat satu respon dengan lebih dari satu prediktor, maka regresi tersebut disebut regresi nonparametrik *spline* multivariabel (Afa dkk., 2018).

terdiri dari segmen-segmen polinomial yang berbeda ditautkan pada knot k_r dan kontinu sehingga bersifat fleksibel dibandingkan polinomial biasa & Sihombing, 2020). Titik knot adalah titik perpaduan bersama dari



potongan-potongan tersebut atau titik yang menunjukkan terjadinya perubahan-perubahan perilaku kurva pada interval-interval yang berbeda (Jao dkk., 2022). Jika diberikan model nonparametrik,

$$y_i = f(x_i) + \varepsilon_i ; i = 1,2,3, \dots, n \quad (2.6)$$

Maka model umum regresi *spline truncated* orde ke-p dengan satu prediktor adalah sebagai berikut:

$$y_i = \sum_{l=0}^p \beta_l X_i^l + \sum_{k=1}^r \beta_{p+k} (X_i - K_k)_+^p + \varepsilon_i \quad (2.7)$$

dengan β adalah parameter, l adalah orde dengan $l = 1, 2, \dots, p$, K adalah titik knot dengan $k = 1, 2, 3, \dots, r$, dan $i = 1, 2, \dots, n$ adalah banyaknya pengamatan.

Sedangkan bentuk umum regresi *spline truncated* orde ke-p dengan multiprediktor adalah sebagai berikut:

$$y_i = \beta_0 + \sum_{j=1}^m \left(\sum_{l=1}^p \beta_{jl} X_{ji}^l + \sum_{k=1}^r \beta_{j(p+k)} (X_{ji} - K_{jk})_+^p \right) + \varepsilon_{ji} \quad (2.8)$$

dengan $j = 1, 2, \dots, m$ adalah banyaknya variabel prediktor, dan

$$(X_{ji} - K_{jk})_+^p = \begin{cases} (X_{ji} - K_{jk})^p & , x_{ji} > K_{jk} \\ 0 & , x_{ji} \leq K_{jk} \end{cases}$$

Metode yang dapat digunakan untuk memilih jumlah titik knot optimal adalah dengan menggunakan metode *Generalized Cross Validation* (GCV). Nilai GCV dapat diperoleh menggunakan Persamaan (2.9) (Budiantara, 2006).

$$GCV[K_{jk}] = \frac{n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\left(n^{-1} \cdot tr \left[\mathbf{I} - \mathbf{A}[K_{jk}] \right] \right)^2} \quad (2.9)$$

dengan, $\mathbf{A}[K_{jk}] = \mathbf{X}_j[K_{jk}] (\mathbf{X}_j^T[K_{jk}] \mathbf{X}_j[K_{jk}])^{-1} \mathbf{X}_j^T[K_{jk}]$. Pemilihan titik knot yang optimal dilakukan dengan melihat nilai GCV dari masing-masing orde dan titik knot kemudian dipilih nilai $GCV[K_{jk}]$ yang paling minimum (Jao dkk., 2022).

2.4 Analisis Data Spasial

Data spasial merupakan data yang disajikan dalam posisi geografis dari suatu objek, berkaitan dengan lokasi, bentuk, dan hubungan diantaranya dalam ruang

nyajian data geografis dilakukan dengan menggunakan titik, garis, dan data spasial dapat berupa data diskrit atau kontinu dan dapat juga memiliki spasial beraturan (*regular*) maupun tak beraturan (*irregular*) (Alfiana, 2010).



Analisis data spasial tidak dapat dilakukan secara global, artinya setiap lokasi mempunyai karakteristik sendiri. Sebagian besar pendekatan analisisnya merupakan eksplorasi data yang disajikan dalam bentuk peta tematik (Pfeiffer dkk., 2008). Pada model regresi spasial terdapat dua efek spasial yaitu, ketergantungan spasial (*spatial dependence*) dan keragaman spasial (*spatial heterogeneity*) (Anselin, 2001).

Adanya ketergantungan spasial menunjukkan bahwa pengamatan pada lokasi yang satu dipengaruhi oleh pengamatan di lokasi yang lain. Oleh karena itu, perlu dilakukan identifikasi kebenaran efek spasial pada data yang digunakan. Salah satu pengujian yang dapat digunakan untuk mengidentifikasi adanya ketergantungan spasial, yaitu dengan pengujian *Moran's I*. Pengujian *Moran's I* merupakan pengujian yang dilakukan untuk melihat pengaruh pengamatan di suatu lokasi terhadap pengamatan di lokasi lain yang letaknya saling berdekatan. Sedangkan, keragaman spasial merujuk pada adanya keberagaman dalam hubungan secara kewilayahan (Dewi & Zain, 2016).

Berdasarkan tipe data, pemodelan spasial dapat dibedakan menjadi pemodelan dengan pendekatan titik dan area (Pertwi dkk., 2012). Data titik menunjukkan lokasi yang berupa titik, misalnya berupa titik pada longitude (garis bujur) dan latitude (garis lintang). Data garis digunakan untuk menggambarkan suatu hal yang memiliki jalur panjang, bukan suatu area, misalnya garis kontur, jaringan jalan, sungai listrik, dan sebagainya. Data area menunjukkan lokasi yang berupa luasan, seperti suatu negara, kabupaten, kota dan sebagainya (Lutfiani dkk., 2019).

2.5 Uji Heterogenitas Spasial dan Temporal

Heterogenitas spasial merujuk pada adanya keberagaman dalam hubungan secara kewilayahan. Pada hampir setiap kasus dianggap bahwa akan didapatkan hubungan yang berbeda pada setiap lokasi pengamatan. Akibatnya, parameter global yang diduga dari data geografis tidak menggambarkan dengan baik fenomena geografis pada lokasi tertentu (Dewi & Zain, 2016). Pengujian heterogenitas spasial dapat dilakukan menggunakan uji *Breusch-Pagan*, dengan rumus sebagai berikut:



i. Hipotesis

$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma$ (Tidak terdapat keragaman spasial)

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

ii. Statistik Uji

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (2.10)$$

dengan elemen faktor \mathbf{f} adalah

$$f_i = \frac{e_i^2}{\sigma^2} - 1$$

dengan

e_i : galat untuk pengamatan ke- i

σ^2 : ragam galat e_i

\mathbf{Z} : matriks berukuran $n \times (p + 1)$ berisi vektor dari yang telah terstandarisasi (z) untuk setiap lokasi dan p merupakan banyaknya variabel prediktor.

iii. Kriteria keputusan

Tolak H_0 jika $BP > \chi^2(p)$ atau jika $p - \text{value} < \alpha$, dengan p adalah banyaknya prediktor.

Heterogenitas temporal merujuk pada adanya perubahan nilai dari waktu ke waktu. Jika terdapat heterogenitas temporal pada data, maka dapat disimpulkan bahwa terdapat perbedaan karakteristik data setiap tahun (Sholihin dkk., 2017). Pemodelan GTWR akan lebih baik jika pada data terdindikasi adanya heterogenitas temporal dan spasial secara bersamaan (Debataraja dkk., 2021). Uji heterogenitas temporal dapat dilakukan dengan melihat *boxplot*. Keragaman dapat dilihat pada ukuran *box* dan juga perubahan *outlier* pada tiap waktu. Perubahan nilai outlier dan besar *box* mengindikasikan bahwa terdapat heterogenitas temporal pada data (Sholihin dkk., 2017).

2.6 Geographically and Temporally Weighted Regression



GTWR adalah salah satu analisis regresi yang merupakan pengembangan dari GTWR dan dapat digunakan untuk memodelkan data yang terindikasi adanya heterogenitas temporal dan spasial secara bersamaan (Wang, 2006). Berbeda dengan

GWR, model GTWR akan menghasilkan model untuk setiap lokasi dan waktu sehingga akan didapatkan model yang lebih representatif (Djuraidah dkk., 2021). Model GTWR menggabungkan informasi spasial dan temporal ke dalam matriks pembobot dalam mengidentifikasi adanya keragaman dari sisi spasial dan temporal (Debataraja dkk., 2021). Model GTWR dapat ditulis sebagai berikut:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{j=1}^m \beta_j(u_i, v_i, t_i) x_{ij} + \varepsilon_i ; i = 1, 2, 3, \dots, n \quad (2.11)$$

dengan

- y_i : Nilai observasi variabel respon untuk lokasi ke- i
- (u_i, v_i, t_i) : Menyatakan koordinat letak geografis (*longitude, latitude*) dari lokasi pengamatan ke- i dan waktu ke- i
- $\beta_0(u_i, v_i, t_i)$: Nilai *intercept* model GWR pada lokasi dan waktu pengamatan ke- i
- $\beta_j(u_i, v_i, t_i)$: Koefisien regresi variabel prediktor ke- j pada lokasi dan waktu pengamatan ke- i
- x_{ij} : Nilai observasi variabel prediktor ke- j pada lokasi pengamatan ke- i
- ε_i : Error pengamatan ke- i yang diasumsikan identik, independen dan berdistribusi normal dengan mean nol dan varian konstan σ^2

Estimasi parameter model GTWR dilakukan dengan metode Weighted Least Square (WLS) yaitu dengan memberikan penimbang/pembobot yang berbeda pada setiap lokasi dan waktu pengamatan. Bentuk estimasi parameter dari model GTWR untuk setiap lokasi dituliskan pada Persamaan (2.12) berikut (Adiningrum dkk., 2018):

$$\hat{\beta}(u_i, v_i, t_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{y} \quad (2.12)$$

dengan $\mathbf{W}(u_i, v_i, t_i)$ adalah matriks pembobot spasio-temporal observasi ke- i .

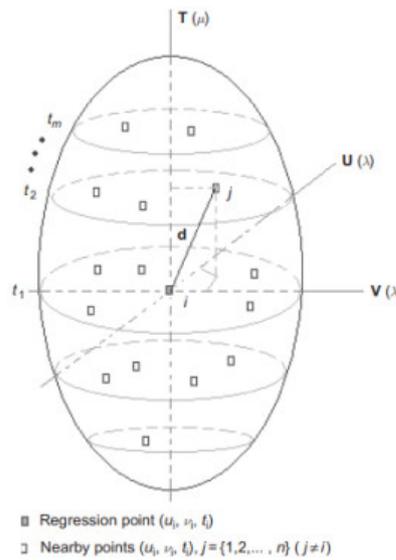
2.7 Jarak Spasio-Temporal

Peran pembobot pada model GTWR sangat penting karena nilai pembobot ini berkaitan dengan letak data observasi satu dengan lainnya (Dewi & Zain, 2016). Matriks pembobot spasio-temporal (\mathbf{W}) dapat diperoleh berdasarkan informasi jarak dari lokasi pengamatan (neighborhood) atau jarak antara satu lokasi dan waktu pengamatan dengan lokasi dan waktu pengamatan yang lain. Elemen matriks pembobot GTWR,



yaitu W_{ij} ditentukan berdasarkan kedekatan titik regresi i dengan titik pengamatan j . Titik pengamatan yang lebih dekat ke titik regresi diberi bobot lebih besar daripada titik pengamatan yang lebih jauh (Widiyanti dkk., 2014).

Definisi kedekatan spasial-temporal dalam sistem koordinat menjadi hal yang penting karena kedekatan titik observasi pada GTWR terdiri dari dua unsur, yaitu spasial dan temporal (Adiningrum dkk., 2018). Dalam memahami jarak spasio-temporal digunakan sistem koordinat elipsoida untuk mengukur jarak antar observasi. Ilustrasi jarak spasio-temporal adalah sebagai berikut:



Gambar 2. 1 Ilustrasi Jarak Spasio-Temporal (Huang dkk., 2010)

Fungsi jarak spasio-temporal terdiri dari fungsi jarak spasial (dS) dan fungsi jarak temporal (dT) sebagai berikut:

$$(d_{ij}^{ST}) = \lambda(d_{ij}^S)^2 + \mu(d_{ij}^T)^2 \tag{2.13}$$

dengan λ dan μ adalah skala penyeimbang antara efek spasial dan temporal (Huang dkk., 2010). Sehingga diperoleh jarak jarak Euclidean sebagai berikut:

$$(d_{ij}^{ST}) = \lambda \{ (u_i - u_j)^2 + (v_i - v_j)^2 \} + \mu (t_i - t_j)^2 \tag{2.14}$$

Misalkan τ merupakan parameter rasio μ/λ dengan $\lambda \neq 0$, maka Persamaan (2.14) dapat dituliskan sebagai berikut:

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

Parameter τ memiliki peran yang sangat penting dalam membentuk matriks jarak spasio-temporal. Parameter τ berperan dalam memperbesar atau



memperkecil efek temporal agar sesuai dengan efek spasial. Parameter τ optimal dapat diperoleh dengan melihat nilai Koefisien Determinasi (R^2).

2.8 Matriks Pembobot

Pembobot merupakan matriks diagonal yang elemen-elemen diagonalnya merupakan sebuah nilai bobot dari setiap titik lokasi pengamatan. Fungsi dari matriks pembobot adalah untuk menentukan atau menaksir parameter yang berbeda pada setiap titik lokasi pengamatan. Salah satu metode pembobotan yang biasa digunakan adalah fungsi kernel (Lutfiani dkk., 2019). Fungsi kernel terbagi menjadi beberapa fungsi sebagai berikut:

i. Fungsi *Gaussian*

$$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right) \tag{2.16}$$

ii. Fungsi *Exponential*

$$w_{ij} = \sqrt{\exp\left(-\left(\frac{d_{ij}}{h}\right)^2\right)} \tag{2.17}$$

iii. Fungsi *Tricube*

$$w_{ij} = \begin{cases} \left[\left(1 - \frac{d_{ij}}{h}\right)^3\right]^3 & ; \text{jika } d_{ij} < h \\ 0 & ; \text{lainnya} \end{cases} \tag{2.18}$$

iv. Fungsi *Bi-square*

$$w_{ij} = \begin{cases} \left[\left(1 - \frac{d_{ij}}{h}\right)^2\right]^2 & ; \text{jika } d_{ij} < h \\ 0 & ; \text{lainnya} \end{cases} \tag{2.19}$$

dengan d_{ij} adalah jarak *euclidean* antara observasi ke- i dengan observasi ke- j dan h adalah radius dari titik pusat observasi ke- i (Rahman, 2021).

Pemilihan bandwidth optimum menjadi sangat penting karena akan mempengaruhi ketepatan model terhadap data, yaitu mengatur varians dan bias dari

Salah satu metode yang digunakan untuk menentukan bandwidth optimum metode cross validation (CV) (Dewi & Zain, 2016). Nilai *Cross Validation* at diperoleh dengan rumus pada Persamaan (2.20).



$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(h))^2 \tag{2.20}$$

2.8 Nonparametric Geographically Weighted Regression

Model nonparametrik pada data spasial dengan pendekatan *spline truncated* merupakan pengembangan dari regresi nonparametrik yang estimasi parameter berlaku lokal pada setiap lokasi pengamatan. *Spline truncated* merupakan suatu metode yang digunakan untuk memecahkan masalah regresi spasial untuk kasus kurva regresi tidak diketahui (Sifriyani, Budiantara, dkk., 2018). Asumsi yang digunakan pada model NGWR adalah *error* berdistribusi normal dengan rata-rata nol dan variansi $\sigma^2(u_i, v_i)$ pada setiap lokasi. Salah satu faktor penting dalam pemodelan adalah koordinat lokasi (u_i, v_i) karena akan digunakan untuk pembentukan matriks pembobot yang digunakan dalam menaksir parameter (Serena dkk., 2021).

Misalkan diberikan data $(X_{1i}, X_{2i}, \dots, X_{mi})$ dan hubungan antara $(x_{1i}, x_{2i}, \dots, x_{mi})$ dan y diasumsikan mengikuti model regresi nonparametrik sebagai berikut:

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{mi}) + \varepsilon_i; \quad i = 1, 2, 3, \dots, n$$

dengan y_i adalah variabel respon, $f(x_{1i}, x_{2i}, \dots, x_{mi})$ adalah fungsi regresi yang tidak diketahui, dan ε_i adalah galat dari model. Maka fungsi regresi $f(x_{1i}, x_{2i}, \dots, x_{mi})$ yang diestimasi dengan pendekatan *spline truncated* pada setiap titik pengamatan (u_i, v_i) didefinisikan sebagai berikut:

$$f = \beta_0(u_i, v_i) + \sum_{j=1}^m \sum_{l=1}^p \beta_{jl}(u_i, v_i) \cdot x_{ji}^l + \sum_{j=1}^m \sum_{k=1}^r \beta_{j,p+k}(u_i, v_i) \cdot (x_{ji} - K_{jk})_+^p$$

Secara matematis hubungan antara variabel respon dan variabel prediktor pada observasi ke- i untuk model NGWR dinyatakan sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^m \sum_{l=1}^p \beta_{jl}(u_i, v_i) \cdot x_{ji}^l + \sum_{j=1}^m \sum_{k=1}^r \beta_{j,p+k}(u_i, v_i) \cdot (x_{ji} - K_{jk})_+^p + \varepsilon_i \tag{2.21}$$

: variabel respon, dengan $i = 1, 2, 3, \dots, n$

: variabel prediktor, dengan $j = 1, 2, 3, \dots, m$



- K_{jk} : titik knot, dengan $k = 1, 2, 3, \dots, r$
- $\beta_{jl}(u_i, v_i)$: parameter regresi, dengan $l = 1, 2, 3, \dots, p$
- $\beta_{j,p+k}(u_i, v_i)$: parameter regresi *spline truncated*

2.9 Pengujian Kesesuaian Model

Uji kesesuaian model (*Goodness of Fit*) dilakukan untuk mengetahui model terbaik antara model NGTWR menggunakan estimator *spline truncated* dan regresi global. Pengujian kesesuaian model dapat dilakukan dengan prosedur sebagai berikut (Sifriyani, Budiantara, dkk., 2018):

i. Hipotesis

$$H_0: \beta_{jl}(u_i, v_i, t_i) = \beta_{jl} \text{ atau } \beta_{j,p+k}(u_i, v_i, t_i) = \beta_{j,p+k}$$

$$H_1: \text{Minimal ada satu } \beta_{jl}(u_i, v_i, t_i) \neq \beta_{jl} \text{ atau } \beta_{j,p+k}(u_i, v_i, t_i) \neq \beta_{j,p+k}$$

ii. Statistik Uji

$$V = \frac{\left(\frac{\mathbf{y}^T \mathbf{S} \mathbf{y}}{n - mp - 1} \right)}{\left(\frac{\mathbf{y}^T \mathbf{D}(u_i, v_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \boldsymbol{\xi})^T (\mathbf{I} - \boldsymbol{\xi}))} \right)} \quad (2.22)$$

dengan matriks $\mathbf{S} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$, $\mathbf{D}(u_i, v_i, t_i) = (\mathbf{I} - \boldsymbol{\xi})^T (\mathbf{I} - \boldsymbol{\xi})$, dan $\boldsymbol{\xi} = \mathbf{X}(\mathbf{X}'\mathbf{W}(u_i, v_i, t_i)\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}(u_i, v_i, t_i)$ dengan \mathbf{y} adalah vektor variabel respon dan \mathbf{I} adalah matriks identitas.

iii. Kriteria Penolakan

$$\text{Tolak } H_0 \text{ jika } F_{hitung}(V^*) > F_{(\alpha; db_1; db_2)} \text{ atau } p\text{-value} < \alpha$$

2.10 Pengujian Signifikansi Parameter

2.10.1 Uji Signifikansi Parameter Simultan

Pengujian signifikansi parameter secara simultan dapat dilakukan dengan prosedur sebagai berikut (Sifriyani dkk., 2018b):

i. Hipotesis

$$H_0: \beta_{11}(u_i, v_i, t_i) = \beta_{12}(u_i, v_i, t_i) = \dots = \beta_{mp}(u_i, v_i, t_i) =$$

$$\beta_{j,p+k}(u_i, v_i, t_i) = 0$$

$$\text{Ada satu } \beta_{jl}(u_i, v_i, t_i) \neq 0 \text{ atau } \beta_{j,p+k}(u_i, v_i, t_i) \neq 0$$



ii. Statistik Uji

$$V^* = \frac{\left(\frac{\mathbf{y}^T \mathbf{M}(u_i, v_i, t_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \mathbf{B}_\omega)^T (\mathbf{I} - \mathbf{B}_\omega))} \right)}{\left(\frac{\mathbf{y}^T \mathbf{D}(u_i, v_i, t_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \boldsymbol{\xi})^T (\mathbf{I} - \boldsymbol{\xi}))} \right)} \quad (2.23)$$

dengan $\mathbf{M}(u_i, v_i, t_i) = (\mathbf{I} - \mathbf{B}_\omega)^T (\mathbf{I} - \mathbf{B}_\omega)$ dengan \mathbf{y} adalah vektor variabel respon dan \mathbf{I} adalah matriks identitas.

iii. Kriteria Penolakan

Tolak H_0 jika $F_{hitung}(V^*) > F_{(\alpha; db_1; db_2)}$ atau $p - value < \alpha$

2.10.2 Uji Signifikansi Parameter Parsial

Pengujian signifikansi parameter dilakukan secara parsial dengan prosedur sebagai berikut (Sifriyani, Budiantara, dkk., 2018):

i. Hipotesis

$$H_0 : \beta_{jl} = 0 \text{ dan } \beta_{j,p+k} = 0$$

$$H_1 : \beta_{jl}(u_i, v_i, t_i) \neq 0 \text{ atau } \beta_{j,p+k}(u_i, v_i, t_i) \neq 0$$

ii. Statistik Uji

$$t = \frac{\hat{\eta}(u_i, v_i, t_i)}{SE(\hat{\eta}(u_i, v_i, t_i))} \quad (2.24)$$

dengan $\hat{\eta}(u_i, v_i, t_i)$ adalah estimator dari $\eta(u_i, v_i, t_i)$ dan $SE(\hat{\eta}(u_i, v_i, t_i))$ adalah elemen diagonal ke $k + 1$ dari matriks $(\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \hat{\sigma}(u_i, v_i, t_i)$ dan \mathbf{X} adalah matriks ukuran $n \times [1 + (d \times p) + (d \times r)]$ yang memuat variabel prediktor dan variabel prediktor yang mempunyai komponen *spline truncated*.

iii. Kriteria Penolakan

Tolak H_0 jika $|t| > t_{(\frac{\alpha}{2}; ab)}$ atau $p - value < \alpha$

