

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis regresi linear merupakan salah satu analisis statistik yang dapat menjelaskan hubungan antarvariabel. Terdapat dua jenis variabel dalam analisis regresi linear, yaitu variabel respon (y) sebagai variabel yang dipengaruhi dan variabel prediktor (x) sebagai variabel yang memengaruhi. Hasil dari analisis regresi dinyatakan dalam persamaan matematis yang disebut sebagai model regresi (Walpole, 1982). Model regresi yang terbentuk tersebut bersifat global karena diberlakukan sama di semua lokasi pengamatan atau hubungan antarvariabel respon dan prediktor dianggap stasioner (Monalisa, 2022). Namun, setiap lokasi memiliki karakteristik yang berbeda sehingga analisis berdasarkan pada pemodelan regresi yang bersifat global belum cukup dalam menjelaskan data yang sebenarnya. Hal ini terkait dengan adanya efek spasial yang memungkinkan terjadinya keragaman atau heterogenitas spasial. Kondisi ini menunjukkan adanya keragaman antarlokasi sehingga setiap lokasi pengamatan berbeda-beda dalam hal struktur dan parameter hubungan (Arif dkk., 2019). Analisis *Geographically Weighted Regression* (GWR) merupakan perluasan regresi global yang setiap lokasi pengamatannya menggunakan pembobot yang berbeda sehingga mampu mengakomodir permasalahan tersebut (Daulay & Simamora, 2023).

Model GWR mengevaluasi heterogenitas spasial dan menghasilkan estimasi parameter model yang bersifat lokal untuk setiap lokasi pengamatan (Fotheringham dkk., 2002). Akan tetapi, model ini hanya mengakomodasi efek heterogenitas spasial tanpa mempertimbangkan adanya pengaruh unsur waktu atau efek temporal. Oleh karena itu, digunakan model *Geographically and Temporally Weighted Regression* (GTWR) yang merupakan pengembangan GWR sehingga tidak hanya mempertimbangkan efek heterogenitas spasial, tetapi juga efek temporal. Parameter model yang dihasilkan menggunakan GTWR pun bersifat lokal untuk setiap lokasi dan waktu pengamatan (Bai dkk., 2022). Sementara itu, terkadang tidak semua variabel berpengaruh secara lokal karena terdapat kemungkinan variabel yang berpengaruh secara global. Terkait dengan hal ini, model *Mixed Geographically and Temporally Weighted Regression* (MGTWR) digunakan untuk menangani variasi pengaruh dari variabel prediktor (Yang dkk., 2021).

Model MGTWR merupakan hasil penggabungan antara model regresi linear dengan GTWR yang dapat digunakan untuk menangani data yang memiliki pengaruh spasial dan temporal, serta mengidentifikasi variabel yang bersifat lokal dan global (Djuraidah dkk., 2021a). Melalui model MGTWR, diperoleh estimasi parameter yang bersifat global dan sebagian lainnya bersifat lokal sesuai dengan kondisi lokasi dan waktu pengamatan data (Djuraidah dkk., 2021b). Model ini dirancang untuk menangani heterogenitas spasial dan temporal yang berbeda dari variabel prediktor sehingga mampu memberikan estimasi koefisien yang lebih akurat dan relevan beserta informasi yang lebih mendetail akan efek variabel prediktor terhadap variabel respon (Hong dkk., 2021). Meskipun demikian, analisis spasial MGTWR masih

memiliki kelemahan, salah satunya tidak dapat mengatasi multikolinearitas sehingga pemodelan yang dilakukan kurang optimal (Suritman dkk., 2023).

Multikolinearitas merupakan kondisi yang terjadi ketika dua atau lebih variabel prediktor dalam suatu model regresi memiliki korelasi yang tinggi (Tsagris & Pandis, 2021). Adanya masalah multikolinearitas menyebabkan nilai galat baku untuk koefisien regresi menjadi lebih besar dari yang sebenarnya dan menyebabkan tanda pada koefisien regresi berlawanan dengan teori (Lestari dkk., 2020). Salah satu solusi untuk mengatasi masalah multikolinearitas adalah dengan memberi penalti pada koefisien regresi menggunakan regresi ridge. Regresi ini memungkinkan prediksi model dengan mengurangi ragam estimasi koefisien melalui penambahan penalti (Ölmez & Yaganoglu, 2023). Pengembangan model MGTWR yang menggunakan pendekatan regresi ridge dalam analisis untuk mengatasi masalah multikolinearitas disebut *Mixed Geographically and Temporally Weighted Ridge Regression* (MGTWRR). Pemodelan dengan MGTWRR diharapkan mengatasi heterogenitas spasial dan temporal, terutama dalam kondisi multikolinearitas sehingga hasil estimasi parameter yang diperoleh menjadi lebih akurat. Adapun pengembangan model MGTWRR dapat ditinjau melalui pengaplikasiannya pada data tingkat pengangguran terbuka di Provinsi Sulawesi Selatan.

Tingkat Pengangguran Terbuka (TPT) adalah indikator penting dalam analisis pasar penyerapan tenaga kerja yang menunjukkan proporsi angkatan kerja yang tidak memiliki pekerjaan, tetapi aktif mencari pekerjaan. Penyerapan tenaga kerja di Indonesia telah mengalami peningkatan positif yang tercermin dari penurunan TPT sebesar 1.17% dalam tiga tahun terakhir (Kemnaker, 2024). Sulawesi Selatan merupakan salah satu provinsi dengan TPT yang relatif rendah, dalam hal ini 4.9% (Bappenas, 2024). Meskipun demikian, masih diperlukan perhatian untuk menurunkan angka pengangguran lebih lanjut dan meningkatkan kesempatan kerja berkualitas di Sulawesi Selatan. Salah satu cara untuk menurunkan angka ini adalah dengan mengetahui faktor-faktor yang memengaruhi TPT di Sulawesi Selatan. Faktor-faktor yang memengaruhi pengangguran sangat kompleks karena menyangkut berbagai sektor, termasuk sosial, ekonomi, dan pendidikan yang saling berhubungan (Adriyanto dkk., 2020). Analisis menggunakan regresi spasial, khususnya MGTWRR merupakan salah satu analisis yang dapat digunakan untuk mengetahui faktor-faktor yang berpengaruh terhadap TPT.

Penerapan regresi spasial yang mempertimbangkan variabel lokal dan global dengan pendekatan regresi ridge telah dilakukan oleh Suritman dkk. (2023) dalam memodelkan kasus kemiskinan di Sulawesi Selatan pada tahun 2016, tetapi tidak mengakomodasi efek temporal di dalamnya. Hasil penelitian menunjukkan bahwa analisis regresi spasial dengan pendekatan regresi ridge lebih baik dibandingkan tanpa pendekatan tersebut. Penelitian lainnya dilakukan oleh Prajanati dkk. (2022) dalam memodelkan indeks pembangunan manusia di Nusa Tenggara Barat dengan hasil penelitian bahwa model analisis spasial tanpa pendekatan regresi ridge tidak lebih baik dibandingkan model dengan pendekatan tersebut. Penerapan

MGTWR pun telah dilakukan oleh Liu dkk. (2017) dalam mengeksplorasi variasi spasial temporal dari perspektif global dan lokal melalui perbandingan MGTWR dengan MGWR dan GWR. Hasil penelitian tersebut menunjukkan bahwa model MGTWR memberikan hasil yang lebih akurat dibandingkan model lainnya. Penelitian-penelitian sebelumnya hanya terbatas pada aplikasi regresi spasial dengan pendekatan regresi ridge tanpa mengakomodasi temporal sehingga pada penelitian ini baik spasial maupun temporal diakomodasi dengan pendekatan regresi ridge untuk mengatasi multikolinearitas.

Data TPT beserta faktor-faktor yang diduga memengaruhi angka tersebut di Sulawesi Selatan telah dikumpulkan untuk setiap kabupaten/kota dalam lima tahun terakhir dengan latar geografis yang berbeda. Hal ini memungkinkan adanya keragaman nilai TPT secara spasial. Begitu pun faktor-faktor yang berpengaruh akan berbeda bergantung pada karakteristik kabupaten/kota. Namun, terdapat kemungkinan bahwa beberapa faktor berpengaruh secara global di semua lokasi pengamatan. Oleh karena itu, penelitian ini akan berfokus pada pembentukan model MGTWRR dan identifikasi faktor-faktor yang memengaruhi TPT untuk setiap kabupaten/kota di Sulawesi Selatan pada tahun 2019-2023.

1.2 Tujuan dan Manfaat

Berdasarkan latar belakang, tujuan penelitian ini sebagai berikut:

1. Memperoleh estimasi parameter model *Mixed Geographically and Temporally Weighted Ridge Regression* untuk data tingkat pengangguran terbuka di Sulawesi Selatan tahun 2019-2023.
2. Memperoleh faktor-faktor yang berpengaruh secara signifikan terhadap tingkat pengangguran terbuka di Sulawesi Selatan tahun 2019-2023.

Adapun manfaat yang ingin dicapai sebagai berikut:

1. Meningkatkan pengetahuan dan menambah wawasan pembaca maupun penulis tentang pemodelan menggunakan MGTWRR serta penerapannya pada data TPT di Sulawesi Selatan.
2. Memberikan dasar acuan bagi pembuat kebijakan dalam merumuskan strategi yang lebih tepat untuk menurunkan angka pengangguran di Sulawesi Selatan.

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut:

1. Data yang digunakan adalah data TPT tingkat kabupaten/kota di Sulawesi Selatan tahun 2019-2023 dengan tujuh variabel prediktor yang diduga memengaruhi TPT.
2. Penentuan variabel global berdasarkan persentase estimasi parameter GTWR yang berada dalam interval kepercayaan regresi global, dalam hal ini lebih dari 70%.

3. Penentuan jarak antarlokasi pengamatan menggunakan jarak *euclidean*, sedangkan penentuan fungsi pembobot di antara *fixed kernel* berdasarkan nilai *Cross Validation* (CV) minimum.
4. Penentuan model regresi terbaik berdasarkan nilai *Akaike Information Criterion* (AIC) dan *Adjusted R²*.

1.4 Landasan Teori

1.4.1 Regresi Linear

Analisis regresi linear merupakan analisis atau studi terkait model hubungan antara suatu variabel respon (y) dengan satu atau lebih variabel prediktor (x). Adapun analisis regresi linear dapat dibagi menjadi dua bentuk, yaitu regresi linear sederhana dan linear berganda. Secara umum, regresi linear dapat dinyatakan seperti pada Persamaan (1) (Sartika dkk., 2020):

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

dengan

- y_i : variabel respon yang diamati
- x_{ik} : variabel prediktor ke- k pada pengamatan ke- i
- β_0 : konstanta
- β_k : koefisien regresi dari variabel prediktor ke- k
- ε_i : galat pengamatan ke- i
- n : banyak pengamatan

Persamaan (1) juga dapat dinyatakan dalam bentuk matriks:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2)$$

atau dalam bentuk Persamaan (3):

$$Y = X\beta + \varepsilon \quad (3)$$

dengan Y sebagai vektor kolom ($n \times 1$) pengamatan, X sebagai matriks pengamatan berukuran ($n \times (p + 1)$), β sebagai vektor kolom dari koefisien regresi berukuran $((p + 1) \times 1)$, dan ε sebagai vektor kolom dari galat berukuran ($n \times 1$).

Metode yang dapat digunakan dalam mengestimasi parameter β adalah metode estimasi kuadrat terkecil atau yang juga dikenal sebagai *Ordinary Least Square* (OLS) melalui Persamaan (4):

$$\hat{\beta} = (X^t X)^{-1} X^t Y \quad (4)$$

dengan asumsi bahwa estimasi parameter regresi berlaku untuk setiap pengamatan atau bersifat secara global (Draper & Smith, 1998). Estimasi parameter dengan OLS dilakukan hanya ketika karakteristik data memenuhi asumsi klasik.

1.4.2 Heterogenitas Spasial

Heterogenitas spasial merupakan suatu kondisi ketika variabel prediktor yang sama memberikan respon yang tidak sama untuk setiap lokasi dalam satu wilayah

pengamatan (Sanusi dkk., 2023). Kondisi ini menunjukkan adanya keragaman antarlokasi sehingga setiap lokasi pengamatan berbeda-beda dalam hal struktur dan parameter hubungan (Arif dkk., 2019). Identifikasi heterogenitas spasial dapat dilakukan menggunakan uji Breusch-Pagan (BP) dengan hipotesis uji sebagai berikut.

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

$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma_j^2 \text{ dengan } i \neq j$ (Terdapat gejala heterogenitas spasial)

Statistik uji yang digunakan pada uji BP dinyatakan pada Persamaan (5).

$$BP = \frac{1}{2} \mathbf{f}^t \mathbf{Z}(\mathbf{Z}^t \mathbf{Z})^{-1} \mathbf{Z}^t \mathbf{f} \sim \chi^2_{(\alpha, p)} \quad (5)$$

dengan $\mathbf{f}$ sebagai vektor ($n \times 1$) yang elemennya $f_i = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1\right)$, ε_i^2 sebagai galat untuk observasi ke- i dari hasil estimasi regresi menggunakan OLS, σ^2 sebagai ragam dari hasil estimasi regresi menggunakan OLS, $\mathbf{Z}$ sebagai matriks variabel prediktor berukuran ($n \times (p + 1)$), dan p sebagai banyaknya variabel prediktor. Pengambilan keputusan pada uji BP adalah menolak H_0 jika nilai $BP > \chi^2_{\alpha, p}$ atau $p - \text{value} < \alpha$ (Rahman dkk., 2023).

1.4.3 Geographically Weighted Regression

Model GWR merupakan model yang dikembangkan dari regresi linear dengan mempertimbangkan lokasi pengamatan sebagai pembobot (Marizal & Monalisa, 2022). Adanya pertimbangan aspek geografi tersebut mengindikasikan bahwa GWR diterapkan pada data yang memiliki pengaruh heterogenitas spasial (Zhao dkk., 2021). Setiap lokasi memiliki karakteristik yang memengaruhi hubungan antarvariabel dan GWR mengakomodasi variasi tersebut dengan membiarkan parameter model bervariasi secara lokal sesuai dengan kondisi di setiap lokasi. Oleh karena itu, parameter model pada GWR bersifat lokal atau bervariasi secara spasial (Trevisan dkk., 2021). Secara umum, model matematis dari GWR dapat dilihat pada Persamaan (6).

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n ; k = 1, 2, \dots, p \quad (6)$$

dengan y_i sebagai variabel respon pada lokasi pengamatan ke- i , (u_i, v_i) sebagai koordinat (*longitude*, *latitude*) lokasi ke- i , $\beta_0(u_i, v_i)$ sebagai *intercept* pada lokasi pengamatan ke- i , $\beta_k(u_i, v_i)$ sebagai koefisien regresi variabel prediktor ke- k pada pengamatan ke- i yang menunjukkan variabilitas spasial, x_{ik} sebagai variabel prediktor ke- k pada lokasi pengamatan ke- i , dan ε_i sebagai galat pengamatan ke- i (Saputra & Radam, 2023).

Pemodelan dengan GWR membutuhkan pembobot spasial (w_{ij}) karena berperan penting dalam estimasi parameter. Estimasi parameter titik spasial cenderung lebih dipengaruhi oleh titik-titik yang berada dekat dengan lokasi tersebut dibandingkan titik-titik yang berada lebih jauh (Moller & Vihrs, 2022). Adapun pembobot spasial ini bertujuan untuk menghasilkan estimasi parameter yang

bervariasi bergantung pada jarak antarlokasi pengamatan. Jarak tersebut dihitung berdasarkan lokasi geografis atau titik koordinat masing-masing pengamatan dan biasanya menggunakan jarak *euclidean*. Jarak ini dinotasikan dengan d_{ij} menyatakan jarak antara lokasi i pada koordinat (u_i, v_i) dan lokasi j pada koordinat (u_j, v_j) yang dapat diperoleh melalui Persamaan (7) (Primasari dkk., 2022).

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (7)$$

Secara umum, skema pembobotan spasial dapat dilakukan dengan menggunakan fungsi kernel untuk mendapatkan estimasi parameter model GWR (Fotheringham dkk., 2002). Terdapat beberapa fungsi kernel yang dapat digunakan dalam pembobotan, di antaranya:

1. Fungsi Kernel *Exponential*

$$w_j(u_i, v_i) = \sqrt{\exp\left(-\left(\frac{d_{ij}}{h}\right)^2\right)} \quad (8)$$

2. Fungsi Kernel *Gaussian*

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right) \quad (9)$$

3. Fungsi Kernel *Bisquare*

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^2\right)^2, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \quad (10)$$

4. Fungsi Kernel *Tricube*

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^3\right)^3, & \text{untuk } d_{ij} \leq h \\ 0, & \text{untuk } d_{ij} > h \end{cases} \quad (11)$$

dengan

$w_j(u_i, v_i)$: bobot lokasi ke- j pada lokasi pengamatan ke- i

h : parameter *bandwidth*

d_{ij} : jarak *euclidean*

Terkait dengan penggunaan fungsi kernel, terdapat *bandwidth* yang perlu diperhatikan. Secara konsep, *bandwidth* adalah jarak (h) yang mewakili radius suatu lingkaran terhadap titik pusat lokasi dan digunakan sebagai dasar penentuan bobot untuk setiap pengamatan (Fadliana & Drajat, 2021). Lokasi titik-titik pengamatan yang berada dalam radius lingkaran dianggap masih berpengaruh terhadap pembentukan model di lokasi pengamatan ke- i (Taek dkk., 2023). Oleh karena itu, dalam model GWR, pemilihan *bandwidth* optimum sangat penting. Metode CV adalah salah satu metode yang umumnya digunakan dalam memilih *bandwidth* optimum berdasarkan nilai CV yang paling minimum (Weku dkk., 2022). Secara matematis, pemilihan *bandwidth* optimum menggunakan metode CV dapat dinyatakan dalam Persamaan (12) (Fotheringham dkk., 2002).

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\#i}(h))^2 \quad (12)$$

dengan $\hat{y}_{\#i}(h)$ sebagai nilai estimasi dari lokasi pengamatan ke- i yang diperoleh tanpa mempertimbangkan lokasi pengamatan itu sendiri.

1.4.4 Geographically and Temporally Weighted Regression

Model GTWR adalah pengembangan dari model GWR yang tidak dapat menangani data non-stasioner baik secara spasial maupun temporal secara bersamaan. GTWR menggabungkan informasi spasial dan temporal dalam matriks pembobot untuk mengidentifikasi variabilitas spasial dan temporal (Sifriyani dkk., 2022). Model GTWR dapat dinyatakan sebagai berikut (He & Yang, 2023):

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i ; i = 1, 2, \dots, n ; k = 1, 2, \dots, p \quad (13)$$

dengan

- y_i : Variabel respon untuk lokasi pengamatan ke- i
- (u_i, v_i, t_i) : Koordinat letak geografis (*longitude, latitude*) dan waktu dari lokasi pengamatan ke- i
- $\beta_0(u_i, v_i, t_i)$: *Intercept* model pada lokasi pengamatan ke- (u_i, v_i) dan waktu pengamatan ke- t_i
- $\beta_k(u_i, v_i, t_i)$: Koefisien regresi variabel prediktor ke- k pada lokasi pengamatan ke- (u_i, v_i) dan waktu pengamatan ke- t_i
- x_{ij} : Variabel prediktor ke- k pada lokasi pengamatan ke- i
- ε_i : Galat pada lokasi pengamatan ke- i

Parameter model yang dihasilkan pada model GTWR bersifat lokal untuk setiap lokasi dan waktu pengamatan (Debataraja dkk., 2021). Adapun estimasi parameter model GTWR dilakukan menggunakan metode kuadrat terkecil terboboti atau *Weighted Least Square* (WLS) yang memberikan pembobot berbeda untuk setiap lokasi dan waktu pengamatan. Estimasi parameter model GTWR untuk setiap lokasi dan waktu pengamatan dilakukan melalui operasi matriks seperti pada Persamaan (14):

$$\hat{\beta}(u_i, v_i, t_i) = (X^t W(u_i, v_i, t_i) X)^{-1} X^t W(u_i, v_i, t_i) Y \quad (14)$$

dengan $W(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in})$ sebagai matriks pembobot, dalam hal ini matriks diagonal dari bobot jarak spasial temporal pada lokasi pengamatan (u_i, v_i) dan waktu ke- t . Subskrip n menunjukkan jumlah pengamatan (Ma dkk., 2021).

1.4.5 Jarak Spasial-Temporal

Kedekatan spasial dan temporal adalah unsur utama dalam model GTWR sehingga definisi dan pengukuran kedekatan ini dalam sistem koordinat menjadi hal yang penting karena kedekatan titik pengamatan terdiri dari kedua unsur tersebut. Pentingnya kedekatan atau jarak spasial-temporal terletak pada kemampuan model yang dirancang dalam menangkap variabilitas yang terjadi baik secara spasial

maupun temporal (Sifriyani dkk., 2022). Jarak spasial-temporal ini dapat dihitung menggunakan sistem koordinat elipsoidal untuk menentukan interval spasial dan temporal antara titik data pengamatan dengan titik data di sekitarnya. Spesifikasi yang tepat dari fungsi atenuasi jarak spasial-temporal digunakan untuk menghitung matriks pembobot secara ruang dan waktu (Liu dkk., 2020). Pembentukan fungsi jarak spasial-temporal (d_{ij}^{ST}) melibatkan kombinasi antara fungsi jarak spasial (d_{ij}^S) dan fungsi jarak temporal (d_{ij}^T) dengan menggunakan Persamaan (15):

$$(d_{ij}^{ST})^2 = \lambda^* (d_{ij}^S)^2 + \mu (d_{ij}^T)^2 \quad (15)$$

dengan λ^* dan μ sebagai parameter untuk menyeimbangkan pengaruh antara jarak spasial dan temporal, $(d_{ij}^S)^2 = (u_i - u_j)^2 + (v_i - v_j)^2$, dan $(d_{ij}^T)^2 = (t_i - t_j)^2$ sehingga diperoleh jarak *euclidean* spasial-temporal (Bai dkk., 2022):

$$(d_{ij}^{ST})^2 = \lambda^* \{(u_i - u_j)^2 + (v_i - v_j)^2\} + \mu (t_i - t_j)^2 \quad (16)$$

Misalkan τ adalah parameter rasio yang didefinisikan sebagai $\frac{\mu}{\lambda^*}$ dengan kondisi $\lambda^* \neq 0$, maka diperoleh Persamaan (17):

$$\frac{(d_{ij}^{ST})^2}{\lambda^*} = \{(u_i - u_j)^2 + (v_i - v_j)^2\} + \tau (t_i - t_j)^2 \quad (17)$$

Parameter τ berfungsi untuk meningkatkan atau mengurangi efek jarak temporal terhadap jarak spasial. Parameter ini dapat diperoleh menggunakan kriteria CV minimum dengan menginisiasi nilai awal seperti pada Persamaan (18) (Sifriyani dkk., 2022):

$$CV(\tau) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(\tau))^2 \quad (18)$$

1.4.6 Mixed Geographically and Temporally Weighted Regression

Model GTWR hanya mempertimbangkan pengaruh lokal dari variabel prediktor secara spasial-temporal terhadap variabel respon. Sementara itu, dalam beberapa kasus, model tidak hanya mempertimbangkan pengaruh lokal saja, tetapi juga terdapat pengaruh global dari variabel prediktor (Nuramaliyah dkk., 2019). Oleh karena itu, Liu dkk. (2017) melakukan pengembangan model GTWR dengan mengadopsi pengaruh lokal dan global dari variabel prediktor yang kemudian disebut dengan model MGTWR. Adapun model MGTWR dengan asumsi bahwa intersep bersifat lokal dinyatakan dalam Persamaan (19):

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=r+1}^p \beta_l(u_i, v_i, t_i) x_{ik} + \sum_{k=1}^r \beta_g x_{ik} + \varepsilon_i ; \quad i = 1, 2, \dots, n \quad (19)$$

dengan

- y_i : variabel respon
- x_{ik} : variabel prediktor ke- k pada lokasi pengamatan ke- i
- $\beta_0(u_i, v_i, t_i)$: *Intercept* model pada lokasi pengamatan ke- (u_i, v_i) dan waktu pengamatan ke- t_i

- β_g : Koefisien regresi global
 $\beta_l(u_i, v_i, t_i)$: Koefisien regresi lokal pada lokasi pengamatan ke- i
 r : jumlah variabel prediktor dengan pengaruh global
 p : jumlah variabel prediktor dengan pengaruh lokal
 ε : Galat pada lokasi pengamatan ke- i

Estimasi parameter model MGTWR dilakukan dengan menggunakan metode WLS, dalam hal ini pembobot yang berbeda diberikan untuk setiap lokasi pengamatan. Estimasi parameter model dapat dituliskan dalam bentuk matriks seperti pada Persamaan (20):

$$Y = X_l \beta_l(u_i, v_i, t_i) + X_g \beta_g + \varepsilon \quad (20)$$

dengan X_l sebagai matriks variabel prediktor lokal, X_g sebagai matriks variabel prediktor global, $\beta_l(u_i, v_i, t_i)$ sebagai vektor parameter variabel prediktor lokal pada lokasi pengamatan ke- i , β_g sebagai vektor parameter variabel prediktor global, dan ε sebagai galat pengamatan.

Melalui cara yang sama pada model MGWR, estimasi parameter model MGWR juga dilakukan dengan metode estimasi dua langkah, yaitu untuk $\beta_l(u_i, v_i, t_i)$ dan β_g (Yang dkk., 2021). Langkah pertama adalah menuliskan model MGTWR dalam bentuk GTWR seperti pada Persamaan (21):

$$\tilde{Y} = Y - X_g \beta_g = X_l \beta_l(u_i, v_i, t_i) + \varepsilon \quad (21)$$

Sehingga diperoleh estimasi parameter model MGTWR dapat dilihat pada Persamaan (22):

$$\hat{\beta}_l(u_i, v_i, t_i) = [X_l^t W(u_i, v_i, t_i) X_l]^{-1} X_l^t W(u_i, v_i, t_i) \tilde{Y} \quad (22)$$

Misalkan $x_{li}^t = (1, x_{1i}, x_{2i}, \dots, x_{ri})$ adalah elemen baris ke- i dari matriks X_l , maka nilai estimasi untuk $\tilde{Y}$ pada (u_i, v_i, t_i) untuk seluruh pengamatan dapat diperoleh dengan Persamaan (23):

$$\hat{\tilde{Y}} = (\hat{\tilde{Y}}_1, \hat{\tilde{Y}}_2, \dots, \hat{\tilde{Y}}_n)^t = S_l \tilde{Y} = S_l (Y - X_g \beta_g) \quad (23)$$

dengan

$$S_l = \begin{bmatrix} x_{l1}^t (X_l^t W(u_1, v_1, t_1) X_l)^{-1} X_{l1}^t W(u_1, v_1, t_1) \\ x_{l2}^t (X_l^t W(u_2, v_2, t_2) X_l)^{-1} X_{l2}^t W(u_2, v_2, t_2) \\ \vdots \\ x_{ln}^t (X_l^t W(u_n, v_n, t_n) X_l)^{-1} X_{ln}^t W(u_n, v_n, t_n) \end{bmatrix}$$

dengan $x_{l1}^t (X_l^t W(u_1, v_1, t_1) X_l)^{-1} X_{l1}^t W(u_1, v_1, t_1)$ menghasilkan vektor $(1 \times n)$ sehingga matriks S_l berukuran $(n \times n)$.

Estimasi $\hat{\beta}_l(u_i, v_i, t_i)$ kemudian disubstitusikan ke dalam model MGTWR Persamaan (20) sehingga diperoleh Persamaan (24):

$$\hat{\beta}_g = [X_g^t (I - S_l)^t (I - S_l) X_g]^{-1} X_g^t (I - S_l)^t (I - S_l) Y \quad (24)$$

Estimasi $\hat{\beta}_g$ disubstitusi ke Persamaan (25) sehingga diperoleh estimasi untuk koefisien lokal pada lokasi pengamatan (u_i, v_i, t_i) :

$$\hat{\beta}_l(u_i, v_i, t_i) = [X_l^t W(u_i, v_i, t_i) X_l]^{-1} X_l^t W(u_i, v_i, t_i) (Y - X_g \hat{\beta}_g) \quad (25)$$

Oleh karena itu, nilai estimasi variabel respon untuk n lokasi pengamatan dinyatakan melalui Persamaan (26):

$$\hat{Y} = SY \quad (26)$$

dengan

$$S = S_l + (I - S_l) X_g [X_g^t (I - S_l)^t (I - S_l) X_g]^{-1} X_g^t (I - S_l)^t (I - S_l) \quad (27)$$

Estimator $\hat{\beta}_l(u_i, v_i, t_i)$ dan $\hat{\beta}_g$ merupakan estimator tak bias untuk $\beta_l(u_i, v_i, t_i)$ dan β_g (Liu dkk., 2017).

1.4.7 Multikolinearitas

Multikolinearitas merupakan istilah dari asumsi yang menunjukkan hubungan linear antarvariabel prediktor (Shaheen dkk., 2023). Hal ini terjadi ketika model regresi menggunakan beberapa variabel prediktor yang berkorelasi secara signifikan satu sama lain (Shrestha, 2020). Adanya multikolinearitas menyebabkan interval kepercayaan dan hasil regresi yang bias untuk setiap variabel prediktor sehingga model yang dihasilkan kurang tepat (Hosmer dkk., 2013). Selain itu, multikolinearitas dapat menyebabkan inflasi atau deflasi galat standar dari koefisien sehingga koefisien dapat menjadi tidak signifikan atau signifikan secara palsu. Hal ini pun dapat mengakibatkan perubahan tanda koefisien, yaitu pengaruh variabel yang seharusnya negatif dapat menjadi positif begitu pun sebaliknya. Oleh karena itu, multikolinearitas harus dipertimbangkan atau diperbaiki (Gokmen dkk., 2020).

Multikolinearitas dapat diidentifikasi melalui nilai *Variance Inflation Factor* (VIF). Nilai VIF digunakan untuk mengukur peningkatan nilai varians koefisien regresi yang diestimasi jika variabel prediktor saling berkorelasi. Nilai VIF dinyatakan dalam Persamaan (28).

$$VIF_k = \frac{1}{1 - R_k^2} ; k = 1, 2, \dots, p \quad (28)$$

dengan R_k^2 merupakan koefisien determinasi antarvariabel prediktor dan $(1 - R_k^2)$ merupakan nilai toleransi. Semakin rendah toleransi, semakin besar kemungkinan adanya multikolinearitas, begitu pun sebaliknya. Adapun nilai $VIF > 10$ mengindikasikan bahwa koefisien regresi diestimasi secara lemah karena terdapat multikolinearitas (Kim, 2019).

1.4.8 Regresi Ridge

Regresi ridge merupakan salah satu metode regresi yang dapat digunakan untuk mengatasi masalah multikolinearitas. Metode ini adalah pengembangan dari metode *Least Squares* (LS) yang meminimalkan jumlah kuadrat dari kesalahan dengan menambahkan pengaturan atau tetapan bias (Prajanati dkk., 2022). Tetapan tersebut bertindak sebagai penalti yang diberikan pada ukuran koefisien regresi untuk mengendalikan kestabilan penduga kuadrat terkecil yang mengalami multikolinearitas sehingga hasil yang diperoleh memiliki ragam yang lebih kecil daripada hasil dengan menggunakan metode kuadrat terkecil (Chen dkk., 2020). Penambahan tetapan berupa bilangan positif kecil, dalam hal ini suku λ pada $(X^t X)$

dilakukan agar varians estimasi yang diperoleh lebih kecil sehingga multikolinearitas melemah (Saputro dkk., 2021). Secara matematis, estimasi parameter regresi ridge dinyatakan dalam Persamaan (29) atau (30):

$$(X^t X + \lambda I) \hat{\beta}_R = X^t Y \quad (29)$$

atau

$$\hat{\beta}_{RR} = (X^t X + \lambda I)^{-1} X^t Y \quad (30)$$

dengan $0 < \lambda \leq 1$ sebagai konstanta atau parameter bias yang dipilih untuk memastikan bahwa nilai $\hat{\beta}_{RR}$ stabil. Nilai konstanta tersebut memberikan *mean square error* yang lebih rendah dibandingkan metode LS (Ölmez & Yaganoglu, 2023).

Pemilihan tetapan nilai λ dapat menggunakan metode Estimator Lawless & Wang (LW) dengan tetapan bias global mengikuti Persamaan (31):

$$\lambda = \frac{p \hat{\theta}}{\hat{\beta}_{OLS}^t (X^t X) \hat{\beta}_{OLS}} \quad (31)$$

dan tetapan bias lokal mengikuti Persamaan (32):

$$\lambda(u_i, v_i, t_i) = \frac{p \hat{\theta}}{\hat{\beta}_{WLS_i}^t (X^t X) \hat{\beta}_{WLS_i}} \quad (32)$$

dengan p sebagai jumlah variabel prediktor, $\hat{\theta}$ sebagai *mean square error*, $\hat{\beta}_{OLS}$ merupakan estimasi yang diperoleh menggunakan metode OLS, dan $\hat{\beta}_{WLS_i}$ merupakan estimasi yang diperoleh pada pengamatan ke- i dengan metode WLS melalui model MGTWR (Lawless & Wang, 1976).

1.4.9 Pengujian Signifikansi Parameter

Uji signifikansi parameter dilakukan untuk mengetahui pengaruh dari variabel global dan lokal prediktor terhadap respon pada model MGTWR. Pengujian dilakukan sebanyak dua kali, yaitu untuk variabel global dan lokal. Pengujian signifikansi parameter pada variabel global dilakukan dengan hipotesis:

$$H_0 : \beta_k = 0 \text{ untuk setiap } k = 1, 2, \dots, r$$

$$H_1 : \beta_k \neq 0$$

Statistik uji:

$$t_{ghit} = \frac{\hat{\beta}_k}{\hat{\sigma} \sqrt{g_{kk}}} \sim t_{(\alpha/2, df)} \quad (33)$$

dengan g_{kk} sebagai elemen diagonal ke- k dari matriks GG^t . Matriks $G = [X_g^t (I - S_l)^t (I - S_l) X_g]^t X_g^t (I - S_l)^t (I - S_l)$ dan $\hat{\sigma}^2 = \frac{Y^t (I - S)^t (I - S) Y}{tr((I - S)^t (I - S))}$. H_0 ditolak ketika $|t_{ghit}| > t_{\alpha/2, df}$ pada taraf signifikansi α dengan $df = \frac{u_1^2}{u_2}$ dan $u_i = tr([(I - S)^t (I - S)]^i)$, $i = 1, 2$.

Uji signifikansi parameter selanjutnya untuk mengetahui variabel lokal yang berpengaruh signifikan terhadap respon pada model MGWR. Untuk menguji signifikansi suatu variabel lokal digunakan hipotesis yaitu:

$H_0 : \beta_k(u_i, v_i, t_i) = 0$ untuk setiap $k = r + 1, r + 2, \dots, p$ dan $i = 1, 2, \dots, n$

$H_1 : \beta_k(u_i, v_i, t_i) \neq 0$

Statistik uji:

$$t_{hit} = \frac{\hat{\beta}_k(u_i, v_i, t_i)}{\hat{\sigma} \sqrt{m_{kk}}} \sim t_{(\alpha/2, df)} \quad (34)$$

dengan m_{kk} sebagai elemen diagonal ke- k dari matriks $\mathbf{MM}^t$. Matriks $\mathbf{M} = [\mathbf{X}_l^t \mathbf{W}(u_i, v_i, t_i) \mathbf{X}_l]^{-1} \mathbf{X}_l^t \mathbf{W}(u_i, v_i, t_i) (\mathbf{I} - \mathbf{X}_g \mathbf{G})$. Pengambilan keputusan tolak H_0 jika $|t_{hit}| > t_{\alpha/2, df}$ pada taraf signifikansi α dengan $df = \frac{u_1^2}{u_2}$ dan $u_i = \text{tr}[(\mathbf{I} - \mathbf{S})^t (\mathbf{I} - \mathbf{S})^i], i = 1, 2$.

1.4.10 Pemilihan Model Terbaik

Penentuan model terbaik dapat dilakukan dengan berbagai metode, dua di antaranya yang umum digunakan adalah AIC dan *Adjusted R*². Nilai AIC didefinisikan melalui Persamaan (35):

$$AIC_c = 2n \ln(\hat{\sigma}) + n \ln(2\pi) + n + \text{tr}(\mathbf{S}) \quad (35)$$

dengan subskrip c menunjukkan bahwa AIC terkoreksi, $\hat{\sigma}$ merupakan estimator nilai standar deviasi dari galat, dan $\mathbf{S}$ merupakan matriks hasil proyeksi dari model. Nilai AIC terkecil atau minimum mengindikasikan model terbaik (Primasari dkk., 2022).

Adapun nilai *Adjusted R*² dapat diperoleh menggunakan Persamaan (36):

$$R_{adj}^2 = \frac{(n-1)R^2 - p}{n - p + 1} \quad (36)$$

dengan $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ merupakan koefisien determinasi, n merupakan jumlah pengamatan, dan p merupakan jumlah variabel prediktor (Chung, 2010).

1.4.11 Tingkat Pengangguran Terbuka

Pengangguran merupakan suatu istilah untuk orang yang tidak bekerja sama sekali, bekerja kurang dua hari dalam seminggu, maupun sedang mencari pekerjaan secara aktif sebagai akibat dari ketidakseimbangan angka angkatan kerja dengan jumlah lapangan kerja yang tersedia (Rokhim, 2023). Indikator untuk mengukur masalah pengangguran disebut TPT yang didefinisikan sebagai persentase jumlah pengangguran terhadap jumlah angkatan kerja (Nurjannah dkk., 2022). Nilai TPT dapat diukur menggunakan Persamaan (37):

$$TPT = \frac{\text{Jumlah pengangguran}}{\text{Jumlah angkatan kerja}} \times 100\% \quad (37)$$

Indikator ini tidak hanya mengukur jumlah pengangguran, tetapi juga memberikan gambaran tentang dinamika pasar tenaga kerja dan kondisi ekonomi suatu wilayah. Tingginya TPT menunjukkan bahwa ada banyak tenaga kerja yang belum terserap oleh pasar sehingga mengindikasikan masalah dalam perekonomian, seperti kurangnya kesempatan kerja, ketidaksesuaian keterampilan dengan kebutuhan industri, atau kondisi ekonomi yang tidak stabil. Berdasarkan penelitian-penelitian

sebelumnya, terdapat beberapa faktor dari berbagai sektor yang diduga memengaruhi TPT, yaitu indeks pembangunan manusia dengan beberapa dimensi pembentuk, termasuk rata-rata lama sekolah dan umur harapan hidup (Darmawan & Mifrahi, 2022). Selain itu, kepadatan penduduk, produk domestik regional bruto, tingkat partisipasi angkatan kerja, dan jumlah penduduk miskin juga diduga memengaruhi TPT (Fahanisa, 2023; Mahmud & Pasaribu, 2021; Sore dkk., 2023).

BAB II

METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder terkait pengangguran yang diperoleh dari laman Badan Pusat Statistik (BPS) Provinsi Sulawesi Selatan tahun 2019-2023 dengan jumlah amatan sebanyak 120. Adapun unit penelitian pada tingkat kabupaten/kota yang terdiri dari 21 kabupaten dan 3 kota. Data untuk beberapa amatan dapat dilihat pada Lampiran 1. Variabel dalam penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. Variabel penelitian

Variabel	Keterangan	Satuan	Definisi
y	Tingkat Pengangguran Terbuka	Persentase	Persentase jumlah pengangguran terhadap jumlah angkatan kerja.
x_1	Indeks Pembangunan Manusia	Indeks	Indikator komposit yang mengukur capaian pembangunan kualitas hidup manusia dalam tiga dimensi dasar: umur panjang dan hidup sehat, pengetahuan, serta standar hidup layak.
x_2	Rata-Rata Lama Sekolah	Tahun	Rata-rata jumlah tahun yang dihabiskan oleh penduduk berusia 15 tahun ke atas untuk menempuh pendidikan
x_3	Umur Harapan Hidup	Tahun	Rata-rata jumlah tahun yang masih akan dijalani oleh seseorang pada usia tertentu dengan asumsi tingkat kematian tetap.
x_4	Kepadatan Penduduk per km^2	Jiwa	Jumlah rata-rata penduduk per kilometer persegi wilayah.
x_5	Produk Domestik Regional Bruto Atas Dasar Harga Berlaku	Milyar Rupiah	Nilai total barang dan jasa yang dihasilkan dalam satu wilayah selama periode tertentu dan dihitung berdasarkan harga pasar tanpa memperhitungkan inflasi.
x_6	Tingkat Partisipasi Angkatan Kerja	Persentase	Persentase jumlah angkatan kerja (penduduk usia kerja yang bekerja atau sedang mencari pekerjaan) terhadap total jumlah penduduk usia kerja.

Variabel	Keterangan	Satuan	Definisi
x_7	Jumlah Penduduk Miskin	Ribu Jiwa	Jumlah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan di bawah garis kemiskinan.

2.2 Analisis Data

Metode analisis yang digunakan adalah metode pemodelan dengan MGTWRR. Analisis ini dilakukan melalui tahapan-tahapan sebagai berikut:

1. Melakukan eksplorasi data pada variabel respon berdasarkan sudut geografis dengan menggunakan peta tematik.
2. Melakukan pendeteksian heterogenitas spasial menggunakan uji BP berdasarkan Persamaan (5) dan heterogenitas temporal menggunakan *box plot*.
3. Melakukan pendeteksian multikolinearitas menggunakan nilai VIF berdasarkan Persamaan (28).
4. Melakukan standarisasi data menggunakan *scaling* dan *centering* untuk variabel prediktor dengan persamaan:

$$x_{ik}^* = \frac{x_{ik} - \bar{x}_k}{S_{x_{ik}}}$$

$$\text{dengan } S_{x_{ik}} = \sqrt{\frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}{n-1}}.$$

5. Menentukan variabel global dan lokal menggunakan interval kepercayaan untuk setiap parameter dari model regresi global.
6. Analisis Model MGTWR
 - 1) Menentukan *bandwidth* optimum dan fungsi pembobot spasial menggunakan nilai CV minimum berdasarkan Persamaan (12).
 - 2) Menentukan parameter rasio spasial temporal (τ) berdasarkan Persamaan (18), parameter spasial (λ^*), dan parameter temporal (μ) dengan meminimumkan nilai CV mengikuti persamaan:

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

- 3) Menentukan jarak lokasi amatan dan waktu menggunakan Persamaan (15), *bandwidth* optimum spasial temporal (h_{st}), dan fungsi pembobot spasial temporal.
- 4) Membentuk matriks pembobot spasial-temporal menggunakan fungsi pembobot spasial temporal.
- 5) Estimasi Parameter model MGTWR menggunakan Persamaan (24) dan Persamaan (25).
7. Menghitung tetapan bias (λ) global dan lokal menggunakan metode estimator LW berdasarkan Persamaan (31) dan Persamaan (32).
8. Melakukan estimasi parameter global dan lokal MGTWRR.

9. Melakukan pengujian signifikansi parameter global dan lokal MGTWRR menggunakan Persamaan (33) dan Persamaan (34).
10. Melakukan pemilihan model terbaik menggunakan nilai AIC berdasarkan Persamaan (35) dan *Adjusted R²* berdasarkan Persamaan (36) antara model MGTWR dan MGTWRR.
11. Melakukan transformasi balik estimasi parameter model ke parameter awal menggunakan persamaan:

$$\beta_k = \frac{\beta_k^*}{s_{x_{ik}}} \text{ dan } \beta_0 = \bar{y} - \sum_{k=1}^p \beta_k \bar{x}_k$$

12. Menginterpretasi model dan membuat kesimpulan.