

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis regresi spasial merupakan pengembangan dari analisis regresi linear yang mempertimbangkan pengaruh spasial (lokasi) pada data. Salah satu pengaruh spasial pada data yaitu terjadinya heterogenitas spasial dengan kondisi data di lokasi yang satu dengan lokasi yang lain tidak sama dari segi geografis (Anselin, 1988). Dampak dari heterogenitas spasial adalah parameter regresi bervariasi secara spasial atau disebut juga terjadi nonstasioneritas spasial pada parameter regresi (Nuryulianingdyah, 2022). Jenis regresi spasial yang digunakan untuk mengatasi data yang mengalami heterogenitas spasial adalah model *Geographically Weighted Regression* (GWR) (Fotheringham dkk., 2003). Model GWR menggunakan faktor geografis sebagai variabel independen yang dapat mempengaruhi variabel dependen (Fotheringham dkk., 2002). Model GWR hanya menggunakan data spasial pada satu periode waktu, padahal data spasial biasanya melibatkan beberapa pengamatan waktu. Oleh karena itu, data spasial seharusnya dianalisis dengan mempertimbangkan pengaruh temporal (waktu) (Sifriyani, Rasjid, dkk., 2022).

Komponen waktu dalam model GWR dapat ditambahkan melalui model *Geographically and Temporally Weighted Regression* (GTWR) (Fotheringham dkk., 2015). Model GTWR menggabungkan informasi spasial dan temporal pada matriks pembobot untuk mengidentifikasi adanya heterogenitas spasial dan temporal (Conita & Purwaningsih, 2017). Namun, model GTWR mengasumsikan bahwa jarak spasial dan temporal tidak saling memengaruhi sehingga interaksi antara keduanya terabaikan. Hal ini dapat menyebabkan terjadinya pemodelan yang kurang tepat (Ania dkk., 2024). Model *Improved Geographically and Temporally Weighted Regression* (IGTWR) mengatasi kelemahan tersebut dengan memodifikasi fungsi jarak spasial-temporal pada GTWR dengan mempertimbangkan interaksi fungsi jarak spasial-temporal (Wu dkk., 2014). Selain itu, Analisis data spasial-temporal tidak dapat mengatasi masalah multikolinearitas sehingga pemodelan menjadi kurang optimal (Arthayanti dkk., 2017).

Multikolinearitas adalah suatu kondisi adanya satu atau lebih variabel independen yang berkorelasi tinggi dengan variabel independen lainnya di setiap amatan (Abdullah dkk., 2017). Menurut Hastie dkk. (2009) regresi *ridge* dapat menangani data yang mengalami masalah multikolinearitas. Regresi *ridge* adalah metode untuk mengendalikan kestabilan penduga kuadrat terkecil yang mengalami multikolinearitas dengan cara menambahkan suatu koefisien tetapan bias bernilai



positif sehingga hasil yang diperoleh memiliki ragam data hasil dengan menggunakan metode kuadrat terkecil (Yulita, 2022). Analisis data spasial-temporal dengan modifikasi fungsi jarak spasial-temporal untuk mengatasi masalah multikolinearitas dapat digunakan IGTWR dengan regresi *ridge* (Arthayanti dkk., 2017) atau *Improved Geographically and Temporally Weighted Ridge Regression* (IGTWR-Ridge).

Penelitian sebelumnya telah menggunakan pendekatan IGTWR untuk menganalisis data spasial-temporal serta regresi *ridge* untuk data yang mengalami masalah multikolinearitas. Misalnya, penelitian oleh Abdullah dkk. (2017) menerapkan *Geographically Weighted Regression* (GWR) dengan regresi *ridge* pada studi kasus angka kematian bayi di Kabupaten Bulukumba. Adapun penelitian oleh Sifriyani dkk. (2022) menggunakan model IGTWR yang diimplementasikan pada kasus COVID-19 di Kalimantan Timur dan menemukan bahwa model IGTWR lebih baik dari model GTWR dan regresi klasik. Selain itu, Ania dkk. (2024) menerapkan IGTWR dengan regresi *robust* menggunakan estimasi-M pada kasus angka kejadian DBD di Sulawesi Selatan. Hasil penelitiannya menunjukkan bahwa model ini lebih efektif dalam memodelkan kasus tersebut karena memiliki nilai *Root Mean Square Error* (RMSE) dan *Mean Absolut Deviation* (MAD) yang lebih kecil serta *Adjusted R-Square* (R^2_{Adj}) yang lebih besar dibandingkan dengan model IGTWR. Penelitian-penelitian terdahulu telah menerapkan model IGTWR serta regresi *ridge* yang berfokus pada data kesehatan sehingga pada penelitian ini akan berfokus pada data sosial-ekonomi yaitu tingkat pengangguran terbuka di Sulawesi Selatan.

Tingkat Pengangguran Terbuka (TPT) merupakan indikator penting yang menunjukkan persentase jumlah pengangguran terhadap jumlah angkatan kerja. TPT yang tinggi mencerminkan ketidakseimbangan antara jumlah angkatan kerja dan ketersediaan lapangan kerja. Berdasarkan publikasi Badan Pusat Statistik (2024), TPT di Sulawesi Selatan pada Agustus 2022 tercatat sebesar 4.51 persen. Namun, angka tersebut meningkat sebesar 0.65 persen sehingga menjadi 5.26 persen pada Februari 2023. Peningkatan ini menunjukkan adanya tekanan di pasar tenaga kerja yang memerlukan perhatian dari pemerintah dalam merancang kebijakan dan program pengendalian angka TPT dengan mengetahui berbagai faktor yang memengaruhinya. Penelitian sebelumnya terkait dengan data TPT menemukan bahwa Indeks pembangunan manusia, jumlah penduduk, produk domestik regional bruto atas dasar harga berlaku, persentase penduduk miskin, umur harapan hidup dan rata-rata lama sekolah berpengaruh secara signifikan terhadap TPT (Sisnita & Prawoto, 2017; Rambe dkk., 2019; Adrian dkk., 2023; Fikriani & Rifai, 2023).

Data TPT serta faktor-faktor yang diduga memengaruhinya diamati di setiap kabupaten/kota di Provinsi Sulawesi Selatan selama tahun 2020-2023. Adanya kemungkinan perbedaan karakteristik perkabupaten/kota selama periode 4 tahun di Sulawesi Selatan sehingga memerlukan pertimbangan pengaruh heterogenitas spasial dan temporal (Amin, 2024). Selain itu, data faktor-faktor yang diduga memengaruhi TPT di Sulawesi Selatan selama periode tersebut memungkinkan adanya masalah multikolinearitas. Berdasarkan uraian diatas, peneliti akan mengkaji



el regresi spasial-temporal yang mempertimbangkan interaksi
-temporal serta mengatasi masalah multikolinearitas yang
ta TPT di Sulawesi Selatan tahun 2020-2023. Oleh karena itu,
il “**Pemodelan *Improved Geographically and Temporally*
gression (Studi Kasus: Tingkat Pengangguran Terbuka di**

1.2 Tujuan dan Manfaat

Tujuan dalam penelitian ini adalah sebagai berikut:

1. Melakukan estimasi parameter model *improved geographically and temporally weighted ridge regression*.
2. Memodelkan *improved geographically and temporally weighted ridge regression* pada data tingkat pengangguran terbuka untuk setiap kabupaten/kota di Provinsi Sulawesi Selatan tahun 2020-2023.
3. Menentukan faktor-faktor yang berpengaruh signifikan terhadap tingkat pengangguran terbuka untuk setiap kabupaten/kota di Provinsi Sulawesi Selatan tahun 2020-2023.

Manfaat dalam penelitian ini adalah sebagai berikut:

1. Sebagai bentuk pengembangan ilmu pengetahuan di bidang statistika tentang analisis *improved geographically and temporally weighted ridge regression* untuk mengatasi data yang mengalami masalah multikolinearitas.
2. Sebagai bantuan kepada Pemerintah Provinsi Sulawesi Selatan dan sektor terkait lainnya untuk dapat menggunakan informasi hasil penelitian sebagai bahan pertimbangan dalam perencanaan program pengendalian faktor-faktor yang memengaruhi tingkat pengangguran terbuka di setiap kabupaten/kota di Sulawesi Selatan.

1.3 Batasan Masalah

Batasan-batasan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah tingkat pengangguran terbuka di Sulawesi Selatan tahun 2020-2023.
2. Penentuan jarak antarlokasi pengamatan menggunakan jarak *euclidean* dan pembentukan fungsi pembobot menggunakan pembobot *fixed* kernel dengan CV optimum.
3. Penentuan model terbaik berdasarkan nilai *Akaike Information Criterion* (AIC), *Root Mean Square Error* (RMSE), dan *Adjusted R-Square* (R^2_{Adj}).

1.4 Landasan Teori

1.4.1 Analisis Regresi

Analisis regresi berkaitan dengan studi mengenai ketergantungan satu variabel, yaitu variabel dependen terhadap satu atau lebih variabel independen (Lutfiani dkk., 2019). Analisis hubungan antara satu variabel independen dan satu variabel dependen disebut analisis regresi linier sederhana, sedangkan analisis hubungan antara beberapa variabel independen dengan satu variabel dependen disebut

berganda (Rawlings dkk., 1998). Bentuk model regresi dengan sebanyak p dan jumlah pengamatan n sesuai Persamaan (1).

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



Y_i : Nilai dari variabel dependen pada pengamatan ke- i
 X_{ik} : Nilai dari variabel independen ke- k pada pengamatan ke- i ($i = 1, 2, \dots, n$)
 β_0 : Nilai intersep model regresi
 β_k : Nilai parameter regresi variabel independen ke- k ($k = 1, 2, \dots, p$)
 ε_i : Residual pengamatan ke- i .

dalam notasi matriks, dapat ditulis sesuai Persamaan (2).

$$Y = X\beta + \varepsilon \quad (2)$$

dengan,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

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

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

1.4.2 Heterogenitas Spasial

Heterogenitas spasial disebabkan karena adanya perbedaan karakteristik setiap lokasi. Pengujian heterogenitas spasial dapat dilakukan dengan menggunakan statistik uji Breusch-Pagan (BP) (Anselin, 1988). Dikatakan data memiliki heterogenitas spasial jika nilai $BP > \chi^2_{\alpha,p}$ atau $p\text{-value} < \alpha$. Adapun hipotesis dalam pengujian ini adalah:

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

$H_1 : \exists i, \sigma_i^2 \neq \sigma^2$ (terdapat heterogenitas spasial)

Statistik uji BP menggunakan Persamaan (3).

$$BP = \left(\frac{1}{2}\right) \mathbf{f}^t \mathbf{Z} (\mathbf{Z}^t \mathbf{Z})^{-1} \mathbf{Z}^t \mathbf{f} \sim \chi^2_{(\alpha,p)} \quad (3)$$

dengan,



rgan $f_i = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1\right)$
 atan ke- i

1), vektor yang sudah dinormalisasi untuk semua pengamatan.

1.4.3 Pembobot Spasial

Pemodelan pada *Geographically Weighted Regression* (GWR) sangat bergantung pada pembobot spasial yang digunakan. Menurut Fotheringham dkk. (2002), pembobot terbentuk dari fungsi kernel yang digunakan untuk mengestimasi parameter pada model GWR. Fungsi kernel umumnya mendefinisikan bobot sedemikian rupa sehingga data amatan yang letaknya dekat ke titik- i memiliki pengaruh yang lebih besar dari titik- i . Pembobot tersebut dilambangkan sebagai W yang merupakan diagonal matriks dari w_j (bobot tiap titik data dengan lokasi ke- j). Fungsi kernel terbagi menjadi beberapa fungsi sebagai berikut:

1. Fungsi Kernel *Gaussian*

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right) \quad (4)$$

2. Fungsi Kernel *Bisquare*

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^2\right)^2, & \text{untuk } d_{ij} < h \\ 0, & \text{untuk } d_{ij} \geq h \end{cases} \quad (5)$$

3. Fungsi Kernel *Exponential*

$$w_{ij}(u_i, v_j) = \exp\left(-\left(\frac{d_{ij}}{h}\right)\right) \quad (6)$$

4. Fungsi Kernel *Tricube*

$$w_j(u_i, v_i) = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^3\right)^3, & \text{untuk } d_{ij} < h \\ 0, & \text{untuk } d_{ij} \geq h \end{cases} \quad (7)$$

dengan d_{ij} adalah jarak *euclidean* antara lokasi pengamatan ke- i dan titik lokasi pengamatan ke- j , sedangkan h adalah *bandwidth* (Fotheringham dkk., 2002). Perhitungan d_{ij} dapat dilakukan menggunakan Persamaan (8).

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (8)$$

dengan,

u_i : *Latitude* pada lokasi pengamatan ke- i

u_j : *Latitude* pada lokasi pengamatan ke- j

v_i : *Longitude* pada lokasi pengamatan ke- i

v_j : *Longitude* pada lokasi pengamatan ke- j



Menurut Fotheringham dkk. (2002), *bandwidth* (h) merupakan ukuran jarak yang dapat mengukur sejauh mana pengaruh suatu lokasi dengan semakin besar nilai h menunjukkan bobot yang terbentuk relatif lebih tinggi yang lebih terpencil, sedangkan semakin kecil nilai h menunjukkan bobot yang terbentuk relatif lebih rendah untuk daerah yang lebih dekat. Pemilihan *bandwidth* yang optimal dapat dilakukan dengan melihat nilai *Cross Validation* (CV) minimum. Perhitungan CV menggunakan Persamaan (9).

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(h))^2 \quad (9)$$

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

1.4.4 Geographically Weighted Regression

Model *Geographically Weighted Regression* (GWR) merupakan salah satu model regresi yang dikembangkan dari model regresi global dengan melibatkan faktor geografis atau parameternya dihitung pada setiap lokasi pengamatan sehingga setiap lokasi pengamatan mempunyai nilai parameter yang berbeda-beda. Variabel dependen dalam model GWR diprediksi dengan variabel independen yang masing-masing koefisien regresinya tergantung pada lokasi data yang diamati. Model GWR adalah salah satu model yang dimunculkan dari metode pendekatan titik yaitu pendekatan berdasarkan posisi koordinat garis lintang (*latitude*) dan garis bujur (*longitude*) (Wahyudi dkk., 2023). Persamaan umum untuk model GWR menggunakan Persamaan (10).

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (10)$$

dengan,

- y_i : Nilai dari variabel dependen pada lokasi pengamatan ke- i
- x_{ik} : Nilai dari variabel independen ke- k pada lokasi pengamatan ke- i dengan $i = 1, 2, \dots, n$
- (u_i, v_i) : Titik koordinat (*latitude*, *longitude*) lokasi pengamatan ke- i
- $\beta_0(u_i, v_i)$: Nilai intersep model regresi pada lokasi pengamatan ke- i
- $\beta_k(u_i, v_i)$: Nilai parameter regresi variabel independen ke- k dengan $k = 1, 2, \dots, p$ untuk setiap lokasi pengamatan ke- i
- ε_i : Residual pengamatan ke- i yang diasumsikan identik, bebas, dan berdistribusi normal dengan rata-rata nol dan varian konstan σ^2 .

1.4.5 Improved Geographically and Temporally Weighted Regression

Model *Geographically and temporally weighted regression* (GTWR) merupakan pengembangan dari model *Geographically Weighted Regression* (GWR) dengan menambahkan unsur waktu (*temporal*). Model GTWR mengakomodasi adanya heterogenitas spasial dan temporal (Musdalifah dkk., 2023). Persamaan umum untuk model GTWR menggunakan Persamaan (11).

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (11)$$



- y_i : Nilai variabel dependen untuk pengamatan ke- i
- x_{ik} : Nilai variabel independen ke- k pada lokasi pengamatan (u_i, v_i) dan waktu pengamatan ke- t
- (u_i, v_i, t_i) : Koordinat lokasi pengamatan ke- i dan waktu pengamatan ke- t
- $\beta_0(u_i, v_i, t_i)$: Nilai intersep model regresi pada lokasi pengamatan ke- i dan waktu pengamatan ke- t
- $\beta_k(u_i, v_i, t_i)$: Nilai parameter regresi variabel independen ke- k pada lokasi ke- (u_i, v_i) dan waktu pengamatan ke- t dengan $k = 1, 2, \dots, p$

x_{ik} : Nilai dari variabel independen ke- k pada pengamatan ke- i dengan
 $k = 1, 2, \dots, p$

ε_i : Residual di pengamatan ke- i

Estimasi parameter dapat ditulis sesuai Persamaan (12).

$$\hat{\beta}(u_i, v_i, t_i) = (X^t W(u_i, v_i, t_i) X)^{-1} X^t W(u_i, v_i, t_i) Y \quad (12)$$

dengan $W(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in})$ adalah matriks pada lokasi pengamatan (u_i, v_i) dan waktu t_i . Elemen diagonal $w_{ij} (1 \leq j \leq n)$ adalah fungsi jarak spasial-temporal pada titik pengamatan (u_i, v_i, t_i) . Jarak spasial-temporal yang dibangun merupakan kombinasi dari fungsi jarak spasial dan fungsi jarak temporal (Harianto dkk., 2021). Fungsi jarak spasial-temporal (d_{ij}^{ST}) merupakan gabungan dari fungsi jarak spasial (d_{ij}^S) dan fungsi jarak temporal (d_{ij}^T), yang dituliskan sesuai Persamaan (13).

$$\begin{aligned} (d_{ij}^S)^2 &= (u_i - u_j)^2 + (v_i - v_j)^2 \\ (d_{ij}^T)^2 &= (t_i - t_j)^2 \\ (d_{ij}^{ST})^2 &= \lambda [(u_i - u_j)^2 + (v_i - v_j)^2] + \mu [(t_i - t_j)^2] \end{aligned} \quad (13)$$

Parameter jarak spasial (λ) dan parameter jarak temporal (μ) digunakan sebagai parameter penyeimbang antara pengaruh lokasi dan waktu pada pengukuran jarak spasial-temporal. Jika $\mu = 0$, maka hanya nonstasioneritas spasial dan jarak spasial yang dipertimbangkan dan model GTWR disederhanakan menjadi model GWR. Jika $\lambda = 0$, maka hanya nonstasioneritas waktu dan jarak waktu yang dipertimbangkan dan model GTWR disederhanakan menjadi model TWR (Ma dkk., 2021). Pada kebanyakan kasus, nilai λ dan μ tidak sama dengan nol. Jika hal tersebut terjadi, maka nonstasioneritas spasial-temporal dan jarak spasial-temporal yang dipertimbangkan sehingga digunakan model GTWR. Berdasarkan Persamaan (13) diperoleh:

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

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

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



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

h_S adalah *bandwidth* spasial, h_T adalah *bandwidth* temporal, dan h_{ST} adalah *bandwidth* spasial-temporal.

Misalkan τ adalah parameter rasio dari $\tau = \frac{\mu}{\lambda}$ dengan $\lambda \neq 0$, maka diperoleh Persamaan (14) (Huang dkk., 2010):

$$\frac{(d_{ij}^{ST})^2}{\lambda} = [(u_i - u_j)^2 + (v_i - v_j)^2] + \tau [(t_i - t_j)^2] \quad (14)$$

Misalnya $\lambda = 1$ yang bertujuan untuk mengurangi parameter yang tidak diketahui. Ada satu parameter yang tidak diketahui yaitu τ . Parameter τ berfungsi untuk menyeimbangkan pengaruh jarak temporal terhadap jarak spasial (Erda dkk., 2019). Parameter ini didapatkan dari metode *Cross Validation* (CV) minimum dengan menginisialisasi nilai τ awal yang dituliskan sesuai Persamaan (15).

$$CV(\tau) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(\tau))^2 \quad (15)$$

dengan $\hat{y}_{\neq i}(\tau)$ adalah nilai prediksi y_i pada pengamatan ke- i yang dihitung dari proses estimasi tanpa melibatkan data ke- i . Selanjutnya estimasi parameter λ dan μ bisa diperoleh dengan metode iteratif berdasarkan hasil estimasi τ yang dihasilkan CV minimum.

Matriks pembobot yang digunakan pada model GTWR menggunakan operator sederhana yaitu penjumlahan '+', operator tersebut digunakan untuk mengukur jarak spasial-temporal dengan kombinasi linier antara jarak spasial dan jarak temporal yang ditunjukkan pada Persamaan (16) (Huang dkk., 2010).

$$d_{ST}^2 = d_S^2 \otimes d_T^2 = \lambda d_S^2 + \mu d_T^2 \quad (16)$$

dengan parameter jarak spasial (λ^*) dan parameter jarak temporal (μ) disesuaikan untuk menyeimbangkan skala yang digunakan dalam mengukur jarak spasial dan temporal pada masing-masing sistem koordinat. Ini setara dengan membentuk matriks bobot spasial (W^S) dan matriks bobot temporal (W^T) secara terpisah kemudian menggabungkannya menjadi matriks bobot spasial-temporal ($W^{ST} = W^S \times W^T$) (Huang dkk., 2010). Namun, spesifikasi ini mengasumsikan bahwa sistem koordinat spasial-temporal bersifat tegak lurus sehingga jarak spasial tidak memengaruhi jarak temporal yang menjadikannya kurang tepat untuk pemodelan interaksi spasial-temporal (Ania dkk., 2024). Pengembangan model GTWR oleh Wu

gunakan operator ' $\otimes$ ' yang didefinisikan dengan cara yang relatif untuk model *Improved Geographically and Temporally Weighted* dengan jarak spasial-temporal sesuai Persamaan (17).

$$d_{ij}^S \otimes d_{ij}^T = \begin{cases} \lambda d_{ij}^S + \mu d_{ij}^T + 2 \sqrt{\lambda \mu d_{ij}^S d_{ij}^T} \cos(\xi) & , t_j \leq t_i \\ \infty & , t_j > t_i \end{cases} \quad (17)$$



dengan t_i dan t_j adalah waktu pengamatan pada lokasi ke- i dan ke- j . Parameter λ , μ , dan $\xi \in [0, \pi]$ adalah parameter penyeimbang yang diperoleh dengan mengoptimalkan koefisien determinasi melalui prosedur CV. Parameter ξ digunakan untuk mengukur interaksi pengaruh lokasi dan waktu. Jika $\xi = 0$, maka interaksi antara lokasi dan waktu maksimal; jika $\xi = \frac{\pi}{2}$, maka tidak ada interaksi (kembali seperti GTWR) (Wu dkk., 2014). Estimasi parameter model GTWR dan model IGTWR sama, yang membedakan yaitu hanya pada matriks pembobotnya. Adapun estimasi parameter model IGTWR dapat ditulis sesuai Persamaan (18).

$$\hat{\beta}(u_i, v_i, t_i) = (X^t W_{IGTWR}(u_i, v_i, t_i) X)^{-1} X^t W_{IGTWR}(u_i, v_i, t_i) Y \quad (18)$$

dengan $W_{IGTWR}(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in})$ adalah matriks pada lokasi pengamatan (u_i, v_i) dan waktu t_i . Elemen diagonal $w_{ij} (1 \leq j \leq n)$ adalah fungsi jarak spasial-temporal pada titik pengamatan (u_i, v_i, t_i) . Spesifikasi pada model IGTWR memperhatikan kausalitas waktu dengan hanya mempertimbangkan pengaruh dari pengamatan yang terjadi pada waktu yang sama atau sebelumnya. Oleh karena itu, jika waktu pengamatan $t_j > t_i$ maka nilai bobot tersebut menjadi nol.

1.4.6 Pendeteksian Multikolinearitas

Pendeteksian multikolinieritas digunakan untuk mengidentifikasi adanya hubungan yang signifikan antarvariabel independen (Syukron & Fahri, 2018). Model regresi yang baik seharusnya bebas multikolinearitas. Nilai *Variance Inflation Factor* (VIF) dapat digunakan sebagai kriteria untuk mendeteksi kasus multikolinieritas (Amalah dkk., 2023). Nilai VIF dapat diperoleh menggunakan Persamaan (19).

$$VIF = \frac{1}{1 - R_k^2} \quad (19)$$

dengan R_k^2 merupakan koefisien determinasi antara x_k dengan variabel independen lainnya. Jika $VIF < 10$, maka dapat diindikasikan bahwa tidak terdapat kasus multikolinearitas. Sebaliknya, jika $VIF > 10$, maka dapat diindikasikan bahwa terdapat kasus multikolinieritas (Amalah dkk., 2023).

1.4.7 Regresi Ridge

Regresi *ridge* adalah metode untuk mengendalikan kestabilan penduga kuadrat terkecil yang mengalami multikolinieritas (Hoerl & Kennard, 2000). Metode regresi *ridge* dilakukan dengan cara menambahkan koefisien tetapan bias positif (λ_R) pada proses estimasi parameter sehingga hasil yang diperoleh memiliki ragam yang lebih kecil daripada hasil dengan menggunakan metode kuadrat terkecil. Metode regresi *ridge* ini didasarkan pada modifikasi metode kuadrat terkecil, yakni dengan



$\lambda_R I$ pada $(X^t X)$ sebelum diinverskan. Estimasi parameter dilakukan pada Persamaan (20).

$$+ \lambda_R I) \hat{\beta}_R = X^t Y \text{ atau } \hat{\beta}_R = (X^t X + \lambda_R I)^{-1} X^t Y \quad (20)$$

di mana λ_R adalah suatu konstanta (parameter bias) yang dipilih sedemikian kecil. Jika $\lambda_R = 0$, maka estimator *ridge* sama dengan estimator

Penambahan konstanta λ_R mengakibatkan nilai-nilai elemen diagonal matriks $(X^t X + \lambda_R I)^{-1}$ menjadi kecil sehingga rata-rata kuadrat residualnya menjadi kecil. Hal ini menunjukkan bahwa taksiran koefisien regresi menjadi stabil. Pemilihan nilai konstan λ_R dapat digunakan metode Estimator Lawless & Wang (LW) yang diperoleh dengan Persamaan (21).

$$\lambda_R(u_i, v_i, t_i) = \frac{p \times (MSE)}{\hat{\beta}_{WLS_i}^t (X^t X) \hat{\beta}_{WLS_i}} \quad (21)$$

dengan p merupakan jumlah parameter $\hat{\beta}$ tanpa $\hat{\beta}_0$, MSE merupakan *Mean Square Error*, $\hat{\beta}_{WLS}$ merupakan estimasi yang diperoleh dengan *Weighted Least Square* (WLS) (Lawless & Wang, 1976).

1.4.8 Tingkat Pengangguran Terbuka

Pengangguran merupakan fenomena umum di negara berkembang, termasuk Indonesia. Menurut Badan Pusat Statistik (BPS), pengangguran mencakup individu yang sedang mencari pekerjaan, mempersiapkan usaha, merasa tidak mungkin mendapatkan pekerjaan atau sudah memiliki pekerjaan namun belum mulai bekerja (BPS, 2025). Umumnya, pengangguran terjadi akibat pertumbuhan angkatan kerja yang lebih cepat dibandingkan dengan pertumbuhan lapangan pekerjaan yang tersedia. Menurut Putra & Faridatussalam (2023), Tingkat Pengangguran Terbuka (TPT) menggambarkan kondisi ketika tenaga kerja tidak terserap secara maksimal yang menyebabkan penduduk tanpa keahlian khusus lebih mudah tersingkir dari pasar kerja. Faktor lain yang memperburuk kondisi ini adalah rendahnya motivasi individu untuk mencari pekerjaan dan mengembangkan keterampilan.

Tingginya angka TPT menjadi tantangan serius dalam perekonomian karena dapat menurunkan produktivitas nasional, menghambat pertumbuhan ekonomi, dan menurunkan pendapatan masyarakat. Secara sosial, pengangguran berdampak pada meningkatnya angka kemiskinan, kriminalitas, serta gangguan kesehatan mental. Selain itu, pengangguran jangka panjang dapat menyebabkan individu kehilangan keterampilan sehingga semakin sulit untuk kembali bekerja. Oleh karena itu, menurunkan angka TPT merupakan prioritas dalam pembangunan ekonomi nasional. Upaya tersebut perlu didukung dengan pemahaman mendalam mengenai faktor-faktor yang memengaruhi TPT. Analisis yang tepat dapat membantu merumuskan kebijakan yang efektif dalam mengurangi pengangguran dan meningkatkan kesejahteraan masyarakat.

Indeks Pembangunan Manusia (IPM) menjadi salah satu indikator penting dalam menilai kualitas hidup dan kesejahteraan masyarakat di suatu wilayah.

awoto (2017), semakin tinggi IPM suatu daerah, maka semakin hidup masyarakatnya. Tingginya IPM juga berkaitan dengan angka pengangguran karena masyarakat memiliki akses yang baik dan peluang kerja yang lebih luas. Dengan kata lain, IPM an sejauh mana masyarakat mampu mengembangkan dirinya



Salah satu faktor yang erat kaitannya dengan IPM adalah persentase penduduk miskin. Menurut Kevin dkk. (2022), kemiskinan tidak hanya terkait dengan jumlah penduduk miskin, tetapi juga persentase penduduk miskin di suatu wilayah tertentu. Perbedaan tingkat kemajuan pembangunan antardaerah menyebabkan ketimpangan dalam angka kemiskinan. Daerah yang mengalami pertumbuhan ekonomi yang lambat cenderung memiliki angka kemiskinan lebih tinggi. Hal ini menunjukkan bahwa pemerataan pembangunan menjadi sangat penting agar masyarakat di seluruh wilayah dapat merasakan manfaat ekonomi secara adil.

Pertumbuhan ekonomi itu sendiri dapat dilihat melalui Produk Domestik Regional Bruto Atas Dasar Harga Berlaku (PDRB ADHB) yang menggambarkan kemampuan suatu wilayah dalam mengelola sumber daya alam dan faktor produksi lainnya. Rambe dkk. (2019) menyatakan bahwa besarnya PDRB ADHB sangat dipengaruhi oleh potensi ekonomi dan kapasitas produksi daerah tersebut. PDRB ADHB yang tinggi mencerminkan tingginya nilai barang dan jasa yang dihasilkan serta potensi peningkatan pendapatan masyarakat. Namun, pertumbuhan PDRB ADHB belum tentu mencerminkan kesejahteraan merata jika tidak diiringi distribusi pendapatan yang adil.

Kesehatan juga merupakan faktor fundamental dalam pembangunan manusia. Umur Harapan hidup (UHH) menjadi indikator utama dalam mengukur keberhasilan pembangunan di bidang kesehatan. Kevin dkk. (2022) menyebutkan bahwa semakin tinggi UHH, maka semakin baik kualitas pelayanan kesehatan dan kesejahteraan masyarakat di wilayah tersebut. Sebaliknya, rendahnya UHH mengindikasikan bahwa pembangunan di sektor kesehatan masih tertinggal dan belum merata. Kesehatan yang baik mendukung produktivitas masyarakat dan memperkuat daya saing tenaga kerja.

Selain kesehatan, pendidikan juga menjadi aspek penting yang memengaruhi angka pengangguran. Rata-rata Lama Sekolah (RLS) menggambarkan lamanya pendidikan formal yang ditempuh oleh penduduk usia 25 tahun ke atas. Siskawati dkk. (2021) menjelaskan bahwa semakin tinggi tingkat pendidikan seseorang, maka semakin terbuka pula peluangnya untuk mendapatkan pekerjaan yang layak. Orang yang berpendidikan tinggi cenderung memiliki preferensi dan fleksibilitas dalam memilih pekerjaan. Namun, mereka juga lebih selektif dan tidak jarang lebih memilih menganggur sementara jika tidak menemukan pekerjaan yang sesuai dengan harapannya.

Faktor terakhir yang tidak kalah penting adalah jumlah penduduk. Sisnita & Prawoto (2017) mengungkapkan bahwa jumlah penduduk yang besar tidak selalu menjadi kekuatan pembangunan apabila tidak diimbangi dengan penyediaan



semua penduduk memiliki kemampuan untuk menghasilkan tenaga kerja menjadi sangat penting. Penduduk yang bekerja akan dapat berkontribusi terhadap pembangunan ekonomi. Oleh karena itu, peningkatan jumlah penduduk dan peningkatan kualitasnya melalui kesehatan sangat diperlukan untuk mendukung pembangunan berkelanjutan.

BAB II

METODOLOGI PENELITIAN

2.1 Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari *website* resmi Badan Pusat Statistika (BPS) Provinsi Sulawesi Selatan pada situs <https://sulsel.bps.go.id/> yang diakses pada tahun 2024. Data yang digunakan sebanyak 24 kabupaten/kota di Provinsi Sulawesi Selatan tahun 2020-2023 yaitu selama 4 tahun periode sehingga pada penelitian ini ada 96 pengamatan seperti pada Lampiran 1. Pada penelitian ini terdapat variabel dependen (Y) dan variabel independen (X) yang ditampilkan pada Tabel 1.

Tabel 1. Variabel penelitian

Variabel	Keterangan	Deskripsi	Satuan
Y	Tingkat Pengangguran Terbuka (TPT)	Persentase jumlah pengangguran terhadap jumlah angkatan kerja. Angkatan kerja adalah penduduk usia 15-65 tahun.	Persen
X_1	Indeks Pembangunan Manusia (IPM)	Indikator komposit untuk mengukur capaian pembangunan kualitas hidup manusia.	Indeks
X_2	Persentase Penduduk Miskin (PPM)	Rasio jumlah penduduk miskin yang berada di bawah garis kemiskinan nasional terhadap jumlah penduduk di setiap kabupaten/kota.	Persen
X_3	Rata-rata Lama Sekolah (RLS)	Rata-rata jumlah tahun belajar penduduk usia 15 tahun keatas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang)	Tahun
X_4	Produk Domestik Regional Bruto Atas Dasar Harga Berlaku (PDARB ADHB)	Nilai total produksi barang dan jasa akhir yang dihasilkan dalam suatu wilayah geografis tertentu selama satu periode tertentu, biasanya dalam satu tahun, diukur dengan harga-harga yang berlaku pada saat produksi tersebut terjadi.	Miliar Rupiah
	Harapan Hidup (UHH)	Rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang selama hidup	Tahun



Variabel	Keterangan	Deskripsi	Satuan
X_6	Jumlah Penduduk (JP)	Jumlah orang yang berdomisili di suatu wilayah atau daerah pada waktu tertentu	Ribu Jiwa

2.2 Metode Analisis

Metode analisis yang digunakan dalam penelitian ini adalah *Improved Geographically and Temporally Weighted Ridge Regression* (IGTWR-Ridge). Pengolahan data dilakukan dengan bantuan *software* Rstudio. Adapun tahapan-tahapan analisis data yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan analisis statistik deskriptif untuk mengeksplorasi sebaran data tingkat pengangguran terbuka di Provinsi Sulawesi Selatan pada tahun 2020-2023.
2. Mengidentifikasi heterogenitas spasial menggunakan uji Breusch-Pagan (BP) berdasarkan Persamaan (3) dan heterogenitas temporal menggunakan *boxplot*.
3. Mengidentifikasi multikolinearitas dengan melihat nilai *Variance Inflation Factor* (VIF) berdasarkan Persamaan (19).
4. Melakukan standarisasi data menggunakan *scaling* dan *centering* pada variabel independen 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. Melakukan pemodelan *Improved Geographically and Temporally Weighted Regression* (IGTWR) sebagai berikut:
 - a) Menghitung *bandwidth* spasial optimum (h_s) menggunakan *Cross Validation* (CV) berdasarkan pendekatan *Geographically Weighted Regression* (GWR).
 - b) Menghitung parameter rasio spasial-temporal optimum (τ) menggunakan CV berdasarkan pendekatan optimasi IGTWR sesuai rumus pada Persamaan (15).
 - c) Menghitung parameter λ dan μ dengan menggunakan pendekatan CV berdasarkan pada fungsi jarak spasial-temporal pada IGTWR sesuai rumus pada Persamaan (17) dan memisalkan nilai $\xi = 0$.
 - d) Menghitung parameter ξ menggunakan pendekatan CV.
 - e) Menghitung *bandwidth* spasial-temporal optimum (h_{sT}) menggunakan CV berdasarkan pendekatan optimasi IGTWR.
 - f) Membentuk matriks pembobot spasial (W_{IGTWR}) menggunakan ukuran jarak spasial dengan interaksi untuk setiap lokasi amatan berdasarkan *exponential*. Estimasi parameter IGTWR menggunakan Persamaan (18). Pemodelan *Improved Geographically and Temporally Weighted Ridge Regression* (IGTWR-Ridge) sebagai berikut:



- a) Menghitung tetapan bias (λ_R) berdasarkan metode estimator Lawless & Wang (LW) menggunakan Persamaan (21) berdasarkan nilai jumlah parameter $\hat{\beta}$ tanpa $\hat{\beta}_0$ serta MSE yang diperoleh dari pemodelan IGTWR.
- b) Melakukan estimasi parameter IGTWR-Ridge menggunakan persamaan:
- $$\hat{\beta}(u_i, v_i, t_i) = [X^t W_{IGTWR}(u_i, v_i, t_i) X + \lambda_R(u_i, v_i, t_i) I]^{-1} X^t W_{IGTWR}(u_i, v_i, t_i) Y$$
- c) Melakukan pemodelan IGTWR-Ridge untuk masing-masing kabupaten/kota menggunakan persamaan:

$$\hat{Y} = x_i^t \hat{\beta}(u_i, v_i, t_i)$$

dengan $x_i^t = (1, x_{i1}, x_{i2}, \dots, x_{ip})$ adalah elemen baris ke- i dari matriks X_i .

7. Melakukan pengujian signifikansi parameter IGTWR-Ridge menggunakan persamaan:

$$t_k = \frac{\hat{\beta}_k(u_i, v_i, t_i)}{\hat{\sigma} \sqrt{C_{kk}}}$$

dengan C_{kk} adalah elemen diagonal ke- k dari matriks CC^t , $C = [X^t W_{IGTWR}(u_i, v_i, t_i) X + \lambda_R(u_i, v_i, t_i) I]^{-1} X^t W_{IGTWR}(u_i, v_i, t_i)$, $\hat{\sigma}^2 = \frac{Y^t(I-L)^t(I-L)Y}{tr((I-L)^t(I-L))}$.

8. Melakukan uji kebaikan model IGTWR-Ridge sebagai berikut:

- a) Menghitung matriks $L_{IGTWR-Ridge}$ menggunakan persamaan:

$$L = \begin{bmatrix} x_1^t [X^t W_{IGTWR}(u_1, v_1, t_1) X + \lambda_R(u_1, v_1, t_1) I]^{-1} X^t W_{IGTWR}(u_1, v_1, t_1) \\ x_2^t [X^t W_{IGTWR}(u_2, v_2, t_2) X + \lambda_R(u_2, v_2, t_2) I]^{-1} X^t W_{IGTWR}(u_2, v_2, t_2) \\ \vdots \\ x_n^t [X^t W_{IGTWR}(u_n, v_n, t_n) X + \lambda_R(u_n, v_n, t_n) I]^{-1} X^t W_{IGTWR}(u_n, v_n, t_n) \end{bmatrix}$$

- b) Menghitung Jumlah Kuadrat Galat (JKG) menggunakan persamaan:

$$JKG_{IGTWR-Ridge} = y^t (I - L_{IGTWR-Ridge})^t (I - L_{IGTWR-Ridge}) y$$

- c) Menghitung Jumlah Kuadrat Total (JKT) menggunakan persamaan:

$$JKT_{IGTWR-Ridge} = y^t y - \frac{1}{n} (y^t J y)$$

dengan J matriks simetris dengan elemen-elemennya bernilai 1.

- d) Menghitung nilai Akaike Information Criterion (AIC), Root Mean Square Error (RMSE), dan Adjusted R-Square (R_{Adj}^2) model IGTWR-Ridge menggunakan persamaan:

$$AIC = -2 \log l(\hat{\beta}) + 2p$$

$$RMSE = \sqrt{\frac{JKG_{IGTWR-Ridge}}{n}}$$

$$R_{Adj}^2 = 1 - \frac{JKG_{IGTWR-Ridge}}{JKT_{IGTWR-Ridge}}$$

yaitu nilai *likelihood* dan p yaitu banyaknya variabel independen informasi balik estimasi parameter model ke bentuk awal persamaan:

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

dan penarikan kesimpulan.

