

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi spasial merupakan pendekatan statistika yang digunakan untuk data yang dipengaruhi oleh faktor lokasi atau wilayah. Pemodelan pada data spasial perlu memperhatikan efek spasial, yaitu heterogenitas spasial dan dependensi spasial (Djuraidah, 2020). Heterogenitas spasial menunjukkan adanya keragaman antar lokasi sedangkan dependensi spasial menunjukkan adanya keterkaitan antar lokasi objek penelitian (Agustina dkk., 2022). Jika terdapat efek heterogenitas maka diselesaikan menggunakan pendekatan titik, tetapi jika terdapat efek dependensi spasial maka diselesaikan menggunakan pendekatan area (Pratiwi dkk., 2018). Metode statistika yang dapat digunakan apabila ditemukan adanya efek heterogenitas spasial adalah *Geographically Weighted Regression* (GWR).

Metode GWR adalah salah satu jenis regresi spasial yang menggunakan pembobotan berdasarkan lokasi geografis, dengan setiap wilayah diwakili oleh titik koordinat geografisnya masing-masing (Oktarina dkk., 2024). Namun, permasalahan yang dihadapi oleh peneliti tidak hanya adanya efek heterogenitas spasial, tetapi juga pengamatan yang dipengaruhi oleh waktu dalam beberapa lokasi berdekatan atau disebut juga dengan efek spasio-temporal. Hal ini menunjukkan adanya variasi yang terjadi dalam dimensi ruang dan waktu secara bersamaan (Liu dkk., 2017). Oleh karena itu, dikembangkanlah metode untuk mengatasi masalah tersebut yakni *Geographically and Temporally Weighted Regression* (GTWR).

Metode GTWR adalah pengembangan dari metode GWR yang memperhitungkan faktor lokasi dan waktu. Keunggulan dari metode GTWR adalah mampu menghasilkan model yang bersifat lokal di setiap lokasi dan waktu (Huang dkk., 2010). Namun, pada pemodelan regresi spasial seperti GWR dan GTWR terkadang ditemukan permasalahan, salah satunya adalah terjadinya multikolinieritas lokal. Hal ini terjadi saat variabel prediktor saling berkorelasi di lokasi tertentu, yang dapat menimbulkan masalah dalam perhitungan koefisien regresi. Oleh karena itu, untuk mengatasi adanya multikolinieritas lokal dapat menggunakan beberapa metode, salah satunya dengan penambahan konsep regresi ridge. Metode ini digunakan untuk mengatasi permasalahan multikolinearitas dengan menghasilkan penaksir bias dari koefisien regresi (Azzahra dkk., 2020). Konstanta bias dari regresi ridge menghasilkan bias yang relatif kecil dan memberikan koefisien regresi yang cenderung stabil. Namun, penggunaan satu konstanta bias diberlakukan untuk seluruh wilayah pengamatan. Padahal tidak semua wilayah yang diamati mengalami masalah multikolinieritas.

Gollini dkk (2015) memperkenalkan model *Locally Compensated Ridge-Geographically Weighted Regression* (LCR-GWR) dengan menggunakan satu koefisien bias untuk suatu wilayah tertentu. Hal ini berarti, jika terdapat N wilayah pengamatan maka terdapat n koefisien bias ridge yang berbeda. Metode ini menghasilkan bias ridge secara lokal. Parameter ridge bervariasi di setiap wilayah untuk menyesuaikan dengan dampak kolinearitas antar variabel prediktor, sehingga diharapkan hasil estimasi koefisien dalam model menjadi lebih tepat (Herawati dkk., 2024). Namun, model tersebut belum dapat mengakomodir efek spasio-temporal, sehingga perlu dilakukan pengembangan lebih lanjut yakni *Locally Compensated Ridge-Geographically and Temporally Weighted Regression* (LCR-GTWR). Metode LCR-GTWR dikembangkan untuk analisis data dengan mempertimbangkan variasi lokasi dan waktu serta mengatasi multikolinieritas. Metode ini memungkinkan evaluasi yang lebih tepat pada setiap titik data dengan mengintegrasikan wilayah dan waktu, sehingga menghasilkan estimasi koefisien yang lebih akurat untuk setiap wilayah pengamatan. Salah satu keunggulan utama dari LCR-GTWR adalah kemampuannya dalam menangkap heterogenitas yang ada dalam data serta mengatasi masalah multikolinieritas lokal melalui penerapan koefisien bias ridge yang disesuaikan secara lokal.

Penelitian terdahulu dilakukan oleh Hoang dkk (2017) menggunakan model LCR-GWR dalam menganalisis faktor-faktor yang memengaruhi wabah malaria di Provinsi DakNong, Vietnam. Hasil penelitian menunjukkan bahwa model LCR-GWR dapat meningkatkan kesesuaian hasil serta baik dalam mengukur kejadian malaria di Provinsi DakNong, Vietnam. Penelitian lain dilakukan oleh Fadliana dkk (2020) menggunakan model LCR-GWR dalam menganalisis faktor-faktor yang memengaruhi distribusi prevalensi stunting pada anak usia balita di Provinsi Nusa Tenggara Timur yang mengindikasikan permasalahan multikolinearitas lokal. Hasil penelitian menunjukkan bahwa model LCR-GWR mampu menghasilkan model yang lebih baik dibandingkan GWR model dalam mengatasi masalah multikolinearitas lokal pada kasus stunting di Provinsi Nusa Tenggara Timur. Penelitian terdahulu berfokus pada penerapan LCR-GWR dalam bidang kesehatan dan juga tidak mempertimbangkan efek temporal. Oleh karena itu, pada penelitian ini akan dilakukan analisis di bidang lainnya yaitu bidang ekonomi dan sosial dalam hal ini Tingkat Pengangguran Terbuka (TPT) dengan mempertimbangkan efek temporal.

TPT di Provinsi Sulawesi Selatan menunjukkan tren penurunan yang signifikan dalam beberapa tahun terakhir. Pada Februari 2024, TPT tercatat sebesar 4,90 persen. Angka ini menurun dari besaran 5,26 persen pada Februari 2023. Hal ini mencerminkan penurunan jumlah pengangguran sebanyak 24.621 orang dalam satu tahun, dengan total penduduk usia kerja di Sulawesi Selatan mencapai 7,08 juta orang, di mana 4,56 juta di antaranya merupakan angkatan

kerja. Dari jumlah tersebut, 4,32 juta orang sudah bekerja, sementara 239.589 orang masih menganggur (BPS Sulsel, 2023). Oleh karena itu, dalam memaksimalkan penurunan TPT di Sulawesi Selatan adalah dengan mengetahui faktor-faktor yang diperkirakan dapat memengaruhinya.

Beberapa faktor diduga memengaruhi TPT adalah indeks pembangunan manusia dan persentase penduduk miskin (Zanzibar dkk., 2024). Selain itu, penelitian lainnya juga menyebutkan bahwa TPT juga dipengaruhi oleh rerata lama sekolah, umur harapan hidup, kepadatan penduduk serta produk domestik regional bruto atas dasar harga berlaku (Rambe dkk., 2019; Alvia dkk., 2024; Romadhona dkk., 2024). Faktor-faktor tersebut juga memungkinkan adanya keragaman yang memengaruhi TPT di suatu daerah dengan daerah lainnya. Selain itu, juga menimbulkan adanya faktor yang saling berkorelasi antarwilayah. Berdasarkan uraian tersebut, peneliti akan mengkaji pengembangan model regresi spasial dengan mengatasi multikolinieritas lokal pada efek spasio-temporal yang diaplikasikan pada data Tingkat Pengangguran Terbuka (TPT) Sulawesi Selatan tahun 2020-2023 dengan judul **“Pemodelan *Locally Compensated Ridge-Geographically And Temporally Weighted Regression* Pada Data Tingkat Pengangguran Terbuka di Sulawesi Selatan**”. Model yang diperoleh nantinya dapat menjadi landasan dalam mengidentifikasi faktor-faktor yang berpengaruh signifikan pada TPT di Provinsi Sulawesi Selatan.

1.2 Batasan Penelitian

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan merupakan data temporal dengan rentang 2020 hingga 2023.
2. Metode penentuan jarak antar lokasi pengamatan menggunakan jarak *euclidean* dan fungsi pembobot yang digunakan adalah *Fixed Gaussian Kernel*.
3. Uji Kebaikan model menggunakan kriteria nilai R^2_{adj} , nilai *Root Mean Square Error* (RMSE) dan *Mean Absolut Deviation* (MAD).

1.3 Tujuan Penelitian dan Manfaat Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh estimasi parameter model *Locally Compensated Ridge-Geographically and Temporally Weighted Regression* Tingkat Pengangguran Terbuka di Sulawesi Selatan tahun 2020-2023.
2. Memperoleh faktor-faktor yang memengaruhi Tingkat Pengangguran Terbuka di Sulawesi Selatan menggunakan model *Locally Compensated Ridge-Geographically and Temporally Weighted Regression* tahun 2020-2023.

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Sumber pengetahuan terkait metode *Locally Compensated Ridge-Geographically and Temporally Weighted Regression* serta menjadi rujukan penelitian selanjutnya.
2. Sumber pengetahuan dan informasi mengenai faktor-faktor yang memengaruhi Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan Tahun 2020-2023 sehingga bermanfaat bagi pemerintah dan penentu kebijakan untuk membuat kebijakan lanjutan dalam mengatasi Tingkat Pengangguran Terbuka di setiap Kabupaten/kota.

1.4 Teori

1.4.1 Regresi Linier Berganda

Regresi linier berganda merupakan regresi dengan dua atau lebih variabel prediktor dan merupakan perluasan dari regresi linear sederhana (Daulay & Simamora, 2023). Tujuan analisis ini untuk mengetahui hubungan antara variabel respon dan sekumpulan variabel prediktor yang berkaitan (Ohaegbulem & Iheaka, 2024). Adapun modelnya dapat dituliskan pada Persamaan (1).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (1)$$

Persamaan (1) dapat dituliskan dalam bentuk matriks yaitu pada Persamaan (2).

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

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

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

dengan

- n : jumlah pengamatan
- p : jumlah prediktor
- y_i : nilai observasi variabel respon ke- i
- x_{ik} : nilai observasi variabel prediktor ke- k pada pengamatan ke- i
- β_0 : nilai konstanta atau *intercept* model regresi
- $\beta_1, \beta_2, \dots, \beta_k$: koefisien regresi variabel prediktor ke- k
- ε_i : residual pada pengamatan ke- i , dengan asumsi identik, independen dan berdistribusi normal dengan *mean* nol dan konstanta σ^2 (Fotheringham dkk., 2002).

1.4.2 Pengujian Heterogenitas Spasial

Efek heterogenitas adalah efek yang menunjukkan adanya keragaman antar lokasi. Jadi setiap lokasi mempunyai struktur dan parameter hubungan yang berbeda. Pengujian efek heterogenitas spasial dilakukan dengan uji *Breusch-Pagan* (BP). Hipotesis yang digunakan pada uji ini yaitu:

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

$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ dengan } i \neq j$ (Terjadi Heterogenitas spasial)

Nilai dari uji BP dihitung menggunakan Persamaan (3).

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} + \left(\frac{1}{T} \right) \left(\frac{\boldsymbol{\varepsilon}^T \mathbf{W} \boldsymbol{\varepsilon}}{\sigma^2} \right)^2 \sim \chi_{(\alpha)}^2 \quad (3)$$

dengan vektor $\mathbf{f}$ adalah $f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right)$ dan $T = \text{tr}[\mathbf{W}^T \mathbf{W} + \mathbf{W}^2]$

- BP : statistik uji *Breusch-Pagan*
 $\boldsymbol{\varepsilon}$: vektor residual ε_i berukuran $n \times 1$
 $\mathbf{f}^T$: $(f_1 \ f_2 \ \dots \ f_n)$
 $\mathbf{Z}$: matriks $\mathbf{X}$ yang distandarisasi, berukuran $n \times (p + 1)$
 $\mathbf{W}$: matriks pembobot W_i berukuran $n \times n$
 e_i^2 : residual untuk observasi ke- i dari hasil estimasi regresi dengan menggunakan OLS

Pengambilan keputusan pada uji ini adalah tolak H_0 jika nilai $BP > \chi_{\alpha,p}^2$ atau jika $p\text{-value} < \alpha$ dengan p adalah banyaknya variabel prediktor (Agustina dkk., 2022).

1.4.3 Geographically Weighted Regression

Geographically Weighted Regression (GWR) adalah pengembangan model regresi dengan masing-masing parameter dihitung pada setiap lokasi pengamatan, dengan kata lain setiap lokasi memiliki nilai parameter regresi pengamatan berbeda (Winarso & Yasin, 2016). GWR merupakan metode statistik yang digunakan untuk menganalisis heterogenitas spasial yang merupakan pengembangan dari regresi linier berganda (Fotheringham dkk., 2002). Dalam model GWR, variabel respon ditaksir dengan variabel prediktor yang masing-masing koefisien regresinya tergantung pada lokasi di mana data tersebut diamati. Model GWR dituliskan dalam Persamaan (4).

$$y_i = \beta_0(u_i, v_i) + \sum_k^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (4)$$

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

dengan

- y_i : nilai observasi variabel respon ke- i
 x_{ik} : nilai observasi variabel prediktor ke- k pada pengamatan ke- i

- $\beta_0(u_i, v_i)$: nilai konstanta atau *intercept* model regresi
 $\beta_k(u_i, v_i)$: koefisien regresi variabel prediktor ke- k
 ε_i : residual pada pengamatan ke- i , dengan asumsi identik, independen dan berdistribusi normal dengan *mean* nol dan konstanta σ^2 (Fotheringham dkk., 2002).

Pemodelan menggunakan GWR sangat bergantung pada pembobot spasial yang akan digunakan, sebab nilai pembobot ini mewakili letak data observasi satu dengan lainnya (Fotheringham dkk., 2002). Pembobotan dalam model GWR dapat dilakukan dengan berbagai metode, salah satunya melalui fungsi kernel yang digunakan untuk memperkirakan parameter model. Adapun fungsi kernel yang digunakan yakni gaussian dengan persamaan sebagai berikut:

$$w_j(u_i, v_i) = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{h} \right)^2 \right] \quad (5)$$

dengan h merupakan parameter penghalus (*bandwidth*) dan d_{ij} adalah jarak antara lokasi ke- i dan lokasi ke- j berdasarkan titik koordinat. Nilai d_{ij} dihitung dengan Persamaan (6).

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (6)$$

Parameter penghalus (*bandwidth*) adalah ukuran jarak yang menentukan seberapa besar pengaruh antar lokasi dalam fungsi pembobot (Daulay & Simamora, 2023). Pemilihan parameter penghalus yang optimal dapat dilakukan dengan menggunakan metode *Cross Validation* (CV). Parameter penghalus yang akan dipilih berdasarkan pada nilai CV minimum dengan menggunakan Persamaan (7).

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(h)]^2 \quad (7)$$

dengan

- CV : nilai *Cross Validation*
 y_i : nilai observasi variabel respon pada lokasi ke- i
 $\hat{y}_{\neq i}$: nilai dugaan dari model GWR tanpa melibatkan lokasi ke- i dalam proses estimasi

1.4.4 Geographically and Temporally Weighted Regression

Geographically and Temporally Weighted Regression (GTWR) merupakan pengembangan model dari GWR yang mempertimbangkan adanya variasi spasial (lokasi) dan variasi temporal (waktu) (Chu dkk., 2018). Model GTWR

untuk p variabel prediktor dengan variabel respon pada koordinat (u_i, v_i, t_i) untuk setiap pengamatan. Model GTWR dapat dituliskan pada Persamaan (8).

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i \quad (8)$$

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

dengan

- y_i : nilai observasi variabel respon ke- i
- x_{ik} : nilai observasi variabel prediktor ke- k pada pengamatan ke- i
- $\beta_0(u_i, v_i, t_i)$: nilai konstanta atau *intercept* model regresi
- $\beta_k(u_i, v_i, t_i)$: koefisien regresi variabel prediktor ke- k pada lokasi pengamatan ke- i dan waktu ke- i
- ε_i : residual pada pengamatan ke- i , dengan asumsi identik, independen dan berdistribusi normal dengan *mean* nol dan konstanta σ^2 (Fotheringham dkk., 2002).

Pendugaan parameter model GTWR untuk setiap k variabel pada lokasi ke- i diperoleh dengan menggunakan metode *Weighted Least Square* dengan memberi pembobot berbeda untuk setiap lokasi dan waktu. Pendugaan parameter pada model GTWR dapat dilihat pada Persamaan (9).

$$\hat{\beta}(u_i, v_i, t_i) = (X^T W(u_i, v_i, t_i) X)^{-1} X^T W(u_i, v_i, t_i) Y \quad (9)$$

dengan $W(u_i, v_i, t_i)$ merupakan matriks pembobot spasial yang dituliskan sebagai berikut.

$$W(u_i, v_i, t_i) = \begin{pmatrix} w_0(u_i, v_i, t_i) & 0 & \dots & 0 \\ 0 & w_1(u_i, v_i, t_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n(u_i, v_i, t_i) \end{pmatrix}$$

$W(u_i, v_i, t_i)$ merupakan matriks pembobot pada pengamatan (u_i, v_i) dan waktu ke t_i , sedangkan n adalah jumlah observasi.

1.4.5 Jarak Spasio-Temporal

Menentukan dan mengukur kedekatan spasio-temporal dalam sistem koordinat adalah masalah utama dalam penyusunan model GTWR, karena kedekatan titik observasi melibatkan dua aspek, yaitu kedekatan spasial dan kedekatan temporal (Conita & Purwaningsih, 2017). Jarak spasio-temporal dihitung menggunakan sistem koordinat *ellipsoidal* untuk menentukan seberapa dekat titik regresi dengan titik observasi di sekitarnya (Oktarina dkk., 2024). Kombinasi fungsi jarak spasial (d_{ij}^S) dan fungsi jarak temporal (d_{ij}^T) ditunjukkan pada Persamaan (10).

$$(d_{ij}^{ST})^2 = \lambda^* (d_{ij}^S)^2 + \mu (d_{ij}^T)^2 \quad (10)$$

dengan parameter λ^* dan μ menyatakan faktor skala penyeimbang efek yang berbeda untuk mengukur jarak spasial dan temporal. Dengan melakukan substitusi fungsi jarak *euclidian* menggunakan persamaan sebagai berikut.

$$(d_{ij}^S)^2 = (u_i - u_j)^2 + (v_i - v_j)^2 \quad (11)$$

$$(d_{ij}^T)^2 = (t_i - t_j)^2 \quad (12)$$

dan d_{ij} merupakan jarak antara titik i dan titik j . Sehingga, jarak euclidean spasio-temporal pada Persamaan (10) menjadi Persamaan (13).

$$(d_{ij}^{ST})^2 = \lambda^* \{(u_i - u_j)^2 + (v_i - v_j)^2\} + \mu(t_i - t_j)^2 \quad (13)$$

u_i menunjukkan koordinat latitude pada lokasi ke- i , dan v_i menunjukkan koordinat longitude pada lokasi ke- i . Sehingga W_{ij} dapat dituliskan sebagai berikut.

$$\begin{aligned} W_{ij} &= \exp \left\{ - \left(\frac{\lambda^* \{(u_i - u_j)^2 + (v_i - v_j)^2\} + \mu(t_i - t_j)^2}{h_{ST}^2} \right) \right\} \\ &= \exp \left\{ - \left(\frac{(u_i - u_j)^2 + (v_i - v_j)^2}{h_S^2} + \frac{(t_i - t_j)^2}{h_T^2} \right) \right\} \\ &= \exp \left\{ - \left(\frac{(d_{ij}^S)^2}{h_S^2} + \frac{(d_{ij}^T)^2}{h_T^2} \right) \right\} \\ &= W_{ij}^S \times W_{ij}^T \end{aligned} \quad (14)$$

dengan h_{ST}^2 , $h_S^2 = h_{ST}^2 / \lambda^*$ dan $h_T^2 = h_{ST}^2 / \mu$ adalah masing-masing parameter dari bandwidth spasial temporal, bandwidth spasial, dan bandwidth temporal.

Menurut Liu dkk (2017), misalkan τ merupakan parameter rasio $\frac{\mu}{\lambda}$ dengan $\lambda^* \neq 0$ maka Persamaan (13) dibagi dengan λ^* untuk memunculkan parameter τ sehingga dapat dituliskan menjadi Persamaan (15).

$$\frac{(d_{ij}^{ST})^2}{\lambda^*} = (u_i - u_j)^2 + (v_i - v_j)^2 + \tau(t_i - t_j)^2 \quad (15)$$

Parameter τ digunakan untuk menyesuaikan besar kecilnya pengaruh jarak temporal terhadap jarak spasial. Nilai parameter ini diperoleh menggunakan metode *CV* minimum melalui proses *iterative*, dimulai dengan inisialisasi nilai awal τ melalui Persamaan (16).

$$CV(\tau) = \sum_{i=1}^n [y_{=i} - \hat{y}_{\neq i}(\tau)]^2 \quad (16)$$

Kemudian, estimasi parameter μ dan λ^* diperoleh berdasarkan hasil estimasi τ yang menghasilkan nilai *CV* minimum (Huang dkk., 2010).

1.4.6 Multikolinieritas Lokal

Multikolinieritas lokal merujuk pada kondisi di mana terdapat hubungan linier yang sempurna atau hampir sempurna antara variabel-variabel prediktor pada setiap lokasi pengamatan. Deteksi masalah multikolinearitas lokal dalam model GTWR dapat dilakukan dengan menggunakan kriteria *Variance Inflation Factors*

(VIF) lokal. Terindikasi terjadi multikolinieritas ketika nilai $VIF > 10$ (Belsley, 1991). Nilai VIF lokal dihitung untuk setiap variabel prediktor di masing-masing area pengamatan, dengan matriks pembobot digunakan sebagai alat deteksi kolinearitas dalam model GTWR. Perhitungan nilai VIF dinyatakan pada Persamaan (17).

$$VIF_k(u_i, v_i, t_i) = \frac{1}{1 - R_k^2(u_i, v_i, t_i)} \quad (17)$$

dengan R_k^2 adalah R^2 lokal atau koefisien determinasi antara x_k dengan variabel prediktor lainnya untuk setiap wilayah (He & Zhang, 2024).

Menurut Fotheringham dkk (2002), R^2 lokal dihitung menggunakan persamaan sebagai berikut.

$$\begin{aligned} R^2(u_i, v_i, t_i) &= \frac{JKR^w}{JKT^w} \\ &= \frac{JKT^w - JKG^w}{JKT^w} \\ &= 1 - \frac{JKG^w}{JKT^w} \end{aligned} \quad (18)$$

dengan JKT^w adalah Jumlah Kuadrat Total model GTWR, yang dituliskan pada Persamaan (19) sebagai berikut

$$JKT^w = \sum_j w_j(u_i, v_i, t_i) [y_j - \bar{y}]^2 \quad (19)$$

dan JKG^w adalah Jumlah Kuadrat Galat model GTWR, yang dituliskan pada persamaan berikut.

$$JKG^w = \sum_j w_j(u_i, v_i, t_i) [y_j - \hat{y}_j]^2 \quad (20)$$

Menurut Gujarati & Porter (2015) keberadaan multikolinieritas sempurna di antara variabel prediktor menyebabkan ketidakmampuan untuk memperoleh estimasi yang unik dari semua parameter, serta membuat nilai residual tidak dapat didefinisikan. Akibatnya, tidak ada kesimpulan statistik apapun yang dapat ditarik tentang hasil tersebut dari sampel yang ada. Dengan kata lain, dalam situasi multikolinieritas sempurna, baik estimasi maupun pengujian hipotesis terkait koefisien regresi individual menjadi tidak mungkin dilakukan. Sementara itu, pada kasus multikolinieritas tidak sempurna, estimasi parameter regresi masih dapat dilakukan, namun nilai residual cenderung lebih besar.

1.4.7 Regresi Ridge

Regresi Ridge merupakan suatu metode yang dirancang untuk menstabilkan estimasi parameter regresi yang dipengaruhi oleh adanya multikolinieritas. Konsep ini pertama kali diperkenalkan oleh A.E. Hoerl pada tahun 1962. Metode ini bertujuan untuk mengatasi masalah kondisi buruk yang muncul akibat

tingginya korelasi antar variabel prediktor dalam model regresi. Hal ini dapat menyebabkan matriks $X^T X$ hampir singular, yang pada gilirannya menghasilkan estimasi parameter model regresi yang tidak konsisten (Ölmez & Yağanoğlu, 2023).

Penggunaan metode regresi ridge didasarkan pada modifikasi metode kuadrat terkecil yakni dengan menambahkan suku λI pada $X^T X$ sebelum dilakukan invers, sehingga menyebabkan multikolinieritas dapat diatasi. Estimasi parameter regresi ridge dituliskan pada Persamaan (21).

$$(X^T X + \lambda I) \hat{\beta}_R = X^T Y \text{ atau } \hat{\beta}_R = (X^T X + \lambda I)^{-1} X^T Y \quad (21)$$

dengan nilai $0 < \lambda < 1$ adalah parameter bias ridge yang dihitung sehingga nilai $\hat{\beta}_R$ stabil. Jika nilai $\lambda = 0$ maka estimator ridge bernilai sama dengan estimator kuadrat terkecil (Ifadah, 2011; Akhtar dkk., 2024).

Penambahan koefisien bias ridge mengakibatkan nilai dari elemen diagonal matriks $(X^T X + \lambda I)^{-1}$ menjadi kecil sehingga rata-rata nilai kuadrat residual menjadi kecil. Hal ini menunjukkan bahwa koefisien regresi menjadi stabil. Untuk pemilihan nilai koefisien bias ridge dihitung menggunakan rumus yang dituliskan pada Persamaan (22)

$$\lambda(u_i, v_i, t_i) = \frac{p \hat{\sigma}^2}{\hat{\beta}_{WLS}^T (X^T X) \hat{\beta}_{WLS}} \quad (22)$$

dengan p merupakan jumlah parameter β tanpa β_0 , $\hat{\sigma}^2$ merupakan *Mean square error (MSE)*, dan $\hat{\beta}_{WLS}$ merupakan estimasi yang diperoleh dengan WLS (Lawless & Wang, 1976).

1.4.8 Locally Compensated Ridge-Geographically Weighted Regression

Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR) merupakan pengembangan dari model *Geographically Weighted Ridge Regression* (GWRR) dengan menerapkan satu koefisien bias tertentu untuk setiap wilayah. Dengan demikian, jika terdapat N wilayah pengamatan, maka akan ada n koefisien bias ridge yang memiliki koefisien berbeda. LCR-GWR menghasilkan koefisien bias ridge secara lokal, di mana parameter ridge bervariasi sesuai dengan tingkat kolinearitas antar variabel prediktor pada masing-masing wilayah, sehingga koefisien parameter yang dihasilkan oleh model menjadi lebih akurat (Fadliana dkk., 2019). Model LCR-GWR dituliskan dalam Persamaan (23).

$$y_i = \beta_0(u_i, v_i) + \sum_k^p \beta_k(u_i, v_i, \lambda_i) x_{ik} + \varepsilon_i \quad (23)$$

dengan $\beta_k(u_i, v_i, \lambda_i)$ adalah koefisien regresi lokal untuk variabel prediktor ke- k di lokasi i dan λ_i merupakan nilai tertentu dari koefisien bias ridge pada lokasi ke- i .

1.4.9 Tingkat Pengangguran Terbuka

Tingkat Pengangguran Terbuka (TPT) adalah salah satu indikator ekonomi yang sangat penting, menggambarkan persentase dari angkatan kerja yang tidak memiliki pekerjaan tetapi aktif mencari pekerjaan. TPT mencakup individu yang sedang mencari pekerjaan, mereka yang sedang mempersiapkan usaha, dan mereka yang baru mendapatkan tawaran pekerjaan tetapi belum mulai bekerja (BPS, 2023). Dalam konteks ini, TPT memberikan gambaran tentang kesehatan pasar tenaga kerja di suatu negara atau wilayah.

Beberapa faktor penyebab pengangguran terbuka meliputi ketidakcocokan keterampilan antara pencari kerja dan kebutuhan pasar, di mana banyak lulusan baru tidak memiliki keterampilan yang relevan dengan lowongan pekerjaan yang tersedia. Selain itu, minimnya pengalaman kerja sering kali menjadi hambatan bagi lulusan baru untuk mendapatkan posisi yang diinginkan, karena banyak perusahaan lebih memilih kandidat dengan pengalaman. Diskriminasi dalam proses perekrutan juga dapat menghambat peluang kerja bagi kelompok tertentu, sementara kondisi ekonomi yang tidak stabil dapat menyebabkan perusahaan mengurangi jumlah karyawan atau bahkan melakukan pemutusan hubungan kerja (Rahman & Riani, 2023).

Penurunan angka TPT memiliki signifikansi yang besar untuk peningkatan kesejahteraan masyarakat dan stabilitas ekonomi. Tingkat pengangguran yang tinggi berpotensi menyebabkan peningkatan kemiskinan, ketidakstabilan sosial, dan penurunan daya beli masyarakat (Azzahra dkk., 2024). Oleh karena itu, penting untuk memprioritaskan upaya dalam menciptakan lebih banyak pekerjaan berkualitas, memperluas akses pendidikan dan pelatihan, serta memperkuat kolaborasi antara pemerintah dan sektor swasta dalam pengembangan lapangan kerja. Dengan implementasi strategi yang efektif, diharapkan dapat terwujud kondisi ekonomi yang lebih stabil dan adil bagi seluruh masyarakat Indonesia khususnya di Sulawesi Selatan.

BAB II METODE PENELITIAN

2.1. Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder yang bersumber dari laman *website* resmi Badan Pusat Statistik Provinsi Sulawesi Selatan. Data tersebut merupakan data temporal dengan rentang waktu dari tahun 2020 hingga 2023. Adapun unit pengamatan yakni terdiri dari 24 kabupaten/kota di Sulawesi Selatan dan dapat dilihat pada Lampiran 1.

2.2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas variabel respon Tingkat Pengangguran Terbuka (TPT) dan enam variabel prediktor yang diduga memengaruhi TPT di 24 kabupaten/kota di Provinsi Sulawesi Selatan. Variabel yang digunakan dituliskan dalam Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Definisi	Satuan
Y	Tingkat Pengangguran Terbuka (TPT)	TPT adalah persentase dari angkatan kerja yang tidak memiliki pekerjaan namun aktif mencari pekerjaan.	%
X_1	Indeks Pembangunan Manusia (IPM)	IPM adalah ukuran yang digunakan untuk menilai kualitas hidup dan pembangunan manusia di suatu negara.	Indeks
X_2	Persentase Penduduk Miskin (PPM)	PPM adalah mengukur proporsi penduduk yang hidup di bawah garis kemiskinan	%
X_3	Rerata Lama Sekolah (RLS)	RLS adalah rata-rata jumlah tahun pendidikan yang telah dijalani oleh penduduk dalam kelompok usia tertentu.	Tahun
X_4	Umur Harapan Hidup (UHH)	UHH rata-rata usia yang diharapkan dapat dicapai oleh individu saat lahir, yang mencerminkan kondisi kesehatan dan kualitas hidup di suatu wilayah.	Tahun

Variabel	Keterangan	Definisi	Satuan
X_5	Kepadatan Penduduk (KP)	KP mengukur jumlah penduduk per unit area (biasanya per kilometer persegi) yang memberikan gambaran tentang distribusi populasi di suatu wilayah dan dapat mempengaruhi perencanaan pembangunan, penyediaan layanan publik, dan pengelolaan sumber daya.	Jiwa
X_6	Produk Domestik Regional Bruto Atas Dasar Harga Berlaku (PDRB ADHB)	PDRB ADHB adalah total nilai semua barang dan jasa yang diproduksi dalam suatu daerah selama periode tertentu, dihitung dengan harga yang berlaku pada saat itu.	Rupiah

2.3. Metode Analisis

Metode analisis yang digunakan pada penelitian ini adalah *Locally Compensated Ridge-Geographically and Temporally Weighted Regression*. Pengolahan data dilakukan dengan bantuan *software* R-studio. Adapun tahapan analisis yang dilakukan pada penelitian ini sebagai berikut.

1. Melakukan eksplorasi data dengan menggunakan peta sebaran nilai TPT setiap kabupaten/kota.
2. Mengidentifikasi efek heterogenitas spasial dengan menggunakan uji *Breusch-Pagan* (BP) berdasarkan Persamaan (3). Adapun hipotesisnya sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 \text{ (Tidak terjadi heterogenitas spasial)}$$

$$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ dengan } i \neq j \text{ (Terjadi heterogenitas spasial)}$$
3. Menghitung nilai *Variance Inflation Factor* (VIF) untuk mengetahui multikolinearitas global dan lokal.
4. Analisis model GTWR
 - 1) Penentuan parameter rasio spasio-temporal (τ)
 - a. Menentukan fungsi pembobot dengan menggunakan pembobot *gaussian* pada Persamaan (5)
 - b. Menghitung penduga parameter GTWR
 - c. Menghitung nilai $\hat{y}_\#(\tau) = \mathbf{X}^T \mathbf{t} \hat{\boldsymbol{\beta}}(u_i, v_i, t_i)$ menggunakan nilai τ dengan tidak memasukkan lokasi amatan ke- i .

- d. Meminimumkan nilai *Cross Validation (CV)* berdasarkan τ dengan menggunakan rumus sebagai berikut.

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

- 2) Penentuan parameter spasial (λ) dan parameter (μ)
- Menentukan fungsi pembobot dengan menggunakan pembobot *gaussian* pada Persaman (5)
 - Menghitung penduga parameter GTWR
 - Menghitung nilai $\hat{y}_{\neq}(\lambda^*) = \mathbf{X}^T \mathbf{i} \hat{\boldsymbol{\beta}}(u_i, v_i, t_i)$ menggunakan nilai λ^* dengan tidak memasukkan lokasi amatan ke- i .
 - Meminimumkan nilai *Cross Validation (CV)* berdasarkan λ^* dengan menggunakan rumus sebagai berikut.

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

- 3) Penentuan *bandwidth* optimum
- Menentukan jarak lokasi amatan dan waktu dengan menggunakan Persamaan (13).
 - Menentukan fungsi pembobot spasio-temporal dengan menggunakan pembobot *gaussian* pada Persaman (5)
 - Menghitung penduga parameter GTWR
 - Menghitung nilai $\hat{y}_{\neq}(h_{ST}) = \mathbf{X}^T \mathbf{i} \hat{\boldsymbol{\beta}}(u_i, v_i, t_i)$ menggunakan nilai h_{ST} dengan tidak memasukkan lokasi amatan ke- i .
 - Meminimumkan nilai *Cross Validation (CV)* berdasarkan h_{ST} dengan menggunakan Persamaan (16)

$$CV(h_{ST}) = \sum_{i=1}^n [y_{=i} - \hat{y}_{\neq i}(h_{ST})]^2$$

- 4) Melakukan estimasi parameter model GTWR
5. Analisis GTWR dengan *Locally Compensated Ridge* untuk tiap lokasi dan waktu.
- Melakukan estimasi parameter model *Locally Compensated Ridge-Geographically and Temporally Weighted Regression* (LCR-GTWR)
 - Menentukan koefisien bias ridge untuk masing-masing wilayah $\lambda_i(u_i, v_i, t_i)$ menggunakan Persamaan (22) berdasarkan nilai jumlah parameter $\hat{\beta}$ tanpa $\hat{\beta}_0$ serta MSE yang diperoleh dari pemodelan GTWR.
 - Melakukan pemodelan LCR-GTWR untuk masing-masing kabupaten/kota menggunakan Persamaan (23).
6. Melakukan pengujian secara parsial pada model LCR-GTWR dengan menggunakan rumus berikut.

$$t = \frac{\hat{\beta}_k(u_i, v_i, t_i, \lambda_i)}{se_{\hat{\beta}_k}}$$

dengan hipotesis

$H_0: \beta_k(u_i, v_i, t_i, \lambda_i) = 0$ (parameter model β_k pada lokasi dan waktu (u_i, v_i, t_i) tidak berpengaruh signifikan terhadap model)

$H_1: \beta_k(u_i, v_i, t_i, \lambda_i) \neq 0$ (parameter model β_k pada lokasi dan waktu (u_i, v_i, t_i) berpengaruh signifikan terhadap model)

dengan,

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

7. Pengujian kebaikan model LCR-GTWR dengan menghitung nilai R_{adj}^2 , *Root Mean Square Error* (RMSE) dan *Mean Absolut Deviation* (MAD) dengan rumus sebagai berikut.

$$R_{adj}^2 = 1 - \frac{(1 - R)(n - 1)}{n - k - 1}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$MAD = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

8. Melakukan interpretasi model yang diperoleh.
9. Menarik kesimpulan.