PEMODELAN *GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION* DENGAN *PRINCIPAL COMPONENT ANALYSIS* PADA DATA YANG MENGANDUNG MULTIKOLINEARITAS
(Studi kasus: Indeks Ketahanan Pangan di Provinsi Sulawesi Selatan pada Tahun 2021-2023)

NADIA NAZWADIAH CAESAR

H051201027

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

2024

PEMODELAN *GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION* DENGAN *PRINCIPAL COMPONENT ANALYSIS* PADA DATA YANG MENGANDUNG MULTIKOLINEARITAS

(Studi kasus: Indeks Ketahanan Pangan di Provinsi Sulawesi Selatan pada Tahun 2021-2023)

NADIA NAZWADIAH CAESAR

H051201027

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

2024

PEMODELAN *GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION* DENGAN *PRINCIPAL COMPONENT ANALYSIS* PADA DATA YANG MENGANDUNG MULTIKOLINEARITAS
(Studi kasus: Indeks Ketahanan Pangan di Provinsi Sulawesi Selatan pada Tahun 2021-2023)

NADIA NAZWADIAH CAESAR

H051201027

Skripsi

Diajukan sebagai salah satu syarat untuk mencapai gelar Sarjana Statistika

Program Studi Statistika

pada

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

NOVEMBER 2024

PEMODELAN *GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION* DENGAN *PRINCIPAL COMPONENT ANALYSIS* PADA DATA YANG MENGANDUNG MULTIKOLINEARITAS
(Studi kasus: Indeks Ketahanan Pangan di Provinsi Sulawesi Selatan pada Tahun 2021-2023)

yang disusun dan diajukan oleh

NADIA NAZWADIAH CAESAR

H051201027

Skripsi,

telah dipertahankan di depan Panitia Ujian Sarjana Statistika pada tanggal 29 November 2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:	Mengetahui:
Pembimbing Tugas Akhir,	Ketua Program Studi,
Siswanto, S.Si., M.Si.	**Dr. Anna Islamiyati, S.Si., M.Si.**
NIP. 199201072019031012	NIP. 19770808 200501 2 002

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Pemodelan *Geographically and Temporally Weighted Regression* Dengan *Principal Component Analysis* pada Data yang Mengandung Multikolinearitas (Studi kasus: Indeks Ketahanan Pangan di Provinsi Sulawesi Selatan pada Tahun 2021-2023)" adalah benar karya saya dengan arahan dari Siswanto, S.Si., M.Si. sebagai Pembimbing Utama. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 29 November 2024

Nadia Nazwadiah Caesar

NIM. H051201027

UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Allah SWT. atas segala limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan sahabatnya.

Penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada **Bapak Siswanto, S.Si., M.Si.** selaku Pembimbing Utama yang telah memberikan begitu banyak ilmu, waktu, dan dorongan semangat kepada penulis selama proses penulisan tugas akhir ini. Terima kasih kepada **Ibu Anisa, S.Si., M.Si.** dan **Bapak Andi Kresna Jaya, S.Si., M.Si.** selaku Tim Penguji yang senantiasa memberikan saran dan kritikan dalam penyempurnaan penulisan tugas akhir ini. Terima kasih kepada **Pimpinan Universitas Hasanuddin, Departemen Statistika, Jajaran Dosen,** dan **Staf Departemen Statistika** yang telah memfasilitasi, memberikan ilmu bermanfaat, dan membantu penulis selama menempuh studi.

Penghargaan dan ucapan terima kasih yang tulus juga penulis ucapkan kepada orang tua terkasih **Aten Djou** dan **Nisma Abdurrahman** yang telah memberikan dukungan penuh serta pengorbanan luar biasa yang telah menemani langkah penulis dengan doa dan restu mulianya. Ucapan terima kasih juga penulis haturkan kepada kedua adik tercinta penulis, **Naila Syakirah** dan **Ahmad Wildan Al-Furqon Djou**. Penulis juga mengucapkan terima kasih kepada teman-teman Mblo (**Aliffa Cahya Potutu**, **Dwi Angriani Habie**, **Nabilah Zaizafun Katili**, **Rayhandani Limonu**, dan **Yulia Ibrahim**) serta **Sitti Wanda Paramitha Hemeto**, **Tryana Sefya Saleh**, dan **Ahmad Subhan Hapulu** yang walaupun terpisahkan jarak tetapi selalu mendukung dan mendengarkan keluh kesah penulis.

Ucapan terima kasih juga penulis sampaikan kepada YMMA (**Aulia, Aliah, Jihan,** dan **Nahla)**, Ciwi-Ciwi Among (**Rahmi, Cynthia, Laurine, Ica, Radia, Afika, Aish, Ayu, Isna, Rifdah, Parida, Pebi, Putri**, dan **Stansye**), Among Us (**Alip, Lili, Yoel, Fadlan, Theo, Rais, Bahar, Hakam, Faldi, Ryval,** dkk.), teman-teman MAN di Unhas (**Nurla**. **Rania**, dan **Fatur**), teman-teman Statistika Angkatan 2020, KKNT Pemetaan dan Penataan Lingkungan Tamalanrea Indah Gel. 110, Organisasi Unhas MUN, serta teman-teman dan mentor MSIB PT Paragon Technology & Innovation yang telah berjuang bersama-sama selama masa perkuliahan dan menciptakan banyak kenangan indah yang tak akan bisa penulis lupakan.

Makassar, 29 November 2024

Nadia Nazwadiah Caesar

ABSTRAK

Nadia Nazwadiah Caesar. **Pemodelan *Geographically and Temporally Weighted Regression* Dengan *Principal Component Analysis* pada Data yang Mengandung Multikolinearitas (Studi kasus: Indeks Ketahanan Pangan di Provinsi Sulawesi Selatan pada Tahun 2021-2023)** (dibimbing oleh Siswanto).

Latar Belakang. *Geographically Temporally and Weighted Regression* (GTWR) adalah metode regresi spasial yang menangani heterogenitas spasial dan temporal secara sekaligus pada data. Akan tetapi, data dapat menghadapi masalah pelanggaran asumsi seperti multikolinearitas, yaitu kondisi ketika variabel prediktor dalam model regresi memiliki korelasi yang tinggi dengan variabel prediktor lainnya. Salah satu cara untuk mengatasi multikolinearitas adalah dengan melakukan *Principal Component Analysis* (PCA) pada variabel prediktor. Oleh karena itu, pada penelitian ini dilakukan pemodelan faktor-faktor yang memengaruhi Indeks Ketahanan Pangan untuk setiap kabupaten/kota di Provinsi Sulawesi Selatan pada tahun 2021-2023 menggunakan GTWR dengan PCA. **Tujuan**. Penelitian ini bertujuan untuk memperoleh model terbaik dari model GTWR dan GTWR-PCA pada Indeks Ketahanan Pangan di Sulawesi Selatan pada tahun 2021-2023. **Metode.** Penelitian ini menggunakan metode GTWR dengan PCA. **Hasil.** Data yang mengandung multikolinearitas diatasi dengan PCA sehingga menghasilkan *dataset* baru yang digunakan pada pemodelan. Pemodelan GTWR dengan data yang diatasi dengan PCA menghasilkan 72 model yang mewakili setiap lokasi dan waktu. Model GTWR-PCA menghasilkan model yang lebih baik dibandingkan dengan model GTWR tanpa PCA yakni dengan nilai R^2 sebesar 99.6945% dan *Root Mean Square Error* (RMSE) sebesar 1.5606. **Kesimpulan**. Model terbaik dalam menjelaskan IKP di Provinsi Sulawesi Selatan pada tahun 2021-2023 adalah model GTWR-PCA. Amatan dengan model GTWR-PCA terbaik adalah Kabupaten Kepulauan Selayar pada tahun 2023 dengan nilai R^2 lokal sebesar 99.9759%.

Kata Kunci: Heterogenitas, Ketahanan Pangan, Multikolinearitas, Regresi, Spasial, Temporal

ABSTRACT

Nadia Nazwadiah Caesar. **Geographically and Temporally Weighted Regression with Principal Component Analysis in Modelling of Data with Multicollinearity (Case study: Food Security Index in South Selatan for the Period of 2021-2023)** (supervised by Siswanto).

Background. Geographically Temporally and Weighted Regression (GTWR) is a spatial regression method that handles spatial and temporal heterogeneity simultaneously in data. However, the data may have assumption violation problems such as multicollinearity, which is a condition when the predictor variables in the regression model have a high correlation with other predictor variables. One way to overcome multicollinearity is to perform Principal Component Analysis (PCA) on the predictor variables. Therefore, in this study, the factors influencing the Food Security Index for each regency/city in South Sulawesi Province in 2021-2023 were modeled using GTWR with PCA. **Aim**. This study aims to obtain the best model from the GTWR and GTWR-PCA models on the Food Security Index in South Sulawesi in 2021-2023. **Methods**. This research uses the GTWR method with PCA. **Results**. Data containing multicollinearity was resolved with PCA resulting in a new dataset used in modeling. GTWR modeling with PCA-resolved data resulted in 72 models representing each location and time. The GTWR-PCA model produces a better model than the GTWR model without PCA, with an R^2 value of 99.6945% and a Root Mean Square Error (RMSE) of 0.1430. **Conclusion**. The best model in explaining Food Security Index in South Sulawesi Province in 2021-2023 is the GTWR-PCA model. The observation with the best GTWR-PCA model is Selayar Islands Regency in 2023 with a local R^2 value of 99.9759%.

Keywords: Food Security, Heterogeneity, Multicollinearity, Regression, Spatial, Temporal

DAFTAR ISI

Halaman

DAFTAR TABEL

Nomor Urut Halaman

DAFTAR GAMBAR

DAFTAR LAMPIRAN

DAFTAR LAMBANG/SINGKATAN

Lambang/Singkatan	**Arti dan Penjelasan**
GWR	*Geographically Weighted Regression*
GTWR	*Geographically and Temporally Weighted Regression*
PCA	*Principal Component Analysis*
PCR	*Principal Component Regression*
GWR-PCA	*Geographically Weighted Regression-Principal Component Analysis*
GTWR-PCA	*Geographically and Temporally Weighted Regression-Principal Component Analysis*
IKP	Indeks Ketahanan Pangan
IPM	Indeks Pembangunan Manusia
UHH	Umur Harapan Hidup
RMSE	*Root Mean Square Error*
VIF	*Variance Inflation Factor*
PC	*Principal Component*
WLS	*Weighted Least Square*
CV	*Cross Validation*
x_{ik}	Nilai variabel prediktor ke-k pada pengamatan ke-i
z_{ik}	Nilai variabel prediktor hasil standardisasi ke-k pada pengamatan ke-i
C_{ik}	Nilai komponen utama ke-k pada pengamatan ke-i
$\boldsymbol{\beta}$	Matriks parameter model regresi
$\boldsymbol{W}$	Matriks pembobot
u_i	Latitude atau garis lintang pada lokasi ke-i
v_i	Longitude atau garis bujur pada lokasi ke-i
t_i	Waktu ke-i
γ_S	*Bandwidth* Spasial
γ_T	*Bandwidth* Temporal
γ_{ST}	*Bandwidth* Spasial-Temporal
d_{ij}^{ST}	Jarak Spasial-Temporal
w_{ij}^{ST}	Fungsi Pembobot Spasial-Temporal
λ	Parameter jarak spasial
μ	Parameter jarak temporal
τ	Parameter rasio

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan analisis yang dilakukan untuk melihat pengaruh dari beberapa variabel prediktor terhadap variabel responnya (Yusuf dkk., 2024). Salah satu pengembangan dari analisis regresi adalah analisis yang dilakukan pada data wilayah yaitu analisis regresi spasial (Ramadhani dkk., 2020). Kondisi dari suatu wilayah pengamatan dengan wilayah pengamatan lainnya dapat memengaruhi hasil analisis. Salah satu kondisi yang dapat terjadi adalah heterogenitas spasial, yaitu kondisi yang mengacu pada distribusi yang tidak merata dari berbagai atribut geospasial di dalam wilayah geografis tertentu (Song dkk., 2020). Hal yang menyebabkan heterogenitas spasial adalah perbedaan pengaruh faktor tertentu terhadap hasil di tiap wilayah, sehingga variasi koefisien regresi dari faktor tersebut akan bergantung pada wilayahnya (Haining dan Li, 2021). Heterogenitas spasial dapat diatasi dengan menggunakan metode *Geographically Weighted Regression* (Lu dkk., 2014).

Geographically Weighted Regression (GWR) adalah metode regresi yang membahas non-stasioneritas spasial dalam analisis regresi (Collins, 2010). Berbeda dengan model regresi linier tradisional, estimasi model pada model GWR dilakukan secara lokal pada data spasial (Wheeler, 2021). Data spasial sering kali mengalami perubahan seiring waktu, sehingga Huang dkk. (2010) menambahkan dimensi temporal ke dalam matriks pembobot untuk memperhitungkan ketidakstasioneritas temporalnya juga. Metode ini yang kemudian disebut sebagai metode *Geographically Temporally and Weighted Regression* (GTWR), yaitu metode regresi spasial yang menangani non-stasioneritas spasial dan temporal secara sekaligus pada data (Yang dkk., 2019).

Metode GTWR memiliki beberapa asumsi yang perlu dipenuhi, diantaranya asumsi multikolinearitas, normalitas, dan heterogenitas spasial-temporal (Supianti, 2023). Akan tetapi, dalam praktiknya data yang diperoleh dapat menghadapi masalah pelanggaran asumsi seperti multikolinearitas, yaitu kondisi ketika variabel prediktor dalam model regresi memiliki korelasi yang tinggi dengan variabel prediktor lainnya. Korelasi yang tinggi dapat disebabkan apabila perubahan pada satu variabel diikuti dengan perubahan yang konsisten pada variabel lain (Mandloi, 2018). Multikolinearitas pada data dapat menyebabkan interval kepercayaan yang lebih luas dan hasil estimasi yang kurang kuat berdasarkan variabel prediktornya. Hal ini mengakibatkan model dengan data yang mengandung multikolinearitas dapat memiliki hasil estimasi yang kurang akurat (Shrestha, 2020). Salah satu cara untuk mengatasi multikolinearitas adalah dengan melakukan *Principal Component Analysis* (PCA) pada variabel prediktor (Sulaiman dkk., 2021). Tujuan dari PCA adalah untuk mengidentifikasi sekumpulan fitur yang direduksi ke dalam subruang yang berdimensi lebih rendah, tanpa harus kehilangan banyak informasi yang diperlukan (Kherif dan Latypova, 2019).

Salah satu pengembangan metode PCA adalah metode *Principal Component Regression* (PCR), yaitu analisis regresi yang dilakukan dengan mengubah variabel prediktor menjadi komponen utama hasil PCA (Mikis dkk., 2022). Penggunaan PCA juga dapat diaplikasikan pada bentuk regresi lainnya seperti regresi spasial GWR (Wagner, 2013). Penggabungan metode GWR dengan PCA menunjukkan hasil yang lebih efektif dibandingkan saat menggunakan PCR saja sebab metode GWR-PCA dapat menunjukkan hubungan spasial antar variabel dan pola lokal dari data (Shrestha dan Luo, 2017). Penelitian yang dilakukan oleh Zhu dkk. (2021) membuktikan bahwa metode GWR-PCA menghasilkan kecocokan yang lebih baik dibandingkan metode PCR dalam menganalisis mekanisme dampak sosial-ekonomi terhadap nilai jasa ekosistem di wilayah Beijing-Tianjin-Hebei. Hal ini ditunjukkan dengan kenaikan nilai R^2 sebesar 55.63% pada metode PCR menjadi sebesar 90.14% pada metode GWR-PCA.

Penambahan faktor spasial pada penelitian tersebut berhasil membuktikan adanya peningkatan nilai kebaikan model. Selain itu, penambahan faktor temporal juga terbukti lebih baik pada penelitian yang dilakukan oleh Oktarina dkk. (2024) mengenai Indeks Pembangunan Manusia di Provinsi Bengkulu. Akurasi yang diperoleh untuk pemodelan GTWR pada penelitian ini adalah nilai R^2 sebesar 99.98%. Metode ini terbukti lebih baik dibandingkan pemodelan menggunakan GWR yang hanya memperoleh nilai R^2 sebesar 99.74%. Selanjutnya, perbandingan antara GWR dan GTWR dengan menggunakan PCA telah dilakukan oleh Fu dkk. (2022) mengenai faktor-faktor yang memengaruhi $\mathrm{PM}_{2.5}$ dan PM_{10} di wilayah Heilongjiang pada tahun 2014-2018. Model GTWR-PCA terbukti lebih baik dengan nilai R^2 sebesar 76% pada $\mathrm{PM}_{2.5}$ dan sebesar 71% pada PM_{10} dibandingkan model GWR-PCA yang hanya memiliki nilai R^2 sebesar 68% pada $\mathrm{PM}_{2.5}$ dan sebesar 63% pada PM_{10}.

Permasalahan yang dibahas pada penelitian sebelumnya terkait dengan kondisi sosial-ekonomi serta lingkungan pada wilayah tersebut. Permasalahan lain yang perlu menjadi perhatian adalah ketahanan pangan. Hal ini sejalan dengan Undang-Undang Pangan No 18 tahun 2012 yang menyatakan bahwa pemenuhan pangan merupakan bagian dari hak asasi manusia (PPK-LIPI, 2012). Ketahanan pangan dari suatu wilayah diukur menggunakan sistem penilaian yang disebut Indeks Ketahanan Pangan (IKP), yaitu suatu ukuran yang mengevaluasi ketersediaan, aksesibilitas, dan keterjangkauan pangan suatu wilayah dengan menilai seberapa baik wilayah tersebut dalam menyediakan makanan yang cukup, aman, dan bergizi bagi penduduknya. Kebijakan pangan nasional dapat dirumuskan menggunakan IKP sehingga dapat memenuhi peningkatan ketahanan pangan dan kebutuhan penduduk (Salasa, 2021).

Nilai IKP suatu wilayah dengan wilayah lain dapat memiliki tingkatan yang berbeda-beda. Lima provinsi dengan tingkat ketahanan terbaik di Indonesia pada periode 2023 secara berturut-turut adalah Bali (87.65), Jawa Tengah (84.80), DKI Jakarta (83.30), Sulawesi Selatan (83.36), dan Sumatera Barat (83.36). Berdasarkan informasi tersebut, Provinsi Sulawesi Selatan berhasil menjadi provinsi dengan tingkat ketahanan terbaik di Sulawesi. Selain itu, Provinsi Sulawesi Selatan juga secara konsisten berhasil meningkatkan ketahanan pangannya selama periode 2021 hingga 2023 dengan nilai IKP secara berturut-turut sebesar 80.82, 81.83, dan 83.36. Nilai IKP tersebut mencakup keseluruhan Provinsi Sulawesi Selatan. Akan tetapi, terdapat perbedaan tingkat IKP pada 24 kabupaten/kota di Provinsi Sulawesi Selatan (Badan Pangan Nasional, 2023).

Perbedaan tingkat IKP pada kabupaten/kota di provinsi yang sama dapat disebabkan oleh perbedaan faktor-faktor yang memengaruhi IKP di wilayah tersebut. Adapun beberapa faktor yang dapat memengaruhi kondisi IKP di suatu wilayah diantaranya Indeks Pembangunan Manusia (IPM) (Azizah dan Ratnasari, 2023), Umur Harapan Hidup (UHH) saat lahir (Yulianti dan Ratnasari, 2020), pengeluaran per kapita disesuaikan (Punggodewi dan Pratiwi, 2020), persentase penduduk yang hidup di bawah garis kemiskinan (Muttaqin dkk., 2022), persentase rumah tangga dengan proporsi pengeluaran untuk pangan lebih dari 65% terhadap total pengeluaran (Prasetya, 2024), dan persentase rumah tangga tanpa akses ke air bersih (Zulaika dkk., 2024). Akan tetapi, terdapat faktor yang merupakan turunan dari faktor lainnya, salah satunya adalah faktor IPM yang menggunakan indikator AHH pada perhitungannya (Yektiningsih, 2018). Hal ini menunjukkan bahwa kedua faktor tersebut dapat memiliki korelasi yang tinggi, sebab perubahan pada faktor IPM diikuti oleh perubahan yang konsisten pada variabel AHH (Mandloi, 2018).

Faktor yang memiliki korelasi yang tinggi dengan faktor lainnya dapat menyebabkan terjadinya multikolinearitas pada data. Selain itu, perbedaan karakteristik wilayah diduga menjadi hal yang menyebabkan perbedaan kondisi IKP tersebut. Peningkatan nilai IKP pada tahun ke tahun juga menunjukkan adanya perubahan kondisi IKP dari waktu ke waktu. Perbedaan karakteristik data secara spasial dan temporal ini dapat meningkatkan risiko terjadinya heterogenitas pada data. Oleh karena itu, pada penelitian ini dilakukan pemodelan faktor-faktor yang memengaruhi Indeks Ketahanan Pangan untuk setiap kabupaten/kota di Provinsi Sulawesi Selatan pada tahun 2021-2023 menggunakan *Geographically and Temporally Weighted Regression* dengan *Principal Component Analysis*. Penanganan masalah data pada penelitian ini diharapkan dapat memberikan hasil yang optimum, sehingga dapat memberikan panduan yang sesuai untuk meningkatkan IKP kepada pemerintah dan masyarakat umum.

1.2 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penentuan jarak antar lokasi pengamatan menggunakan jarak *euclidean*.
2. Penentuan model terbaik berdasarkan nilai R^2 tertinggi dan nilai *Root Mean Square Error* (RMSE) terendah.

1.3 Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh model GTWR-PCA pada Indeks Ketahanan Pangan di Sulawesi Selatan pada tahun 2021-2023.
2. Memperoleh model terbaik untuk Indeks Ketahanan Pangan di Sulawesi Selatan pada tahun 2021-2023.

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Memberikan informasi tentang pemodelan GTWR-PCA pada Indeks Ketahanan Pangan di Sulawesi Selatan pada tahun 2021-2023.
2. Memberikan informasi mengenai PCA sebagai salah satu metode analisis multivariat untuk mengatasi multikolinearitas pada data spasial.

1.4 Teori

1.4.1 Regresi Linier Berganda

Analisis linier regresi berganda adalah metode statistik yang digunakan untuk memodelkan hubungan antara variabel respon dan lebih dari satu variabel prediktor dengan cara mencocokkan persamaan linier dari data yang diamati. Analisis ini bertujuan untuk mengukur hubungan sebab-akibat dan menilai pengaruh dari perubahan dalam variabel respon kepada variabel prediktor (Arkes, 2023). Secara umum, pemodelan regresi linier dapat ditulis dalam Persamaan (1).

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \varepsilon_i \tag{1}$$

dengan $i = 1,2, \dots, n$ dan $k = 1,2, \dots, p$.
Keterangan:
y_i : Nilai variabel respon pada pengamatan ke-i
x_{ik} : Nilai variabel prediktor ke-k pada pengamatan ke-i
β_0 : Intersep model regresi
β_k : Parameter regresi variabel prediktor ke-k
ε_i : Variabel Galat
Bentuk matriks dari persamaan umum model regresi linier pada Persamaan (1) dapat ditulis dalam Persamaan (2) (Punggodewi dan Pratiwi, 2020).

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

$$\boldsymbol{Y}_{n\times 1} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2}$$

1.4.2 Pengujian Asumsi

a. Uji Multikolinearitas

Uji multikolinearitas merupakan uji yang dilakukan untuk mengetahui korelasi di antara faktor-faktor yang terdapat dalam regresi linier berganda. Data yang terdapat multikolinearitas dapat menghasilkan hasil pemodelan yang kurang bisa dipercaya. Hal ini dikarenakan multikolinearitas dapat menyebabkan selang kepercayaan yang lebih luas sehingga hasil prediksi yang dihasilkan kurang akurat. Pendeteksian multikolinearitas dapat dilakukan dengan mencari nilai *Variance Inflation Faktor* (VIF), yakni dengan menghitung banyaknya peningkatan variansi dari estimasi koefisien regresi jika variabel-variabel prediktor saling berkorelasi. Hipotesis yang diajukan adalah sebagai berikut:
H_0 : Tidak terdapat multikolinearitas
H_1 : Terdapat multikolinearitas
Nilai dari VIF dapat diperoleh menggunakan Persamaan (3).

$$VIF = \frac{1}{1 - R^2} \tag{3}$$

dengan R^2 adalah koefisien determinansi antar variabel prediktor. Jika diperoleh nilai $VIF > 10$, maka H_0 ditolak yang menandakan bahwa terdapat multikolinearitas di model regresi (Shrestha, 2020).

b. Uji Normalitas

Uji normalitas adalah uji yang sangat penting untuk memvalidasi model regresi, terutama dalam menilai distribusi residual. Berdasarkan beberapa penelitian yang membandingkan kekuatan dan efektivitas berbagai uji normalitas, uji *Shapiro-Wilk* secara konsisten menunjukkan kekuatan yang tinggi di berbagai jenis distribusi (Malá dkk., 2021). Pengujian normalitas dengan uji *Shapiro-Wilk* ditentukan dari korelasi antara data yang diberikan dan kecocokan angka normalnya. Hipotesis yang diajukan adalah sebagai berikut:

H_0 : Data berdistribusi normal
H_1 : Data tidak berdistribusi normal

Statistik uji dalam uji *Shapiro-Wilk* diperoleh menggunakan Persamaan (4).

$$W = \frac{\sum_{i=1}^{n/2} a_{n-i+1}(x_{n-i+1} - x_i)}{\sum_{i=1}^{n}(x_i - \bar{x})} \tag{4}$$

Keterangan:

W : Statistik uji *Shapiro-Wilk*
a_{n-i+1} : Koefisien uji *Shapiro-Wilk*
x_i : Nilai pada pengamatan ke-i

Jika diperoleh nilai $W > W_{tabel}$ atau $p - value < 0.05$, maka H_0 ditolak yang menandakan bahwa data tidak berdistribusi secara normal (Singgih dan Fauzan, 2022).

1.4.3 Pengujian Heterogenitas Spasial-Temporal

Heterogenitas spasial-temporal adalah kondisi ketika terdapat keragaman dalam data wilayah dan waktu, sehingga setiap wilayah dan waktu memiliki struktur dan parameter hubungan yang berbeda. Pengujian heterogenitas spasial-temporal dapat dilakukan dengan menggunakan uji *Breusch-Pagan* (Ania dkk., 2024). Adapun hipotesis yang diajukan pada uji *Breusch-Pagan* adalah sebagai berikut:

H_0 : $\sigma_1^2 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma$ (Terdapat homoskedastisitas)
H_1 : Terdapat satu $\sigma_i^2 \neq \sigma_j^2$ dengan $i \neq j$ (Terdapat heteroskedastisitas)

Perhitungan statistik uji dari uji *Breusch-Pagan* dapat dilakukan menggunakan Persamaan (5).

$$BP = \frac{1}{2}\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{Z}(\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{Z})^{-\boldsymbol{1}}\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{f} \sim \chi^2_{\alpha,p} \tag{5}$$

dengan

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

Keterangan:

BP : Statistik uji *Breusch-Pagan*
$\boldsymbol{Z}$: Matriks yang berisi variabel prediktor dari model regresi
e_i : Residual dari model regresi
σ^2 : Variansi dari residual

Pengambilan keputusan dalam uji *Breusch-Pagan* dilakukan berdasarkan nilai statistik uji BP atau $p - value$. Hipotesis H_0 ditolak apabila nilai $BP > \chi^2_{\alpha,p}$ atau nilai $p - value < \alpha$ yang berarti terjadi heterogenitas spasial dalam data (Agustina dkk., 2022).

1.4.4 *Principal Component Analysis*

Principal Component Analysis (PCA) merupakan metode reduksi dimensi yang memproyeksikan data ke dalam subruang dimensi yang lebih rendah (Kherif dan Latypova, 2019). PCA dapat digunakan untuk mengatasi masalah multikolinearitas, sebab data dari variabel awal yang sebelumnya saling berkorelasi satu dengan lainnya dapat ditransformasi menjadi satu set variabel baru yang saling bebas dan tidak berkorelasi lagi. Pada praktiknya, data yang ditemukan dapat memiliki satuan yang berbeda-beda, sehingga perlu dilakukan standardisasi data menggunakan Persamaan (6) (Delsen dkk., 2017).

$$z_{ki} = \frac{x_{ki} - \bar{x}_k}{\sigma_k} \tag{6}$$

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

Keterangan:

z_{ki} : Nilai variabel prediktor hasil standardisasi ke-k pada pengamatan ke-i

x_{ki} : Nilai variabel prediktor ke-k pada pengamatan ke-i

$\bar{x}_k$: Rata-rata variabel prediktor ke-k

σ_k : Standar deviasi variabel prediktor ke-k

Set nilai variabel baru tidak berkorelasi linier hasil PCA disebut *Principal Component* (PC) atau komponen utama (Kherif dan Latypova, 2019). Komponen utama merupakan beberapa kombinasi linier tertentu dari p variabel dengan vektor eigen sebagai koefisiennya. Vektor eigen diperoleh berdasarkan nilai eigen dari matriks varian-kovarian $\sum$ dan matriks korelasi $\boldsymbol{\rho}$. Perhitungan matriks varian-kovarian dan matriks korelasi secara berurut dapat dilihat pada Persamaan (7) dan Persamaan (8) (Yulianto dan Putriana, 2019).

$$\sum\nolimits_{\boldsymbol{p}\times\boldsymbol{p}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \tag{7}$$

dengan $i, j = 1,2, \dots p$ untuk σ_{ij}

$$\boldsymbol{\rho}_{\boldsymbol{p}\times\boldsymbol{p}} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix} \tag{8}$$

dengan $i, j = 1,2, \dots p$ untuk ρ_{ij}. Jika matriks korelasi memiliki nilai eigen dan vektor eigen berpasangan $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_2, e_2)$ dengan nilai eigen $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, maka komponen utama ke-k dapat didefinisikan dalam Persamaan (9) (Johnson dan Wichern, 2007).

$$C_k = e_{i1}z_1 + e_{i2}z_2 + \cdots e_{ip}z_p \tag{9}$$

Indikator pemilihan jumlah komponen utama diperoleh berdasarkan komponen utama dengan nilai eigen lebih dari satu dan jumlah komponen utama yang menjelaskan lebih dari 75% variansi kumulatif (Panda dkk., 2020). Perhitungan persentase variansi kumulatif yang dijelaskan oleh komponen utama ke-i dapat dilakukan menggunakan Persamaan (10).

$$\frac{\lambda_i}{\sum_{i=1}^{p_i} \lambda_i} \times 100\% \tag{10}$$

dengan $i = 1,2, \dots p$ (Johnson dan Wichern, 2007).

1.4.5 *Principal Component Regression*

Misalkan diberikan data dengan variabel respon y_i dan variabel prediktor x_{ik} yang dibentuk ke dalam model regresi linier pada Persamaan (1). Jika pada variabel prediktor x_{ik} terdapat multikolinearitas, maka terjadi pelanggaran asumsi multikolinearitas. Multikolinearitas disebabkan oleh adanya korelasi yang tinggi antara satu variabel prediktor dengan variabel prediktor lainnya. Pelanggaran asumsi multikolinearitas dapat diatasi dengan menggunakan PCA. Komponen utama C_{ik} hasil PCA yang diperoleh selanjutnya dimodelkan sebagai pengganti variabel prediktor x_{ik} pada model regresi linier, sehingga diperoleh model *Principal Component Regression* (PCR) pada Persamaan (11).

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k C_{ik} + \varepsilon_i \tag{11}$$

dengan

$$i = 1,2, \ldots, n\, ; k = 1,2, \ldots, p$$

Keterangan:

y_i : Nilai variabel respon pada pengamatan ke-i

C_{ik} : Nilai komponen utama ke-k pada pengamatan ke-i

β_0 : Intersep model regresi

β_k : Parameter regresi variabel prediktor ke-k

ε_i : Variabel Galat

1.4.6 *Geographically Weighted Regression*

Geographically Weighted Regression (GWR) merupakan metode regresi spasial yang diusulkan oleh Brundson dkk. (1996) untuk memungkinkan hubungan dalam model regresi bervariasi menurut lokasi. Perbedaan mendasar antara model regresi linier tradisional dan GWR terdapat pada koefisien regresinya. Koefisien regresi pada model regresi linier tradisional konstan di seluruh lokasi, sedangkan pada model GWR diestimasi secara lokal pada titik data yang direferensikan secara spasial. Pemodelan GWR dapat ditulis dalam Persamaan (12) (Wheeler, 2021).

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{12}$$

dengan

$$i = 1,2, \ldots, n\, ; k = 1,2, \ldots, p$$

Keterangan:

y_i : Nilai variabel respon pada lokasi ke-i

(u_i, v_i) : Titik koordinat lokasi ke-i

$\beta_0(u_i, v_i)$: Intersep model regresi pada lokasi ke-i

$\beta_k(u_i, v_i)$: Parameter regresi variabel prediktor ke-k pada lokasi ke-i

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

ε_i : Variabel Galat

Struktur spasial di pemodelan GWR dispesifikasikan melalui penambahan bobot ke dalam data. Pembobot dalam spasial berkaitan dengan asumsi autokorelasi, yakni asumsi bahwa suatu pengamatan dengan lokasi yang berdekatan akan memiliki hasil yang lebih serupa. Perhitungan fungsi pembobot dalam GWR dilakukan dengan menggunakan fungsi kernel. Beberapa fungsi kernel yang sering digunakan dalam pembobotan GWR di antaranya (Wheeler, 2021):

1. Fungsi Kernel *Gaussian*

$$w_{ij}^{S} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}^{S}}{\gamma_S}\right)^2\right) \tag{13}$$

2. Fungsi Kernel *Exponential*

$$w_{ij}^{S} = \exp\left(-\frac{d_{ij}^{S}}{\gamma_S}\right) \tag{14}$$

3. Fungsi Kernel *Bisquare*

$$w_{ij}^{S} = \begin{cases} \left[1-\left(\frac{d_{ij}^{S}}{\gamma_S}\right)^2\right]^2 & ;\ d_{ij}^{S} \le \gamma_S \\ 0 & ; d_{ij}^{S} > \gamma_S \end{cases} \tag{15}$$

4. Fungsi Kernel *Tricube*

$$w_{ij}^{S} = \begin{cases} \left[1-\left(\frac{d_{ij}^{S}}{\gamma_S}\right)^3\right]^3 & ;\ d_{ij}^{S} \le \gamma_S \\ 0 & ; d_{ij}^{S} > \gamma_S \end{cases} \tag{16}$$

dengan γ_S adalah *bandwidth* atau lebar jendela yaitu ukuran jarak fungsi pembobot yang dapat mengukur sejauh mana pengaruh suatu lokasi dengan lokasi lainnya. Adapun, d_{ij}^{S} adalah jarak antar lokasi ke-i dan lokasi ke-j berdasarkan titik koordinat yang diperoleh berdasarkan Persaman (17) (Wheeler, 2021).

$$d_{ij}^{S} = \sqrt{(u_i - v_i)^2 + \left(u_j - v_j\right)^2} \tag{17}$$

1.4.7 *Geographically and Temporally Weighted Regression*

Geographically and Temporally Weighted Regression (GTWR) adalah metode pengembangan dari GWR yang diperkenalkan oleh Huang dkk. (2010) untuk mempertimbangkan ketidakstasioneran spasial dan temporal secara bersamaan. Pemodelan GTWR dapat dilakukan menggunakan Persamaan (18).

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i \tag{18}$$

dengan

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

Keterangan:

y_i	: Nilai variabel respon lokasi ke-i pada waktu ke-i
(u_i, v_i, t_i)	: Titik koordinat lokasi ke-i pada waktu ke-i
$\beta_0(u_i, v_i, t_i)$	: Intersep model regresi pada lokasi ke-i pada waktu ke-i
$\beta_k(u_i, v_i, t_i)$	: Parameter regresi variabel prediktor ke-k pada lokasi dan waktu ke-i
x_{ik}	: Nilai variabel prediktor ke-k pada pengamatan ke-i
ε_i	: Variabel Galat

Berdasarkan metode *Weighted Least Square* (WLS), parameter regresi dari metode GTWR dapat diestimasi dengan meminimalkan Persamaan (19).

$$\hat{\beta}(u_i, v_i, t_i) = \sum_{i,j=1}^{n} w_{ij}^{ST} \left[y_i - \beta_0(u_i, v_i, t_i) - \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) x_{ik} \right]^2 \tag{19}$$

dengan w_{ij}^{ST} adalah pembobot spasial-temporal pada pengamatan ke-i. Pembobot w_{ij}^{ST} pada suatu amatan dapat diperoleh menggunakan matriks pembobot spasial-temporal yang dapat dilihat pada Persamaan (20).

$$\boldsymbol{W}_{n\times n} = diag(w_{i1}^{ST}, w_{i2}^{ST}, \dots, w_{in}^{ST}) \tag{20}$$

Estimasi parameter dari model GTWR dapat dihitung menggunakan matriks pada Persamaan (21) (Ma dkk., 2018).

$$\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)_{(p+1)\times 1} = [\boldsymbol{X^T W X}]^{-1} \boldsymbol{X^T W Y} \tag{21}$$

Pengujian signifikansi parameter secara parsial dari model GTWR dilakukan dengan menghitung nilai t_{hit} menggunakan Persamaan (22).

$$t_{hit} = \frac{\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)}{\hat{\sigma}\sqrt{m_{kk}}} \tag{22}$$

dengan m_{kk} merupakan elemen diagonal ke-k dari matriks $\boldsymbol{MM^T}$ yang dapat dilihat pada Persamaan (23).

$$\boldsymbol{M}_{(p+1)\times n} = [\boldsymbol{X^T W X}]^{-1} \boldsymbol{X^T W} \tag{23}$$

Adapun estimasi standar deviasi $\hat{\sigma}$ diperoleh menggunakan Persamaan (24).

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\delta_i}} \tag{24}$$

dengan

$$\delta_i = tr([(\boldsymbol{I} - \boldsymbol{XM})^T (\boldsymbol{I} - \boldsymbol{XM})]^i) \tag{25}$$

Hipotesis yang diajukan adalah sebagai berikut:
H_0 : $\hat{\beta}(u_i, v_i, t_i) = 0$ (Variabel prediktor x_{ik} tidak signifikan)
H_1 : $\hat{\beta}(u_i, v_i, t_i) \neq 0$ (Variabel prediktor x_{ik} signifikan)
Pengambilan keputusan dilakukan berdasarkan nilai statistik uji t_{hit}. Hipotesis H_0 ditolak apabila nilai $|t_{hit}| > t_{\frac{\alpha}{2},df}$ dengan $df = \frac{\delta_1^2}{\delta_2^2}$ yang berarti variabel prediktor x_{ik} pada lokasi ke-i signifikan (Djuraidah, 2020).

Tahap pertama dari estimasi parameter adalah mengkontruksi matriks pembobot spasial-temporal $\boldsymbol{W}_{n\times n}$ yang bergantung pada jarak spasial-temporal d_{ij}^{ST}. Jarak spasial-temporal menganut konsep yang sama dengan persamaan jarak 3D *Euclidean* yang dapat dilihat pada Persamaan (26) (Ma dkk., 2018).

$$d^2 = x^2 + y^2 + z^2 \tag{26}$$

Berdasarkan hal tersebut, fungsi dari jarak spasial-temporal juga bisa didefinisikan dengan langsung menambahkan jarak temporal ke dalam fungsinya. Oleh karena itu, jarak spasial-temporal dapat dikembangkan menjadi Persamaan (27).

$$d^{ST} = d^S \oplus d^T \tag{27}$$

Jarak spasial d^S dan jarak temporal d^T secara berturut-turut dapat didefinisikan pada Persamaan (28) dan Persamaan (29).

$$d_{ij}^S = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{28}$$

$$d_{ij}^T = \sqrt{\left(t_i - t_j\right)^2} \tag{29}$$

Huang dkk. (2010) mendefinisikan $\oplus$ sebagai fungsi untuk mengkombinasikan jarak spasial d^S dan jarak temporal d^T. Oleh karena itu, jarak spasial-temporal dapat didefinisikan pada Persamaan (30).

$$\left(d_{ij}^{ST}\right)^2 = \lambda\left[\left(u_i - u_j\right)^2 + (v_i - v_j)^2\right] + \mu\left(t_i - t_j\right)^2 \tag{30}$$

dengan λ dan μ merupakan bobot untuk menyeimbangkan efek yang berbeda, sebab jarak spasial dan temporal diukur dengan menggunakan satuan yang berbeda (Ma dkk., 2018).

Fungsi jarak spasial-temporal dapat digunakan untuk menghitung bobot spasial-temporal. Misalnya digunakan fungsi *Gaussian* dengan peluruhan jarak yang dapat dilihat pada Persamaan (31).

$$w_{ij}^{ST} = \exp\left[-\left(\frac{d_{ij}^{ST}}{\gamma_{ST}}\right)^2\right] \tag{31}$$

dengan γ_{ST}^2 merupakan *bandwidth* spasial-temporal. Fungsi pembobot spasial-temporal dapat diperoleh dengan mengkombinasikan jarak spasial-temporal dan fungsi *Gaussian* dengan peluruhan jarak. Perhitungan fungsi pembobot spasial-temporal pada Persamaan (31) dapat didefinisikan ke dalam Persamaan (32).

$$\begin{aligned} w_{ij}^{ST} &= \exp\left\{-\left[\frac{\lambda\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \mu\left(t_i - t_j\right)^2}{(\gamma_{ST})^2}\right]\right\} \\ &= \exp\left\{-\left[\frac{\left(u_i - u_j\right)^2 + (v_i - v_j)^2}{(\gamma_S)^2} + \frac{\left(t_i - t_j\right)^2}{(\gamma_T)^2}\right]\right\} \\ &= \exp\left\{-\left[\frac{\left(d_{ij}^S\right)^2}{(\gamma_S)^2} + \frac{\left(d_{ij}^T\right)^2}{(\gamma_T)^2}\right]\right\} \\ &= \exp\left[-\left(\frac{d_{ij}^S}{\gamma_S}\right)^2\right] \times \exp\left[-\left(\frac{d_{ij}^T}{\gamma_T}\right)^2\right] \\ &= w_{ij}^S \times w_{ij}^T \end{aligned} \tag{32}$$

Bandwidth spasial γ_S dan *bandwidth* temporal γ_T secara berturut-turut dapat didefinisikan pada Persamaan (33) dan Persamaan (34) (Ma dkk., 2018).

$$(\gamma_S)^2 = \frac{(\gamma_{ST})^2}{\lambda} \tag{33}$$

$$(\gamma_T)^2 = \frac{(\gamma_{ST})^2}{\mu} \tag{34}$$

Estimasi *bandwidth* kernel di model GTWR menggunakan *Cross Validation* (CV) dapat dilakukan menggunakan Persamaan (35).

$$CV(\gamma_{ST}) = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(\gamma_{ST})]^2 \tag{35}$$

dengan $\hat{y}_{\neq i}(\gamma_{ST})$ adalah nilai penaksir y_i saat pengamatan ke-i dihilangkan dari estimasi (Ma dkk., 2018). Misalkan parameter rasio τ dapat dirumuskan menjadi Persamaan (36), maka Persamaan (30) dapat ditulis kembali menjadi Persamaan (37).

$$\tau = \frac{\mu}{\lambda};\ \lambda \neq 0 \tag{36}$$

$$\frac{\left(d_{ij}^{ST}\right)^2}{\lambda} = \left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \tau\left(t_i - t_j\right)^2 \tag{37}$$

Parameter rasio τ bertujuan untuk memperbesar atau memperkecil rasio dari jarak temporal terhadap jarak spasial. Pemilihan parameter ini dapat menggunakan metode CV dengan kriteria minimum melalui iterasi inisialisasi nilai τ awal yang dapat dilihat pada Persamaan (38) (Ma dkk., 2018).

$$CV(\tau) = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(\tau)]^2 \tag{38}$$

1.4.8 *Geographically and Temporally Weighted Regression – Principal Component Analysis*

Jika asumsi multikolinearitas tidak terpenuhi pada pemodelan GTWR, maka dapat dilakukan PCA untuk mengatasi multikolinearitas tersebut. Merujuk pada pemodelan PCR pada Persamaan (11), penggunaan PCA model GTWR dapat dimodifikasi menjadi model GTWR-PCA pada Persamaan (39) (Fu dkk., 2022).

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) C_{ik} + \varepsilon_i \tag{39}$$

Keterangan:

y_i	: Nilai variabel respon lokasi ke-i pada waktu ke-i
(u_i, v_i, t_i)	: Titik koordinat lokasi ke-i pada waktu ke-i
$\beta_0(u_i, v_i, t_i)$	: Intersep model regresi pada lokasi ke-i pada waktu ke-i
$\beta_k(u_i, v_i, t_i)$	: Parameter regresi variabel prediktor ke-k pada lokasi dan waktu ke-i
C_{ik}	: Nilai komponen utama ke-k pada pengamatan ke-i
ε_i	: Variabel Galat

Jumlah kuadrat galat dari model GTWR-PCA diperlukan untuk menduga parameter dari model tersebut. Berdasarkan Persamaan (39), diperoleh kuadrat galat dari model GTWR-PCA pada Persamaan (40).

$$\varepsilon_i^2 = \left(y_i - \beta_0(u_i, v_i, t_i) - \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) C_{ik}\right)^2 \tag{40}$$

Pendugaan parameter pada model GTWR-PCA dapat dilakukan dengan metode *Weighted Least Square* (WLS). Estimasi parameter model GTWR-PCA dengan WLS dilakukan dengan meminimumkan jumlah kuadrat galat pada Persamaan (40) yang telah dikalikan dengan pembobot w_{ij}^{ST} sehingga diperoleh Persamaan (41).

$$\sum_{i=1}^{n} w_{ij}^{ST} \varepsilon_i^2 = \sum_{i=1}^{n} w_{ij}^{ST} \left(y_i - \beta_0(u_i, v_i, t_i) - \sum_{k=1}^{p} \beta_k(u_i, v_i, t_i) C_{ik}\right)^2 \tag{41}$$

Persamaan (41) selanjutnya diubah ke dalam bentuk matriks yang dapat dilihat pada Persamaan (42).

$$\boldsymbol{\varepsilon}_{n\times 1}^{\boldsymbol{T}} \boldsymbol{W}_{n\times n} \boldsymbol{\varepsilon}_{n\times 1} = (\boldsymbol{Y} - \boldsymbol{C\beta})^{\boldsymbol{T}} \boldsymbol{W} (\boldsymbol{Y} - \boldsymbol{C\beta}) \tag{42}$$

Persamaan (42) kemudian diturunkan terhadap matriks $\boldsymbol{\beta}_{(p+1)\times 1}$ dan hasilnya disamadenganKan dengan nol sebagai berikut:

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\beta}} (\boldsymbol{\varepsilon}^{\boldsymbol{T}} \boldsymbol{W} \boldsymbol{\varepsilon}) &= 0 \\ -2\boldsymbol{C}^T \boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{C\beta}) &= 0 \\ -2\boldsymbol{C}^T \boldsymbol{WY} + 2\boldsymbol{C}^T \boldsymbol{WC\beta}) &= 0 \\ 2\boldsymbol{C}^T \boldsymbol{WC\beta} &= 2\boldsymbol{C}^T \boldsymbol{WY} \\ \boldsymbol{C}^T \boldsymbol{WC\beta} &= \boldsymbol{C}^T \boldsymbol{WY} \\ \boldsymbol{\beta}_{(p+1)\times 1} &= (\boldsymbol{C}^T \boldsymbol{WC})^{-\boldsymbol{1}} \boldsymbol{C}^T \boldsymbol{WY} \end{aligned}$$

Diketahui $\boldsymbol{\beta}_{(p+1)\times 1} = \widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)$, sehingga diperoleh estimasi parameter GTWR-PCA untuk setiap amatan pada Persamaan (43) (Fu dkk., 2022).

$$\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)_{(p+1)\times 1} = [\boldsymbol{C}^{\boldsymbol{T}} \boldsymbol{WC}]^{-\boldsymbol{1}} \boldsymbol{C}^T \boldsymbol{WY} \tag{43}$$

Pengujian signifikansi parameter secara parsial dari model GTWR-PCA dilakukan dengan menghitung nilai t_{hit} menggunakan Persamaan (44).

$$t_{hit} = \frac{\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)}{\hat{\sigma}\sqrt{m_{kk}}} \tag{44}$$

dengan m_{kk} merupakan elemen diagonal ke-k dari matriks $\boldsymbol{MM^T}$ yang dapat dilihat pada Persamaan (45).

$$\boldsymbol{M}_{(p+1)\times n} = [\boldsymbol{C}^{\boldsymbol{T}} \boldsymbol{WC}]^{-\boldsymbol{1}} \boldsymbol{C}^{\boldsymbol{T}} \boldsymbol{W} \tag{45}$$

Hipotesis yang diajukan adalah sebagai berikut:
H_0 : $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) = 0$ (Komponen utama C_{ik} tidak signifikan)
H_1 : $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) \neq 0$ (Komponen utama C_{ik} signifikan)
Pengambilan keputusan dilakukan berdasarkan nilai statistik uji t_{hit}. Hipotesis H_0 ditolak apabila nilai $|t_{hit}| > t_{\frac{\alpha}{2},df}$ dengan $df = \frac{\delta_1^2}{\delta_2^2}$ yang berarti komponen utama C_{ik} pada lokasi ke-i signifikan (Fu dkk., 2022).

1.4.9 Pengukuran Kebaikan Model

Kebaikan model dari GTWR dan GTWR-PCA dapat diukur berdasarkan berbagai faktor, contohnya seperti yang digunakan oleh Fu dkk. (2022) yaitu nilai R^2 dan *Root Mean Square Error* (RMSE). Model terbaik diperoleh berdasarkan model yang memiliki nilai R^2 terbesar dan RMSE terkecil. Adapun, amatan dengan model terbaik diperoleh berdasarkan model amatan yang memiliki nilai R^2 lokal terbesar. Perhitungan nilai R^2 dan RMSE secara berturut-turut dapat dilihat pada Persamaan (46) dan Persamaan (47). Adapun, perhitungan nilai R^2 lokal dapat dilihat pada Persamaan (48) (Chicco dkk., 2021).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (\bar{y} - y_i)^2} \tag{46}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{47}$$

$$R_{lokal}^2 = 1 - \frac{\sum_{i=1}^{n} w_{ij}^{ST} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} w_{ij}^{ST} (\bar{y} - y_i)^2} \tag{48}$$

Keterangan:

y_i : Nilai variabel respon lokasi ke-i pada waktu ke-i
$\hat{y}_i$: Estimasi nilai variabel respon lokasi ke-i pada waktu ke-i
$\bar{y}$: Rata-rata nilai variabel respon
w_{ij}^{ST} : Pembobot spasial-temporal pada pengamatan ke-i

1.4.10 Indeks Ketahanan Pangan

Berdasarkan Undang-Undang Pangan Nomor 18 Tahun 2012, ketahanan pangan adalah kondisi terpenuhinya kebutuhan pangan, baik di tingkat negara maupun individu, yang ditandai dengan ketersediaan pangan yang cukup dalam jumlah dan kualitas. Pangan yang tercakup dalam kondisi tersebut harus aman, beragam, bergizi, terdistribusi secara merata, mudah diakses, serta sesuai dengan keyakinan agama, nilai, dan budaya masyarakat (PPK-LIPI, 2012). Ketahanan pangan merupakan aspek penting bagi stabilitas nasional, meliputi ketersediaan, akses, dan mutu pangan (Raihan dkk., 2020).

Tingkat ketahanan pangan suatu wilayah beserta faktor-faktor pendukungnya diukur menggunakan sistem penilaian yang disebut Indeks Ketahanan Pangan (IKP). IKP mengacu pada definisi ketahanan pangan serta subsistem yang menyusunnya. Sembilan indikator yang digunakan dalam perhitungan IKP berasal dari tiga aspek utama ketahanan pangan, yaitu ketersediaan, keterjangkauan, dan pemanfaatan pangan. Indikator-indikator tersebut memiliki bobot yang berbeda-beda yang dapat dilihat pada Tabel 1 (Badan Pangan Nasional, 2023).

Tabel 1. Indikator Ketahanan Pangan

No	Indikator	Bobot
Aspek Ketersediaan Pangan		
1.	Rasio konsumsi normatif terhadap produksi bersih beras, jagung, ubi jalar, ubi kayu, dan sagu, serta stok beras pemerintah daerah	0.30
	Sub Total	0.30

No	Indikator	Bobot
Aspek Keterjangkauan Pangan		
2.	Persentase penduduk di bawah garis kemiskinan	0.15
3.	Persentase rumah tangga dengan proporsi pengeluaran untuk pangan lebih dari 65% terhadap total pengeluaran	0.075
4.	Persentase rumah tangga tanpa akses listrik	0.075
	Sub Total	0.30
Aspek Pemanfaatan Pangan		
5.	Rata-rata lama sekolah Perempuan berusia di atas 15 tahun	0.05
6.	Persentase rumah tangga tanpa akses ke air bersih	0.15
7.	Rasio jumlah penduduk per tenaga Kesehatan terhadap tingkat kepadatan penduduk	0.05
8.	Persentase balita *stunting*	0.05
9.	Angka harapan hidup saat lahir	0.10
	Sub Total	0.40

Pemerintah Indonesia mengarahkan kebijakan ketahanan pangan untuk mewujudkan kemandirian pangan dari tingkat rumah tangga hingga nasional, dengan strategi yang diharapkan dapat meningkatkan daya saing di masa depan (Indah P. dan Setyaningsih, 2020). Beberapa faktor yang dapat memengaruhi kondisi IKP di suatu wilayah diantaranya:

a. Indeks Pembangunan Manusia (IPM).
   IPM adalah nilai yang menunjukkan seberapa jauh negara atau wilayah itu telah mencapai sasaran yang ditentukan yaitu angka harapan hidup, pendidikan dasar, dan tingkat pengeluaran dan konsumsi yang telah mencapai standar hidup layak (Azizah dan Ratnasari, 2023).
b. Umur Harapan Hidup (UHH) saat lahir.
   UHH saat lahir adalah rata-rata perkiraan banyak tahun yang dapat ditempuh seseorang selama hidup. UHH diartikan sebagai umur yang mungkin dicapai seseorang yang lahir pada tahun tertentu (Yulianti dan Ratnasari, 2020).
c. Pengeluaran per kapita disesuaikan.
   Pengeluaran per kapita disesuaikan merupakan penyesuaian indeks harga konsumen dan penurunan utilitas marginal. Pengeluaran perkapita dapat digambarkan dari tingkat daya beli masyarakat dan melihat progres pembangunan manusia di suatu wilayah (Punggodewi dan Pratiwi, 2020).
d. Penduduk yang hidup di bawah garis kemiskinan.
   Penduduk yang hidup di bawah garis kemiskinan tidak memiliki daya beli yang memadai untuk memenuhi kebutuhan dasar hidupnya sehingga akan mempengaruhi ketahanan pangan (Muttaqin dkk., 2022).
e. Rumah tangga dengan pengeluaran pangan lebih dari total pengeluaran.
   Rumah tangga dengan proporsi pengeluaran untuk pangan lebih dari 65% terhadap total pengeluaran merupakan indikator proksi dari ketahanan pangan rumah tangga. Teori Engel menyatakan semakin tinggi tingkat pendapatan maka

persentase pengeluaran rumah tangga untuk konsumsi pangan akan semakin menurun (Prasetya, 2024).

f. Rumah tangga tanpa akses ke air bersih.

Rumah tangga tanpa akses ke air bersih memegang peranan yang sangat penting untuk pencapaian ketahanan pangan. Peningkatan akses terhadap fasilitas sanitasi dan air layak minum sangat penting untuk mengurangi masalah kesehatan sehingga dapat memperbaiki gizi (Zulaika dkk., 2024).

BAB II
METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari publikasi “Peta Ketahanan dan Kerentanan Pangan (FSVA)” yang diterbitkan oleh Badan Pangan Nasional dalam kurun waktu 2021-2023 serta data tabel statistik pada laman Badan Pusat Statistik Sulawesi Selatan dalam kurun waktu 2021-2023. Unit pengamatan yang digunakan adalah 24 kabupaten/kota yang berada di Provinsi Sulawesi Selatan. Data yang digunakan dalam penelitian ini dapat dilihat pada Lampiran 1.

2.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas variabel respon Indeks Ketahanan Pangan (IKP) dan enam variabel prediktor yang diduga memengaruhi IKP di setiap kabupaten/kota di Provinsi Sulawesi Selatan. Variabel yang digunakan dapat dilihat pada Tabel 2.

Tabel 2. Variabel Penelitian

Variabel	Keterangan	Satuan
y	Indeks Ketahanan Pangan (IKP)	Indeks
x_1	Indeks Pembangunan Manusia (IPM)	Indeks
x_2	Umur Harapan Hidup Saat Lahir (UHH)	Tahun
x_3	Pengeluaran per Kapita Disesuaikan	Ribu Rupiah/ Orang/Tahun
x_4	Penduduk yang Hidup di Bawah Garis Kemiskinan	Persentase (%)
x_5	Rumah Tangga dengan Proporsi Pengeluaran untuk Pangan Lebih Dari 65% terhadap Total Pengeluaran	Persentase (%)
x_6	Rumah Tangga Tanpa Akses ke Air Bersih	Persentase (%)

2.3 Metode Analisis

Penelitian ini menggunakan pendekatan analisis regresi spasial *Geographically and Temporally Weighted Regression* (GTWR) dengan analisis multivariat *Principal Component Analysis* untuk mengatasi multikolinearitas pada data. Pengolahan data dilakukan dengan bantuan *software* R-Studio. Tahapan analisis data yang dilakukan dalam penelitian ini yaitu sebagai berikut:

1. Melakukan eksplorasi data.
2. Mengidentifikasi multikolinearitas dengan menghitung nilai *Variance Inflaction Factor* (VIF) berdasarkan persamaan berikut:

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

Hipotesis yang diajukan pada uji multikolinearitas dengan nilai VIF adalah sebagai berikut:

H_0 : Tidak terdapat multikolinearitas

H_1 : Terdapat multikolinearitas

Jika diperoleh nilai $VIF > 10$, maka H_0 ditolak yang menandakan bahwa terdapat multikolinearitas pada data.

3. Mengidentifikasi kenormalan dengan Uji *Shapiro Wilk* menggunakan persamaan berikut:

$$W = \frac{\sum_{i=1}^{n/2} a_{n-i+1}(x_{n-i+1} - x_i)}{\sum_{i=1}^{n}(x_i - \bar{x})}$$

Hipotesis yang diajukan pada uji multikolinearitas dengan nilai VIF adalah sebagai berikut:

H_0 : Data berdistribusi normal

H_1 : Data tidak berdistribusi normal

Jika diperoleh nilai $W > W_{tabel}$ atau $p - value < 0.05$, maka H_0 ditolak yang menandakan bahwa data tidak berdistribusi secara normal.

4. Mengidentifikasi efek heterogenitas spasial-temporal dengan uji *Breusch-Pagan* menggunakan persamaan berikut:

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

dengan

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

Hipotesis yang diajukan pada uji *Breusch-Pagan* adalah sebagai berikut:

H_0 : $\sigma_1^2 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma$ (Terdapat homoskedastisitas)

H_1 : Terdapat satu $\sigma_i^2 \neq \sigma_j^2$ dengan $i \neq j$ (Terdapat heteroskedastisitas)

Pengambilan keputusan dalam uji *Breusch-Pagan* dilakukan berdasarkan nilai statistik uji BP atau $p - value$. Hipotesis H_0 ditolak apabila nilai $BP > \chi^2_{\alpha,p}$ atau nilai $p - value < \alpha$ yang berarti terjadi heterogenitas spasial dalam data.

5. Melakukan standardisasi data menggunakan persamaan berikut:

$$z_{ki} = \frac{x_{ki} - \bar{x}_k}{\sigma_k}$$

6. Membentuk matriks korelasi $\boldsymbol{\rho}$ menggunakan persamaan berikut:

$$\boldsymbol{\rho}_{p\times p} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix}$$

7. Membentuk komponen utama menggunakan persamaan berikut:

$$C_k = e_{i1}z_1 + e_{i2}z_2 + \cdots e_{ip}z_p$$

8. Menentukan jumlah komponen utama berdasarkan nilai eigen dan persentase variansi kumulatif yang diperoleh dengan persamaan berikut:

$$\frac{\lambda_i}{\sum_{i=1}^{p_i}\lambda_i} \times 100\%$$

9. Menganalisis model GTWR dan GTWR-PCA dengan langkah-langkah sebagai berikut:
   a. Menentukan jarak spasial-temporal untuk model GTWR dan GTWR-PCA menggunakan persamaan berikut:
   $$\left(d_{ij}^{ST}\right)^2 = \lambda\left[\left(u_i - u_j\right)^2 + (v_i - v_j)^2\right] + \mu\left(t_i - t_j\right)^2$$
   b. Menentukan pembobot optimum untuk model GTWR dan GTWR-PCA menggunakan persamaan berikut:
   $$w_{ij}(u_i, v_i, t_i) = w_{ij}(u_i, v_i) \times w_{ij}(t_i)$$
   c. Menentukan parameter model GTWR dan GTWR-PCA optimum dengan meminimalkan *Cross Validation* (CV).
      1) Memperoleh *bandwidth* spasial-temporal optimum dengan meminimalkan CV pada persamaan berikut:
      $$CV(\gamma_{ST}) = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(\gamma_{ST})]^2$$
      2) Memperoleh parameter rasio τ optimum dengan meminimalkan CV pada persamaan berikut:
      $$CV(\tau) = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(\tau)]^2$$
      3) Memperoleh parameter bobot spasial λ optimum dengan meminimalkan CV pada persamaan berikut:
      $$CV(\lambda) = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(\lambda)]^2$$
      4) Memperoleh parameter bobot temporal μ optimum dengan meminimalkan CV pada persamaan berikut:
      $$CV(\mu) = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(\mu)]^2$$
   d. Memperoleh estimasi parameter model GTWR menggunakan persamaan berikut:
   $$\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)_{(p+1)\times 1} = [\boldsymbol{X^T W X}]^{-\mathbf{1}}\boldsymbol{X^T W Y}$$
   Adapun estimasi parameter model GTWR-PCA diperoleh menggunakan persamaan berikut:
   $$\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)_{(p+1)\times 1} = [\boldsymbol{C^T W C}]^{-\mathbf{1}}\boldsymbol{C}^T\boldsymbol{WY}$$
   e. Menghitung estimasi parameter model GTWR dan GTWR-PCA
   f. Melakukan pengujian parsial untuk setiap parameter model GTWR dan GTWR-PCA menggunakan persamaan berikut:
   $$t_{hit} = \frac{\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i)}{\hat{\sigma}\sqrt{m_{kk}}}$$
   Hipotesis yang diajukan untuk model GTWR adalah sebagai berikut:
   H_0 : $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) = 0$ (Variabel x_{ki} tidak signifikan)
   H_1 : $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) \neq 0$ (Variabel x_{ki} signifikan)

Adapun hipotesis pada model GTWR-PCA adalah sebagai berikut:

H_0 : $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) = 0$ (Komponen utama C_{ki} tidak signifikan)

H_1 : $\widehat{\boldsymbol{\beta}}(u_i, v_i, t_i) \neq 0$ (Komponen utama C_{ki} signifikan)

Pengambilan keputusan dilakukan berdasarkan nilai statistik uji t_{hit}. Hipotesis H_0 ditolak apabila nilai $|t_{hit}| > t_{\frac{\alpha}{2},df}$ dengan $df = \frac{\delta_1^2}{\delta_2^2}$.

10. Pemilihan model terbaik berdasarkan nilai R^2 dan nilai *Root Mean Square Error* (RMSE). Perhitungan nilai R^2 dan RMSE global pada model GTWR secara berturut-turut menggunakan persamaan berikut:

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

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

Adapun perhitungan nilai R^2 lokal pada model GTWR dan GTWR-PCA secara berturut-turut menggunakan persamaan berikut:

$$R_{lokal}^2 = 1 - \frac{\sum_{i=1}^{n} w_{ij}(u_i, v_i, t_i)(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} w_{ij}(u_i, v_i, t_i)(\bar{y} - y_i)^2}$$

11. Melakukan interpretasi model GTWR-PCA.