

LAPORAN SKRIPSI

**PEMODELAN *GEOGRAPHICALLY WEIGHTED REGRESSION*
PRINCIPAL COMPONENT ANALYSIS (GWR-PCA) DENGAN FUNGSI
PEMBOBOT *FIXED GAUSSIAN KERNEL***

(Studi Kasus: Faktor-faktor yang Memengaruhi Distribusi Persentase PDRB
Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024)



REDELLA REFFA HERDIANTI

NIM. 22106010073

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2026

SKRIPSI

**PEMODELAN *GEOGRAPHICALLY WEIGHTED REGRESSION*
PRINCIPAL COMPONENT ANALYSIS (GWR-PCA) DENGAN FUNGSI
PEMBOBOT *FIXED GAUSSIAN KERNEL***

(Studi Kasus: Faktor-faktor yang Memengaruhi Distribusi Persentase PDRB
Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024)

Diajukan untuk memenuhi salah satu syarat memperoleh derajat
Sarjana Matematika



REDELLA REFFA HERDIANTI

NIM. 22106010073

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2026

HALAMAN PENGESAHAN



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1435/Un.02/DST/PP.00.9/07/2026

Tugas Akhir dengan judul : Pemodelan Geographically Weighted Regression Principal Component Analysis (GWR-PCA) dengan Fungsi Pembobot Fixed Gaussian Kernel (Studi Kasus: Faktor-faktor yang Memengaruhi Distribusi Persentase PDRB Menurut Kabupaten/kota di Provinsi Sulawesi Selatan Tahun 2024)

yang dipersiapkan dan disusun oleh:

Nama : REDELLA REFFA HERDIANTI
Nomor Induk Mahasiswa : 22106010073
Telah diujikan pada : Senin, 22 Juni 2026
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Valid ID: 6a4c83a9e6275

Ketua Sidang

Dr. Mohammad Farhan Qudratullah, S.Si., M.Si
SIGNED



Valid ID: 6a3ce845b89a6

Penguji I

Arya Fendha Ibnu Shina, M.Si.
SIGNED



Valid ID: 6a4c7692615e1

Penguji II

Malahayati, S.Si., M.Sc
SIGNED



Valid ID: 6a4c8862f0308

Yogyakarta, 22 Juni 2026
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

SURAT PERNYATAAN KEASLIAN

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Redella Reffa Herdianti
NIM : 22106010073
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 6 Juni 2026



METERAI
TEMPEL
10000
06F4FAOX05U136020

Redella Reffa Herdianti

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSETUJUAN



Universitas Islam Negeri Sunan Kalijaga



FM-UINSK-BM-05-03/R0

SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir
Lamp :

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Redella Reffa Herdianti
NIM : 22106010073
Judul Skripsi : *Pemodelan Geographically Weighted Regression Principal Component Analysis (GWRPCA) dengan Fungsi Pembobot Fixed Gaussian Kernel*
(Studi Kasus: Faktor-faktor yang memengaruhi Distribusi Persentase PDRB Menurut Kabupaten/kota di Provinsi Sulawesi Selatan Tahun 2024)

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 12 Juni 2026

Pembimbing

Dr. Mohammad Farhan Qudratullah, S.Si., M.Si.

NIP. 19790922 200801 1 011

HALAMAN MOTTO

"Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan."

(QS. Al-Insyirah: 5–6)

"Live life to the fullest, and trust Allah with the rest."



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Tugas akhir ini penulis persembahkan kepada:

Allah SWT, atas segala rahmat, pertolongan, kemudahan, dan kesempatan yang diberikan sehingga penulis dapat menyelesaikan tugas akhir ini.

Kedua orang tua penulis, Ibu tercinta yang senantiasa memberikan kasih sayang, doa, dukungan, dan pengorbanan yang tidak terhingga, serta Bapak yang telah berpulang ke rahmatullah. Meskipun tidak dapat menyaksikan pencapaian ini secara langsung, kenangan, doa, dan nilai-nilai kehidupan yang beliau tinggalkan akan selalu menjadi bagian dari perjalanan penulis.

Para dosen, guru, sahabat, dan semua pihak yang telah memberikan ilmu, bantuan, doa, serta dukungan selama proses perkuliahan dan penyusunan tugas akhir ini.

Diri penulis sendiri, yang telah berusaha, bertahan, dan terus melangkah dalam menghadapi berbagai tantangan hingga mampu menyelesaikan tugas akhir ini.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PRAKATA

Puji dan syukur penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah, serta karunia-Nya sehingga penulis diberikan kekuatan, kesehatan, kemudahan, dan kelancaran dalam menyelesaikan skripsi yang berjudul “*Geographically Weighted Regression Prinicipal Component Analysis (GWR-PCA) dengan Fungsi Pembobot Fixed Gaussian Kernel (Studi Kasus: Faktor-faktor yang Memengaruhi Distribusi Persentase PRDB Menurut Kabupaten/kota di Provinsis Sulawesi Selatan Tahun 2024)*” sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika (S.Mat) pada Program Studi Matematika, Fakultas Sains dan Teknologi, UIN Sunan Kalijaga Yogyakarta.

Shalawat serta salam senantiasa tercurah kepada Nabi Muhammad SAW yang telah membawa umat manusia dari zaman kegelapan menuju zaman yang penuh dengan ilmu pengetahuan dan peradaban. Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, dukungan, serta doa dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati penulis menyampaikan terima kasih kepada:

1. Allah SWT yang senantiasa memberikan petunjuk, pertolongan, kekuatan, dan kemudahan kepada penulis dalam setiap proses penyusunan skripsi ini.
2. Bapak Prof. Noorhaidi, S.Ag., M.A., M.Phil., Ph.D., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
3. Ibu Prof. Dr. Dra. Hj. Khurul Wardati, M.Si. selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Ibu Dr. Epha Diana Supandi, S.Si., M.Sc., selaku Ketua Program Studi Matematika UIN Sunan Kalijaga Yogyakarta.
5. Bapak Dr. Mohammad Farhan Qudratullah, S.Si., M.Si., selaku Dosen Pembimbing Skripsi yang telah meluangkan waktu, memberikan bimbingan, arahan, saran, serta motivasi kepada penulis selama proses penyusunan skripsi.

6. Seluruh dosen dan tenaga kependidikan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta yang telah memberikan ilmu, pengalaman, dan pelayanan selama masa studi penulis.
7. Kedua orang tua penulis, khususnya Ibu tercinta yang senantiasa memberikan doa, kasih sayang, dukungan, dan pengorbanan yang tiada henti. Serta kepada almarhum Bapak yang telah berpulang ke rahmatullah di tengah proses penyusunan skripsi ini, terima kasih atas segala cinta dan pelajaran hidup yang telah diberikan serta menjadi bagian dari perjalanan penulis hingga mampu mencapai tahap ini.
8. Yayasan Meriah Bintang yang telah mendukung dan mendampingi perjalanan pendidikan penulis sejak sekolah dasar hingga perguruan tinggi.
9. Sahabat dan teman-teman terdekat penulis yang senantiasa memberikan dukungan, semangat, bantuan, serta menemani berbagai proses selama masa perkuliahan.
10. Seluruh pihak yang tidak dapat disebutkan satu per satu yang telah memberikan doa, dukungan, bantuan, dan motivasi kepada penulis.
11. *Last but not least, I wanna thank me. I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, and I wanna thank me for never quitting.*

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi pengembangan penelitian selanjutnya. Penulis berharap skripsi ini dapat memberikan manfaat serta menambah wawasan bagi pembaca dan menjadi salah satu kontribusi dalam pengembangan ilmu pengetahuan.

Yogyakarta, 12 Juni 2026

Penulis

Redella Reffa Herdianti

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
SURAT PERNYATAAN KEASLIAN	iv
HALAMAN PERSETUJUAN	v
HALAMAN MOTTO	vi
HALAMAN PERSEMBAHAN	vii
PRAKATA	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
DAFTAR SIMBOL	xvi
INTISARI	xviii
ABSTRACT	xix
BAB I PENDAHULUAN	20
1.1 Latar Belakang	20
1.1 Rumusan Masalah	25
1.2 Tujuan Penelitian	25
1.3 Manfaat Penelitian	25
1.4 Batasan Masalah	26
1.5 Tinjauan Pustaka	27
1.6 Sistematika Penulisan	30
BAB II LANDASAN TEORI	32
2.1 Variabel Random	32
2.2 Ekspektasi	34
2.3 Variansi	35
2.4 Kovariansi	36
2.5 Koefisien Korelasi.....	36
2.6 Matriks	38
2.7 Analisis Regresi Linear Berganda.....	45

2.8	Data Spasial.....	58
2.9	Distribusi Persentase PDRB.....	59
2.10	Indikator Pendidikan	60
2.11	Indikator Ekonomi	61
2.12	Indikator Kesehatan	62
2.13	Indikator Ketenagakerjaan	62
2.14	Indikator Ketimpangan dan Kependudukan.....	64
BAB III METODE PENELITIAN		65
3.1	Data dan Jenis Penelitian	65
3.2	Objek Penelitian	65
3.3	Metode Pengumpulan Data	66
3.4	Variabel Penelitian.....	66
3.5	Metode Analisis Data	68
3.6	<i>Flowchart</i> Penelitian	71
BAB IV PEMBAHASAN.....		73
4.1	Asumsi <i>Principal Component Analysis</i> (PCA).....	73
4.2	Standardisasi Data.....	77
4.3	<i>Principal Component Analysis</i> (PCA)	77
4.4	Penentuan Komponen Utama.....	78
4.5	Heterogenitas Spasial.....	79
4.6	<i>Geographically Weighted Regression</i> (GWR).....	81
4.7	Koordinat Spasial	82
4.8	<i>Bandwidth</i> Optimum model GWR.....	83
4.9	Fungsi Pembobot.....	84
4.10	Pemodelan GWRPCA.....	87
4.11	Estimasi Parameter Model GWRPCA	88
4.12	Uji Kesuaian Model GWRPCA	91
4.13	Pengujian Parameter Model GWRPCA Secara Serentak	93
4.14	Pengujian Parameter Model GWRPCA Secara Parsial	94
4.15	Pemilihan Model Terbaik.....	95
BAB V STUDI KASUS		98
5.1	Analisis Statistika Deskriptif	98
5.2	Uji Asumsi PCA.....	109
5.2.1	Pendeteksian Multikolinearitas	109

5.2.2	Uji Kelayakan Data	111
5.3	Pemebentukan Komponen Utama.....	113
5.4	Analisis Regresi Linear Berganda Berbasis PCA (RPCA).....	118
5.5	Estimasi Parameter Model RPCA.....	119
5.6	Uji Hipotesis Model RPCA.....	120
5.6.1	Uji Simultan (Uji F)	120
5.6.2	Uji Parsial (Uji t).....	121
5.7	Uji Asumsi Klasik Model RPCA	123
5.7.1	Uji Normalitas.....	123
5.7.2	Uji Multikolinearitas	124
5.7.3	Uji Autokorelasi	125
5.7.4	Uji Heterokedastisitas	126
5.8	Pemodelan Menggunakan GWRPCA.....	128
5.8.1	Penentuan <i>Bandwidth</i> Optimum	128
5.8.2	Jarak <i>Euclidean</i>	130
5.8.3	Matriks Pembobot dengan Fungsi Pembobot <i>Fixed Gaussian Kernel</i>	132
5.8.4	Estimasi Parameter Model GWRPCA Setiap Kabupaten/Kota..	135
5.8.5	Pengujian Kesesuaian Model GWRPCA	137
5.9	Pemilihan Model Terbaik.....	144
BAB VI	PENUTUP	145
6.1	Kesimpulan	145
6.2	Saran.....	146
DAFTAR PUSTAKA		148
LAMPIRAN		152

DAFTAR TABEL

Tabel 1. 1 Persamaan dan Perbedaan dengan Penelitian Terdahulu	29
Tabel 2. 1 Tingkat Hubungan Dua Variabel	37
Tabel 2. 2 ANOVA Regresi Linear	50
Tabel 3. 1 Definisi Operasional Variabel	66
Tabel 4. 1 Pengujian KMO	74
Tabel 4. 2 Pengujian MSA	75
Tabel 5. 1 Statistika Deskriptif	98
Tabel 5. 2 Hasil Uji VIF	109
Tabel 5. 3 Hasil Uji KMO	111
Tabel 5. 4 Hasil Uji MSA	112
Tabel 5. 5 Nilai Eigen dan Vektor Eigen	113
Tabel 5. 6 Nilai Eigen dan Proporsi Variansi	115
Tabel 5. 7 Nilai Komponen Utama PC1 dan PC2 per Kabupaten/Kota	116
Tabel 5. 8 Nilai Vektor Eigen Komponen Utama PC1 dan PC2	117
Tabel 5. 9 Hasil Estimasi Parameter Model RPCA	119
Tabel 5. 10 Hasil Uji Simultan Model RPCA	121
Tabel 5. 11 Hasil Uji Parsial Model RPCA	122
Tabel 5. 12 Hasil Uji Normalitas	124
Tabel 5. 13 Hasil Uji Multikolinearitas	125
Tabel 5. 14 Hasil Uji Autokorelasi	126
Tabel 5. 15 Hasil Uji Heteroskedastisitas	128
Tabel 5. 16 Perbandingan Bandwidth dan Nilai CV	129
Tabel 5. 17 Hasil Iterasi Pencarian Bandwidth Optimum	129
Tabel 5. 18 Jarak Euclidean Kota Makassar dengan Kabupaten/Kota di Sulawesi Selatan	131
Tabel 5. 19 Matriks Pembobot pada Kota Makassar	134
Tabel 5. 20 Estimasi Parameter Model GWRPCA	135
Tabel 5. 21 Hasil F Test GWRPCA	137
Tabel 5. 22 Hasil Uji Parsial PC1 Model GWRPCA	139
Tabel 5. 23 Hasil Uji Parsial PC2 Model GWRPCA	140
Tabel 5. 24 Perbandingan Model RPCA dan GWRPCA	144

DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir Tahapan Analisis Data	72
Gambar 3. 1 Diagram Alir Tahapan Analisis Data	72
Gambar 4. 1 Diagram Alir Estimasi Parameter GWRPCA.....	91
Gambar 5. 1 Peta Persebaran Distribusi Persentase PDRB Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	99
Gambar 5. 2 Peta Persebaran Rata-rata Lama Sekolah Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	100
Gambar 5. 3 Peta Persebaran Harapan Lama Sekolah Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	101
Gambar 5. 4 Peta Persebaran Pengeluaran per Kapita Disesuaikan Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	102
Gambar 5. 5 Peta Persebaran Persentase Penduduk Miskin Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	103
Gambar 5. 6 Peta Persebaran Umur Harapan Hidup Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	104
Gambar 5. 7 Peta Persebaran Tingkat Pengangguran Terbuka Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	105
Gambar 5. 8 Peta Persebaran Jumlah Angkatan Kerja Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	106
Gambar 5. 9 Peta Persebaran Gini Ratio Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024	107
Gambar 5. 10 Peta Persebaran Jumlah Penduduk Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024.....	108
Gambar 5. 11 Plot Matriks Korelasi Antar Variabel Independen	110
Gambar 5. 12 Scree Plot Nilai Eigen.....	115
Gambar 5. 13 Peta Persebaran Kemampuan Model GWRPCA dalam Menjelaskan Distribusi Persentase PDRB	141
Gambar 5. 14 Peta Sebaran Koefisien Lokal Komponen Utama Pertama (PC1) pada Model GWRPCA Distribusi Persentase PDRB di Provinsi Sulawesi Selatan Tahun 2024	142
Gambar 5. 15 Peta Sebaran Koefisien Lokal Komponen Utama Pertama (PC2) pada Model GWRPCA Distribusi Persentase PDRB di Provinsi Sulawesi Selatan Tahun 2024	143

DAFTAR LAMPIRAN

Lampiran 1 Data Variabel Penelitian	152
Lampiran 2 Garis Bujur dan Lintang Selatan di Provinsi Sulawesi Selatan	153
Lampiran 3 Statistika Deskriptif R Studio	154
Lampiran 4 Uji Multikolinearitas R Studio.....	155
Lampiran 5 Uji KMO dan MSA	155
Lampiran 6 Nilai Eigen dan Vektor Eigen.....	156
Lampiran 7 Proporsi Variansi	156
Lampiran 8 Skor PC1 dan PC2 pada Setiap Kabupaten/kota	157
Lampiran 9 Estimasi Parameter Model RPCA	157
Lampiran 10 Uji Serentak dan Uji Parsial Model RPCA.....	158
Lampiran 11 Uji Normalitas Model RPCA.....	158
Lampiran 12 Uji Multikolinearitas Model RPCA.....	158
Lampiran 13 Uji Autokorelasi Model RPCA.....	159
Lampiran 14 Uji Heterogenitas Spasial Model RPCA.....	159
Lampiran 15 Pemilihan Fungsi Pembobot Terbaik.....	159
Lampiran 16 Jarak Euclidean Antar Lokasi	160
Lampiran 17 Matriks Pembobot.....	163
Lampiran 18 Estimasi Parameter Model GWRPCA	167
Lampiran 19 Uji Serentak GWRPCA	168
Lampiran 20 Uji Parsial GWRPCA	170
Lampiran 21 Peta Persebaran Komponen Utama yang Signifikan	173
Lampiran 22 Pemilihan Model Terbaik	173

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR SIMBOL

$A_{m \times n}$	: matriks dengan baris sebanyak m dan kolom sebanyak n
a_{ij}	: entri matriks baris ke- i dan kolom ke- j
I	: matriks identitas
$tr(A)$	: nilai matriks A
A^{-1}	: invers matriks A
Y_i	: variabel dependen
X_i	: variabel independen
β_0	: intersep
β_1	: koefisien slop (kemiringan), mengukur besarnya perubahan nilai Y_i yang disebabkan bertambahnya nilai X_i
ε_i	: nilai error (galat)
$\hat{Y}$	: nilai estimasi variabel dependen
b_0	: nilai estimasi parameter β_0
b_1	: nilai estimasi parameter β_1
k	: banyaknya variabel independen
Y	: matriks variabel dependen
X	: matriks variabel independen
β	: matriks parameter β
ε	: vektor galat berukuran $n \times 1$
β_j	: nilai koefisien variabel independen ke- j
X_{ij}	: nilai variabel independen ke- j pada pengamatan ke- i
i	: indeks pengamatan ($i = 1, 2, \dots, n$)
j	: indeks variabel independen ($j = 1, 2, \dots, k$)
n	: banyaknya pengamatan
$\hat{\beta}_j$	: penduga parameter regresi ke- j
$SE(\hat{\beta}_j)$	: standar <i>error</i> penduga parameter

C_{jj}	: elemen diagonal ke- j dari matriks $(X^T X)^{-1}$
DW	: nilai <i>Durbin-Watson</i>
ε_{i-1}	: nilai galat (<i>error</i>) ke- $i - 1$
r_{ij}^2	: koefisien korelasi sederhana antara variabel i dan j
a_{ij}^2	: koefisien korelasi parsial antara variabel i dan j
R	: matriks korelasi dari masing-masing variabel prediktor
X_{ki}	: variabel bebas ke- k pada lokasi ke- i
μ_k	: rata-rata variabel bebas ke- k
σ_k	: simpangan baku variabel bebas ke- k
$\mathbf{f}$	: $(f_1, f_2, \dots, f_n)^T$ dengan $f_i = \left(\frac{e_i^2}{\sigma^2}\right) - 1$
e_i	: $y_i - \hat{y}_i$, $\hat{y}_i$ diperoleh dari metode <i>ordinary least square</i> (OLS)
$\mathbf{Z}$	: matriks $n \times p$ variabel prediktor yang sudah distandarkan
(u_i, v_i)	: koordinat letak geografis (<i>longitude, latitude</i>) dari lokasi pengamatan ke- i
$\beta_0(u_i, v_i)$	: konstanta/intersep pada pengamatan lokasi ke- i
$\beta_k(u_i, v_i)$	: nilai koefisien regresi variabel bebas ke- k pada lokasi pengamatan ke- i
$\mathbf{W}(u_i, v_i)$	: diagonal pembobot yang berbeda-beda dari setiap estimasi parameter pada lokasi ke- i
$w_j(u_i, v_i)$	: nilai pembobot lokasi ke- j untuk mengestimasi parameter pada lokasi ke- i
$\hat{y}_{\neq i}(h)$	: nilai estimasi y_i di mana pengamatan pada lokasi (u_i, v_i) dihilangkan dari proses estimasi.
d_{ij}	: jarak <i>eulidean</i> antara lokasi pengamatan ke- i dan ke- j
$\mathbf{PC}$	: matriks komponen utama
PC_{ik}	: skor komponen utama ke- j pada lokasi ke- i
π	: 3,14

INTISARI

PEMODELAN *GEOGRAPHICALLY WEIGHTED REGRESSION* *PRINCIPAL COMPONENT ANALYSIS* (GWR-PCA) DENGAN FUNGSI PEMBOBOT *FIXED GAUSSIAN KERNEL*

(Studi Kasus: Faktor-faktor yang Memengaruhi Distribusi Persentase PDRB
Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2024)

Oleh

Redella Reffa Herdianti

NIM. 22106020073

Geographically Weighted Regression Principal Component Analysis (GWRPCA) merupakan kombinasi metode *Principal Component Analysis* (PCA) dan *Geographically Weighted Regression* (GWR). Penelitian ini bertujuan untuk memodelkan faktor-faktor yang memengaruhi distribusi persentase Produk Domestik Regional Bruto (PDRB) kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024 menggunakan GWRPCA dengan fungsi pembobot *fixed Gaussian kernel*. PCA digunakan untuk mengatasi multikolinearitas antarvariabel independen, sedangkan heterogenitas spasial diuji menggunakan uji Breusch-Pagan. Pemilihan bandwidth optimum dilakukan menggunakan metode *Cross Validation* (CV), dan estimasi parameter dilakukan dengan metode *Weighted Least Squares* (WLS). Hasil PCA menghasilkan dua komponen utama, yaitu PC1 dan PC2, yang mampu menjelaskan 66,27% total variansi data. Hasil uji Breusch-Pagan menunjukkan adanya heterogenitas spasial ($p\text{-value} = 0,003$) sehingga model GWRPCA layak digunakan. Model GWRPCA menghasilkan nilai R^2 sebesar 0,8170, AIC sebesar 120,7012, dan RMSE sebesar 2,8015, yang lebih baik dibandingkan model RPCA. Hasil penelitian menunjukkan bahwa PC1 dan PC2 berpengaruh signifikan terhadap distribusi persentase PDRB pada seluruh kabupaten/kota di Provinsi Sulawesi Selatan.

Kata kunci: Distribusi Persentase PDRB, GWRPCA, Analisis Komponen Utama, *Fixed Gaussian Kernel*, Heterogenitas Spasial.

ABSTRACT

MODELING GEOGRAPHICALLY WEIGHTED REGRESSION PRINCIPAL COMPONENT ANALYSIS (GWRPCA) USING A FIXED GAUSSIAN KERNEL WEIGHTING FUNCTION

(A Case Study of Factors Affecting the Distribution Percentage of Gross Regional Domestic Product (GRDP) by Regencies/Cities in South Sulawesi Province in 2024)

By

Redella Reffa Herdianti

NIM. 22106020073

Geographically Weighted Regression Principal Component Analysis (GWRPCA) is a combination of Principal Component Analysis (PCA) and Geographically Weighted Regression (GWR). This study aims to model the factors affecting the distribution percentage of Gross Regional Domestic Product (GRDP) across regencies/cities in South Sulawesi Province in 2024 using GWRPCA with a fixed Gaussian kernel weighting function. PCA was employed to address multicollinearity among independent variables, while spatial heterogeneity was examined using the Breusch–Pagan test. The optimum bandwidth was determined using the Cross Validation (CV) method, and parameter estimation was carried out using the Weighted Least Squares (WLS) method. The PCA results produced two principal components, namely PC1 and PC2, which explained 66.27% of the total data variance. The Breusch–Pagan test indicated the presence of spatial heterogeneity ($p\text{-value} = 0.003$), supporting the application of the GWRPCA model. The GWRPCA model yielded an R^2 value of 0.8170, an AIC value of 120.7012, and an RMSE value of 2.8015, outperforming the RPCA model. The results showed that PC1 and PC2 had a significant effect on the distribution percentage of GRDP across all regencies/cities in South Sulawesi Province.

Keywords: Distribution Percentage of GRDP, GWRPCA, Principal Component Analysis, Fixed Gaussian Kernel, Spatial Heterogeneity.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Secara etimologi, statistika berasal dari bahasa Latin "*status*", bahasa Inggris "*state*", atau dalam bahasa Belanda "*stat*" yang berarti negara. Selain itu, istilah statistika dapat dilihat dari bahasa Latin modern, yaitu "*statisticum collegium*" yang berarti dewan negara, serta dalam bahasa Italia, yaitu "*statista*" yang berarti negarawan atau politikus. Statistika adalah suatu disiplin ilmu yang mempelajari sekumpulan konsep dan metode pengumpulan, penyajian, analisis, dan interpretasi data, sampai pada pengambilan keputusan pada situasi di mana terdapat ketidakpastian (Qudratullah, 2013). Berdasarkan penggunaannya, statistika dapat dibagi menjadi dua jenis, yaitu statistika deskriptif dan statistika induktif (inferensia). Statistika deskriptif adalah statistika yang mempelajari bagaimana cara mengumpulkan data, mengolah data, dan menganalisis data, sedangkan statistika inferensia adalah statistika yang mempelajari bagaimana cara mengumpulkan data, mengolah data, menyajikan data, menganalisis data, membuat kesimpulan, dan mengambil keputusan (Supandi, 2020).

Statistika dibagi menjadi dua cabang, yaitu statistika murni dan statistika terapan. Statistika murni adalah statistika yang mempelajari ilmu dan pengetahuan statistika untuk perkembangan teori itu sendiri, sedangkan statistika terapan adalah statistika yang diterapkan secara luas dalam rangka memecahkan permasalahan di berbagai bidang kehidupan (Supandi, 2020).

Salah satu metode statistika yang banyak digunakan dalam menganalisis hubungan antar variabel adalah analisis regresi. Analisis regresi merupakan metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respons (Y) dengan satu atau lebih variabel prediktor (X) (Qudratullah, 2013). Model regresi pada umumnya digunakan ketika unit pengamatan bukan merupakan wilayah atau lokasi, karena model tersebut tidak memperhitungkan wilayah atau lokasi yang

selanjutnya disebut efek spasial. Masalah utama jika analisis ini diterapkan pada data yang mengandung informasi lokasi geografis atau spasial, maka akan terjadi heterogenitas spasial (Bivand et al., 2013).

Heterogenitas spasial merupakan keadaan di mana variabel prediktor yang sama memberikan respons yang tidak sama pada lokasi yang berbeda dalam suatu wilayah penelitian, yaitu data yang memuat informasi geografis suatu daerah, sehingga tidak hanya memuat apa yang diukur. Metode untuk mengatasi efek keragaman spasial adalah model *Geographically Weighted Regression* (GWR). Model GWR adalah suatu metode yang membawa kerangka dari model regresi klasik menjadi regresi terboboti, di mana parameter dihitung pada setiap titik wilayah geografis sehingga setiap wilayah geografis memiliki nilai parameter yang berbeda (Fotheringham et al., 2002). Dengan demikian, model ini mampu menangkap variasi hubungan antar variabel yang bersifat lokal dan tidak dapat dijelaskan oleh model regresi biasa. Model GWR secara umum dinyatakan sebagai:

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

di mana y_i merupakan variabel respons pada lokasi ke- i , x_{ik} adalah variabel prediktor ke- k , (u_i, v_i) menyatakan koordinat geografis, $\beta_k(u_i, v_i)$ adalah parameter lokal, dan ε_i merupakan galat (Sari et al., 2016).

Penerapan model GWR sering menghadapi permasalahan berupa adanya multikolinearitas antar variabel prediktor yang dapat memengaruhi keakuratan estimasi parameter. Permasalahan tersebut diatasi dengan menggunakan metode *Principal Component Analysis* (PCA). PCA merupakan suatu teknik reduksi dimensi yang bertujuan untuk mentransformasi variabel-variabel prediktor yang saling berkorelasi menjadi sekumpulan komponen utama yang saling bebas, sehingga informasi yang terkandung dalam data dapat dipertahankan secara optimal (Richard A. Johnson, 2007). Kombinasi antara GWR dan PCA, yang selanjutnya disebut GWR-PCA, memungkinkan analisis spasial yang lebih akurat karena

mampu menangani multikolinearitas sekaligus mempertimbangkan keragaman spasial antar wilayah. Model GWR-PCA dapat dinyatakan dalam bentuk:

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

di mana PC_{ik} merupakan komponen utama ke- k pada lokasi ke- i yang merupakan hasil transformasi dari variabel prediktor asli. Penggunaan komponen utama ini menjadi pembeda utama dibandingkan model GWR, karena variabel prediktor yang digunakan telah bebas dari multikolinearitas (Sari et al., 2016).

Menurut Badan Pusat Statistik (BPS), Produk Domestik Regional Bruto (PDRB) merupakan indikator ekonomi yang digunakan untuk mengukur total nilai tambah barang dan jasa yang dihasilkan dalam suatu wilayah pada periode tertentu, yang mencerminkan tingkat kesejahteraan ekonomi daerah. PDRB juga merefleksikan kapasitas sumber daya ekonomi suatu wilayah, di mana semakin tinggi nilai PDRB menunjukkan semakin besar potensi ekonomi yang dimiliki, dan sebaliknya (Herbiansyah et al., 2022).

Provinsi Sulawesi Selatan merupakan salah satu wilayah di Indonesia dengan perkembangan ekonomi yang cukup pesat. Sulawesi Selatan dipilih sebagai wilayah penelitian karena berperan sebagai pusat pertumbuhan ekonomi di Kawasan Timur Indonesia, dengan nilai PDRB yang konsisten menjadi salah satu tertinggi di luar Pulau Jawa (Badan Pusat Statistik, 2025a). Selain itu, Sulawesi Selatan memiliki karakteristik wilayah yang beragam, mencakup kawasan perkotaan, pesisir, dan pegunungan, sehingga menghasilkan variasi kondisi sosial, ekonomi, dan demografi yang kaya antar kabupaten/kota. Kondisi ini menjadikan Sulawesi Selatan sebagai objek yang relevan untuk analisis spasial yang mampu menangkap heterogenitas lokal. Salah satu indikator yang umum digunakan untuk menggambarkan kondisi perekonomian daerah adalah Produk Domestik Regional Bruto (PDRB). Berdasarkan data Badan Pusat Statistik, nilai PDRB per kapita Sulawesi Selatan atas dasar harga berlaku (ADHB) pada tahun 2020 sebesar 55,68

juta rupiah. Nilai tersebut meningkat menjadi 59,50 juta rupiah pada tahun 2021 dan terus menunjukkan tren positif hingga mencapai 73,57 juta rupiah pada tahun 2024 seiring dengan pulihnya aktivitas perekonomian pascapandemi COVID-19 (Badan Pusat Statistik, 2025a).

Berdasarkan harga konstan tahun 2010, PDRB per kapita Sulawesi Selatan juga mengalami peningkatan, yaitu dari 36,25 juta rupiah pada tahun 2020 menjadi 41,86 juta rupiah pada tahun 2024. Peningkatan tersebut menunjukkan bahwa kondisi perekonomian di Sulawesi Selatan cenderung mengalami pertumbuhan yang positif (Badan Pusat Statistik, 2025a). Namun demikian, distribusi persentase PDRB antar kabupaten/kota di Sulawesi Selatan masih menunjukkan ketimpangan. Beberapa wilayah memiliki porsi ekonomi yang lebih besar dibanding wilayah lainnya, seperti Makassar yang menjadi pusat kegiatan ekonomi dan perdagangan di Sulawesi Selatan.

Ketimpangan distribusi persentase PDRB antar wilayah tersebut mendorong perlunya analisis yang mampu menangkap perbedaan karakteristik lokal. Oleh karena itu, diperlukan metode analisis spasial yang dapat mengakomodasi variasi hubungan antar variabel pada setiap lokasi. Selain itu, keberadaan multikolinearitas antar variabel independen juga perlu diperhatikan dalam pemodelan. Secara teoritis, beberapa variabel yang digunakan dalam penelitian ini memiliki keterkaitan konseptual yang erat, seperti rata-rata lama sekolah dan harapan lama sekolah yang sama-sama merupakan indikator pendidikan, serta umur harapan hidup dan pengeluaran per kapita yang keduanya mencerminkan dimensi kesejahteraan masyarakat. Keterkaitan ini mengindikasikan bahwa variabel-variabel tersebut berpotensi mengandung informasi yang saling tumpang tindih. Indikasi tersebut diperkuat oleh hasil eksplorasi awal data yang menunjukkan adanya korelasi tinggi antar beberapa variabel prediktor, sehingga diperlukan pendekatan yang mampu mengatasi multikolinearitas dalam pemodelan.

Berdasarkan penjabaran di atas, akan dilakukan penelitian mengenai faktor-faktor yang memengaruhi Distribusi Persentase PDRB secara spasial menurut

kabupaten/kota di Provinsi Sulawesi Selatan dengan menggunakan model *Geographically Weighted Regression* berbasis *Principal Component Analysis* (GWRPCA). Dalam penelitian ini, fungsi pembobot yang digunakan adalah *fixed Gaussian kernel*. Fungsi *Gaussian* dipilih karena memiliki sifat kontinu dan menghasilkan bobot yang menurun secara halus (*smooth decay*) seiring bertambahnya jarak antar lokasi, sehingga tidak terjadi perubahan bobot secara mendadak seperti pada fungsi *bisquare* yang memberikan bobot nol mutlak di luar batas bandwidth (Brunsdon et al., 1998). Adapun penentuan nilai bandwidth optimal dilakukan menggunakan metode *Cross-Validation* (CV), di mana skema *fixed* menghasilkan nilai CV yang lebih baik dibandingkan skema *adaptive* pada data penelitian ini, sehingga skema *fixed* dipilih sebagai yang paling sesuai untuk memodelkan hubungan spasial antar kabupaten/kota di Provinsi Sulawesi Selatan. Untuk mengevaluasi performa model, digunakan tiga indikator kebaikan model, yaitu Koefisien Determinasi (R^2), *Akaike Information Criterion* (AIC), dan *Root Mean Square Error* (RMSE). R^2 digunakan untuk mengukur seberapa besar variasi variabel respons yang mampu dijelaskan oleh model (Fotheringham et al., 2002). AIC digunakan untuk membandingkan kualitas antar model dengan mempertimbangkan kompleksitas model, di mana model dengan nilai AIC lebih kecil dianggap lebih baik (Aike, 1974). Adapun RMSE digunakan untuk mengukur rata-rata besarnya kesalahan prediksi model dalam satuan yang sama dengan variabel respons (Chai et al., 2014).

Variabel independen yang digunakan meliputi rata-rata lama sekolah, pengeluaran per kapita, persentase penduduk miskin, harapan lama sekolah, umur harapan hidup, jumlah angkatan kerja, tingkat pengangguran terbuka, *gini ratio*, dan jumlah penduduk. Berdasarkan hal tersebut, penelitian ini berjudul "Pemodelan *Geographically Weighted Regression Principal Component Analysis* (GWR-PCA) dengan Fungsi Pembobot *Fixed Gaussian Kernel*" pada studi kasus faktor-faktor yang memengaruhi Distribusi Persentase PDRB menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024.

1.1 Rumusan Masalah

Berdasarkan uraian pada latar belakang di atas, maka dapat dirumuskan permasalahan dalam penelitian ini sebagai berikut:

1. Bagaimana langkah-langkah pemodelan *Geographically Weighted Regression* berbasis *Principal Component Analysis* (GWRPCA) dengan fungsi pembobot *fixed Gaussian kernel*?
2. Faktor-faktor apa saja yang memengaruhi Distribusi Persentase PDRB menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024?
3. Bagaimana interpretasi model GWRPCA berdasarkan faktor-faktor yang memengaruhi Distribusi Persentase PDRB menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024?

1.2 Tujuan Penelitian

Berdasarkan rumusan masalah dan batasan masalah yang telah diuraikan, penulis merumuskan beberapa tujuan yang ingin dicapai dalam penelitian ini yaitu:

1. Mengetahui langkah-langkah pemodelan *Geographically Weighted Regression* berbasis *Principal Component Analysis* (GWRPCA) dengan fungsi pembobot *fixed Gaussian kernel*;
2. Mengetahui faktor-faktor yang memengaruhi Distribusi Persentase PDRB menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024 menggunakan model GWRPCA; dan
3. Mengetahui interpretasi model GWRPCA berdasarkan faktor-faktor yang memengaruhi Distribusi Persentase PDRB menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024.

1.3 Manfaat Penelitian

Manfaat penelitian yang ingin diperoleh dari penelitian ini adalah:

1. Bagi penulis, penelitian ini dapat menambah wawasan dan keterampilan dalam menerapkan ilmu statistika untuk memodelkan Distribusi Persentase PDRB menggunakan metode *Geographically Weighted Regression* berbasis

Principal Component Analysis (GWRPCA) dengan pembobot fixed Gaussian kernel.

2. Bagi pembaca, penelitian ini dapat menjadi sumber informasi mengenai faktor-faktor yang memengaruhi Distribusi Persentase PDRB menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024, sehingga dapat dijadikan referensi untuk penelitian selanjutnya.
3. Bagi instansi, penelitian ini dapat digunakan sebagai bahan pertimbangan dalam perumusan kebijakan pembangunan ekonomi daerah, serta sebagai sumber dalam pengembangan keilmuan, khususnya di bidang statistika dan analisis spasial.

1.4 Batasan Masalah

Batasan masalah dalam penelitian sangat diperlukan agar ruang lingkup dalam penelitian ini tidak keluar dari permasalahan. Oleh karena itu penulis memberikan batasan masalah dalam penelitian ini sebagai berikut:

1. Fungsi pembobot yang digunakan adalah *Fixed Gaussian Kernel*.
2. Penaksiran parameter menggunakan *Weighted Least Squares (WLS)*.
3. Kriteria penentuan *bandwidth optimum* menggunakan *Cross Validation (CV)*.
4. Pemilihan model terbaik menggunakan nilai koefisien determinasi (R^2), *Akaike Information Criterion (AIC)*, dan *Root Mean Square Error (RMSE)*.
5. Penelitian ini difokuskan pada data Distribusi Persentase Produk Domestik Regional Bruto (PDRB) kabupaten/kota di Provinsi Sulawesi Selatan dengan peubah prediktor yang diduga memengaruhinya.
6. Reduksi dimensi variabel prediktor dilakukan menggunakan metode *Principal Component Analysis (PCA)* sebelum pemodelan GWR.
7. Aplikasi pengolahan data yang digunakan adalah R Studio versi 4.5.3, Microsoft Excel, dan QGIS versi 4.0.2.

1.5 Tinjauan Pustaka

Tinjauan pustaka pada penelitian ini dideskripsikan dan ditelaah melalui buku, tesis, skripsi, jurnal, maupun sumber lainnya yang relevan dengan penelitian yang dilakukan saat ini. Adapun penelitian-penelitian terdahulu yang digunakan adalah:

1. Penelitian oleh Mohd Syafrizal et al. (2024) yang berjudul *Modeling Gross Enrollment Rate for Higher Education in Central Java Province Using Principal Component Geographically Weighted Regression Approach*. Metode yang digunakan adalah model GWRPCA dengan pembobot fungsi kernel *Fixed Gaussian*, penaksiran parameter menggunakan metode WLS, uji heterogenitas spasial menggunakan uji Breusch-Pagan, serta uji multikolinearitas menggunakan VIF. Pemilihan *bandwidth* optimum menggunakan kriteria *Cross Validation* (CV). Ukuran kebaikan model menggunakan nilai AIC dan R^2 . Hasil yang diperoleh adalah model GWRPCA dengan tiga komponen utama lebih baik dibandingkan model regresi linier berganda maupun GWR tanpa PCA, dengan nilai R^2 sebesar 0,962 dan nilai AIC terkecil sebesar 169,43.
2. Penelitian oleh Al Azies & Ariyanto (2024) yang berjudul *GWRPCA Algorithmic Framework: Analyzing Spatial Dynamics of Poverty in East Java Province*. Metode yang digunakan adalah model GWRPCA dengan pembobot fungsi kernel *Adaptive Gaussian* berdasarkan nilai *Cross Validation* (CV) terkecil, serta uji heterogenitas spasial dilakukan dengan uji Breusch-Pagan. Ukuran kebaikan model menggunakan nilai R^2 . Hasil penelitian menunjukkan bahwa model GWRPCA dengan nilai R^2 sebesar 85,10% lebih baik digunakan dibandingkan model RPCA ($R^2 = 72,84\%$) dalam memodelkan kemiskinan di Provinsi Jawa Timur.
3. Penelitian oleh Sari et al. (2016) yang berjudul *Geographically Weighted Regression Principal Component Analysis (GWRPCA) pada Pemodelan Pendapatan Asli Daerah di Jawa Tengah*. Metode yang digunakan adalah

model GWRPCA dengan pembobot kernel *fixed Gaussian*, penaksiran parameter menggunakan metode WLS, serta uji kecukupan data PCA menggunakan KMO dan MSA. Pemilihan *bandwidth* optimum menggunakan kriteria CV. Ukuran kebaikan model menggunakan nilai AIC serta nilai R^2 . Hasil yang diperoleh adalah model GWRPCA lebih baik digunakan untuk pemodelan PAD di Jawa Tengah dengan nilai R^2 sebesar 0,9929 dan nilai AIC terkecil.

4. Penelitian oleh Harris et al. (2011) yang berjudul *Geographically Weighted Principal Components Analysis*. Metode yang digunakan adalah GWPCA dengan pembobot fungsi kernel *Fixed Bisquare* dan *Adaptive Bisquare*, serta pemilihan *bandwidth* optimum menggunakan metode *leave-one-out Cross Validation* berdasarkan statistik *Goodness of Fit* (GOF). Ukuran kebaikan model menggunakan nilai GOF dan persentase total variansi yang dijelaskan oleh komponen utama lokal.
5. Penelitian oleh Pratnyaningrum et al. (2015) yang berjudul *Pemodelan Persentase Balita Gizi Buruk di Jawa Tengah dengan Pendekatan Geographically Weighted Regression Principal Components Analysis (GWRPCA)*. Metode yang digunakan adalah model GWRPCA dengan pembobot fungsi kernel *Adaptive Bisquare*, penaksiran parameter menggunakan metode WLS, serta uji asumsi PCA menggunakan KMO, MSA, dan *Bartlett Test of Sphericity*. Ukuran kebaikan model menggunakan R^2 global dan R^2 lokal. Hasil yang diperoleh adalah dua komponen utama PC1 dan PC2 mampu menerangkan total varian data sebesar 78,43% dengan koefisien determinasi global sebesar 0,6313.

Tabel 1. 1 Persamaan dan Perbedaan dengan Penelitian Terdahulu

No	Peneliti	Model	Pembobot	Metode	Kebaikan Model	Studi Kasus
1	Mohd Syafrizal et al. (2024)	GWRPCA	<i>Fixed Gaussian Kernel</i>	WLS, CV	AIC, R^2	Angka Partisipasi Murni Perguruan Tinggi di Jawa Tengah Tahun 2021
2	Al Azies & Ariyanto (2024)	GWRPCA	<i>Adaptive Gaussian Kernel</i>	CV	R^2	Kemiskinan di Provinsi Jawa Timur Tahun 2016
3	Sari et al. (2016)	GWRPCA	<i>Fixed Gaussian Kernel</i>	WLS, CV	AIC, R^2	Pendapatan Asli Daerah di Jawa Tengah Tahun 2012–2013
4	Harris et al. (2011)	GWPCA	<i>Fixed dan Adaptive Bisquare Kernel</i>	CV	GOF	Struktur Sosial di Greater Dublin, Irlandia
5	Pratnyaningrum et al. (2015)	GWRPCA	<i>Adaptive Bisquare Kernel</i>	WLS, CV	R^2	Persentase Balita Gizi Buruk di Jawa Tengah

Berdasarkan Tabel 1.1, diperoleh persamaan antara penelitian terdahulu dengan penelitian yang akan dikaji penulis saat ini, yaitu penggunaan model *Geographically Weighted Regression–Principal Component Analysis* (GWR-PCA). Adapun perbedaan penelitian ini dengan penelitian sebelumnya terletak pada studi kasus yang digunakan serta ukuran kebaikan model yang dikaji, yaitu menggunakan nilai AIC, RMSE, dan R^2 .

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini digambarkan sebagai berikut, untuk mempermudah pemahaman penulisan secara keseluruhan.

BAB I PENDAHULUAN

Bab ini menguraikan latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, tinjauan pustaka, serta sistematika penulisan yang digunakan dalam penelitian.

BAB II LANDASAN TEORI

Bab ini membahas teori-teori yang mendukung penelitian, meliputi variabel random, ekspektasi, variansi, kovariansi, koefisien korelasi, matriks, model regresi linear berganda, estimasi parameter regresi linear berganda, uji signifikansi model regresi, uji asumsi klasik, kriteria kebaikan model, data spasial, serta konsep dan indikator yang digunakan dalam penelitian, yaitu distribusi persentase PDRB, indikator pendidikan, indikator ekonomi, indikator kesehatan, indikator ketenagakerjaan, serta indikator ketimpangan dan kependudukan.

BAB III METODE PENELITIAN

Bab ini menjelaskan data dan jenis penelitian, objek penelitian, metode pengumpulan data, variabel penelitian, metode analisis data, serta *flowchart* penelitian yang digunakan sebagai pedoman pelaksanaan penelitian.

BAB IV PEMBAHASAN

Bab ini membahas pengembangan metode yang digunakan dalam penelitian, meliputi pengujian multikolinearitas, pengujian asumsi *Principal Component*

Analysis (PCA), standardisasi data, pembentukan komponen utama, pengujian heterogenitas spasial, pembentukan model *Geographically Weighted Regression* (GWR), estimasi parameter model GWR, penentuan *bandwidth* optimum, penentuan fungsi pembobot, pembentukan model *Geographically Weighted Regression–Principal Component Analysis* (GWRPCA), estimasi parameter model GWRPCA, pengujian kesesuaian model, pengujian parameter model secara serentak dan parsial, serta pemilihan model terbaik menggunakan koefisien determinasi (R^2), *Akaike Information Criterion* (AIC), dan *Root Mean Square Error* (RMSE).

BAB V STUDI KASUS

Bab ini membahas penerapan metode yang telah dikembangkan pada data penelitian. Analisis yang dilakukan meliputi statistika deskriptif, pengujian asumsi PCA, pembentukan komponen utama, pemodelan Regresi Linear Berganda berbasis *Principal Component Analysis* (RPCA), pengujian asumsi klasik model RPCA, pemodelan *Geographically Weighted Regression–Principal Component Analysis* (GWRPCA), pengujian kesesuaian model, serta pemilihan model terbaik berdasarkan nilai R^2 , AIC, dan RMSE.

BAB VI PENUTUP

Bab ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian serta saran yang dapat digunakan sebagai bahan pertimbangan untuk penelitian selanjutnya.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

Pemodelan GWRPCA dengan fungsi pembobot *fixed Gaussian kernel* diawali dengan analisis statistika deskriptif untuk memberikan gambaran umum karakteristik data penelitian. Selanjutnya dilakukan deteksi multikolinearitas menggunakan *Variance Inflation Factor* (VIF), yang menunjukkan bahwa variabel rata-rata lama sekolah, jumlah angkatan kerja, dan jumlah penduduk terindikasi mengalami multikolinearitas. Kelayakan data diuji menggunakan KMO, MSA, dan *Bartlett Test of Sphericity* guna mengatasi hal tersebut, dan hasilnya menyatakan data layak untuk dianalisis menggunakan PCA. Analisis komponen utama kemudian dilakukan dengan menstandarisasi data, membentuk matriks ragam-peragam, serta menghitung nilai eigen dan vektor eigen, sehingga diperoleh dua komponen utama yang secara kumulatif mampu menjelaskan 66,27% total variansi data. Kedua komponen utama tersebut selanjutnya digunakan sebagai variabel independen baru dalam membangun model RPCA.

Keberadaan heterogenitas spasial dideteksi menggunakan uji *Breusch-Pagan*, yang hasilnya membenarkan penggunaan GWRPCA. Koordinat spasial masing-masing kabupaten/kota kemudian ditentukan dan jarak *Euclidean* antar lokasi pengamatan dihitung. Fungsi pembobot terbaik dipilih menggunakan metode *Cross Validation*, di mana fungsi pembobot *fixed Gaussian kernel* dengan *bandwidth* optimum sebesar 4,040536 menghasilkan nilai CV terkecil sehingga dipilih sebagai fungsi pembobot yang paling sesuai. Parameter model GWRPCA pada setiap lokasi diestimasi menggunakan metode *Weighted Least Squares*, sehingga diperoleh koefisien regresi lokal yang berbeda untuk setiap kabupaten/kota. Model GWRPCA terpilih sebagai model terbaik berdasarkan nilai R^2 , AIC, dan RMSE yang lebih baik dibandingkan model RPCA.

Kesembilan variabel independen tereduksi menjadi dua komponen utama melalui analisis PCA. PC1 dibentuk oleh kontribusi seluruh sembilan variabel dengan nilai vektor *eigen* yang relatif seragam, di mana variabel Harapan Lama Sekolah, Rata-rata Lama Sekolah, dan Tingkat Pengangguran Terbuka menjadi variabel paling dominan, sehingga PC1 diinterpretasikan sebagai dimensi kualitas sumber daya manusia dan pembangunan manusia. PC2 didominasi secara mencolok oleh Jumlah Angkatan Kerja dan Jumlah Penduduk dengan lompatan nilai vektor *eigen* yang cukup besar dibandingkan variabel lainnya, sehingga PC2 diinterpretasikan sebagai dimensi skala ketenagakerjaan dan kependudukan. Berdasarkan hasil uji parsial lokal model GWRPCA, kedua komponen utama tersebut berpengaruh signifikan terhadap distribusi persentase PDRB di seluruh 24 kabupaten/kota di Provinsi Sulawesi Selatan tahun 2024.

Model GWRPCA menghasilkan koefisien regresi lokal yang berbeda untuk setiap kabupaten/kota, mencerminkan adanya variasi spasial dalam hubungan antara komponen utama dan distribusi persentase PDRB. Koefisien lokal komponen pertama bernilai positif di seluruh wilayah dengan rentang yang relatif sempit, di mana nilai tertinggi terdapat di Takalar dan Kota Makassar, mengindikasikan bahwa peningkatan kualitas sumber daya manusia dan kesejahteraan di wilayah tersebut memberikan dampak lebih besar terhadap distribusi persentase PDRB. Koefisien lokal komponen kedua bernilai negatif di seluruh wilayah, dengan nilai paling negatif terdapat di Luwu Utara, mengindikasikan bahwa wilayah tersebut paling sensitif terhadap perbedaan skala kependudukan. Nilai Local R^2 yang berkisar antara 0,8126 hingga 0,8208 menunjukkan bahwa model GWRPCA mampu menjelaskan lebih dari 81% variasi distribusi persentase PDRB di seluruh kabupaten/kota Provinsi Sulawesi Selatan.

6.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, saran yang dapat diberikan adalah sebagai berikut:

1. Penelitian selanjutnya dapat mempertimbangkan penambahan jumlah komponen utama hingga proporsi variansi kumulatif mencapai minimal 80% agar informasi yang dipertahankan dari variabel asal lebih besar.
2. Penelitian selanjutnya dapat menggunakan fungsi pembobot lain, seperti *adaptive Gaussian kernel* atau *bisquare kernel*, untuk memperoleh model yang lebih optimal.
3. Penelitian selanjutnya dapat menggunakan data dengan jumlah observasi yang lebih banyak atau cakupan wilayah yang lebih luas agar variasi spasial antarwilayah dapat tergambarkan dengan lebih jelas.
4. Penelitian selanjutnya dapat mempertimbangkan kriteria pemilihan bandwidth lain, seperti *Akaike Information Criterion corrected* (AICc), selain *Cross Validation* (CV), untuk memperoleh bandwidth yang lebih optimal.

DAFTAR PUSTAKA

- Aike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Al Azies, H., & Ariyanto, N. (2024). Gwrpca Algorithmic Framework: Analyzing Spatial Dynamics of Poverty in East Java Province. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(1), 126–136. <https://doi.org/10.23960/jitet.v12i1.3945>
- Anggraini, W. P., & Warsitasari, W. D. (2011). Pengaruh Tingkat Pengangguran Terbuka, Rasio Gini, Kemiskinan dan Upah Minimum terhadap Pertumbuhan Ekonomi di Provinsi Jawa Timur. *Journal of Islamic Economic and Business (JIEB)*, 05(02).
- Anton, H., & Rorres, C. (2010). *Elementary Linear Algebra by Howard Anton 10th edition*.
- Ardhani, R., Herrhyanto, N., & Agustina, F. (2023). Multiscale Geographically Weighted Regression (MGWR) untuk Memodelkan Nilai Angka Buta Huruf di Provinsi Sumatera Selatan Tahun 2021. *Jurnal EurekaMatika*, 11(2), 89–98. <https://doi.org/10.17509/jem.v11i2.63140>
- Badan Pusat Statistik. (2025a). *Produk Domestik Regional Bruto Provinsi Sulawesi Selatan Menurut Lapangan Usaha 2020-2024* (Vol. 11).
- Badan Pusat Statistik. (2025b). *Stistik Pendidikan Provinsi Sulawesi Selatan 2024* (Vol. 10).
- Bain, L. J., & Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*. Duxbury Press.
- Ben-Said, M. (2021). Spatial point-pattern analysis as a powerful tool in identifying pattern-process relationships in plant ecology: an updated review. *Ecological Processes*, 10(1). <https://doi.org/10.1186/s13717-021-00314-4>
- Bivand, R. S., Pebesma, E., & Gomez-Rubio, V. (2013). Applied Spatial Data Analysis with R. In *Spatial Demography* (Vol. 1, Issue 2). <https://doi.org/10.1007/bf03354901>

- Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). *Geographically weighted regression modelling spatial non-stationarity*. 431–443.
- Caraka, R. E., & Yasin, H. (2017). *Geographically Weighted Regression (GWR) Sebuah pendekatan Regresi Geografis* (Vol. 32, Issue 3).
- Chai, T., Draxler, R. R., & Prediction, C. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 2005, 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Fadli, M. R., Goejantoro, R., & Wasono. (2018). Pemodelan Geographically Weighted Regression (GWR) Dengan Fungsi Pembobot Tricube Terhadap Angka Kematian Ibu (AKI) Di Kabupaten Kutai Kartanegara Tahun 2015. *Jurnal Eksponensial*, 9(1), 11–18. <http://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/209>
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression: the analysis of spatially varying realtionships*. John Wiley & Sons, Ltd.
- Ghozali, I. (2021). *Aplikasi Analisis Multivariate dengan Progran IBM SPSS 26*. Universitas Diponegoro.
- Hadi, A. E. S., & Hasmarini, M. I. (2026). Analisis Pengaruh Jumlah Penduduk , Investasi (Penanaman Modal Asing), Pendidikan dan Kemiskinan Terhadap Produk Domestik Regional Bruto Di Indonesia. *Indonesian Journal of Digital Business*, 5(4), 2658–2667.
- Härdle, W. K., & Simar, L. (2019). Applied Multivariate Statistical Analysis. In *Applied Multivariate Statistical Analysis*. <https://doi.org/10.1007/978-3-030-26006-4>
- Harris, P., Brunsdon, C., & Charlton, M. (2011). Geographically weighted principal components analysis. *International Journal of Geographical Information Science*, 25(10), 1717–1736. <https://doi.org/10.1080/13658816.2011.554838>
- Herbiansyah, T., Yahya, I., Baharuddin, Agusrawati, Ruslan, & Program, L. L.

- (2022). Pemodelan Produk Domestik Regional Bruto di Indonesia dengan Geographically Weighted Regression. *Seminar Nasional Sains Dan Terapan VI*, 6(April), 23–33.
- Huda, N., & Indahsari, K. (2021). Pengaruh Rata-Rata Lama Sekolah, Angka Harapan Hidup dan Penegeluaran Perkapita terhadap Pertumbuhan Ekonomi Provinsi Jawa Timur 2014-2018. *Buletin Ekonomika Pembangunan*, 2(1), 55–66.
- Jolliffe, I. T. (2002). *Principal Component Analysis, Second Edition*. Springer.
- Jr., J. F. H., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7th Editio).
- Kaiser, H. F. (1974). *An Index of Factorial Simplicity*. 39(1), 31–36.
- Mohd Syafrizal, Widyanti Rahayu, & Dania Siregar. (2024). Modeling Gross Enrollment Rate for Higher Education in Central Java Province Using Principal Component Geographically Weighted Regression Approach. *Jurnal Statistika Dan Aplikasinya*, 8(1), 41–53. <https://doi.org/10.21009/jsa.08104>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons, Ltd.
- N. Gujarati, D., & Porter, D. C. (2009). *Basic Econometrics Fifth Edition*. Douglas Reiner.
- Noya van Delsen, M. S., Wattimena, A. Z., & Saputri, S. (2017). Use of Principal Component Analysis (PCA) Methods to Reduce Inflation Factors in Ambon City. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 11(2), 109–118.
- Nugroho, S. (2008). *Statistika Multivariat Terapan*.
- Perwitasari, I. D., Radjab, M., & Latief, I. (2023). Pengaruh Pendidikan, PDRB, dan Gini Ratio terhadap Kemiskinan di Sulawesi Selatan. *Journal of Scientech Research and Development*, 5(2), 223–238.
- Pratnyaningrum, N., Yasin, H., & Hoyyi, A. (2015). Pemodelan Persentase Balita Gizi Buruk di Jawa Tengah dengan Pendekatan Geographically Weighted

- Regression Principal Components Analysis (GWRPCA). *Jurnal Gaussian*, 4(2), 171–180. <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- Qudratullah, M. F. (2013). *Analisis regresi terapan: teori, contoh kasus, dan aplikasi dengan SPSS* ((Ed.1)). CV Andi Offset.
- Richard A. Johnson, D. W. W. (2007). *Applied Multivariate Statistical Analysis (6th Edition)*.
- Santoso, S. (2017). *Statistik Multivariat dengan SPSS*. PT Elex Media Komputindo.
- Sari, N., Yasin, H., & Prahutama, A. (2016). Geographically Weighted Regression Principal Component Analysis (Gwrpca) Pada Pemodelan Pendapatan Asli Daerah Di Jawa Tengah. *Jurnal Gaussian*, 5(4), 717–726. <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- Schwarz, G. (2978). Estimating The Dimension of a Model. *The Annals of Statistics*, 6, 461–464.
- Selma, L., Rahim, A., Astuty, S., & Muhammad, I. (2026). Determinan Faktor-Faktor Yang Mempengaruhi Pengangguran Terbuka di Berbagai Kabupaten / Kota di Sulawesi Selatan Pendahuluan. *Jurnal Ilmu Ekonom*, 9(1), 805–815.
- Surbakti, C. A., Sinaga, K. T., & Damanik, A. L. (2025). Analisis Pengaruh PDRB Per Kapita , Umur Harapan Hidup , dan Harapan Lama Sekolah Terhadap Indeks Pembangunan Manusia di Provinsi Sulawesi Utara. *Jurnal Penelitian Ilmu-Ilmu Sosial*, 2(June), 513–524. <https://doi.org/https://doi.org/10.5281/enodenodo.15692817>
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2011). *Probability & Statistics for Engineers & Scientist*. Pearson.
- Widiyanti, K. D., Fadlurohman, A., & Purwanto, D. (2025). Geographically Weighted Regression Modeling Using Fixed and Adaptive Kernel Weights for the Human Development Index Case in West Java Province. *Journal Of Data Insights*, 3(2), 92–103. <https://doi.org/10.26714/jodi.v3i2.887>