PEMODELAN *STATISTICAL DOWNSCALING* MENGGUNAKAN *ROBUST JACKKNIFE RIDGE REGRESSION* DENGAN *ESTIMATOR GENERALIZED-M* UNTUK PERAMALAN CURAH HUJAN DI KABUPATEN PANGKEP

RUSLINDA
H051201045

PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

PEMODELAN *STATISTICAL DOWNSCALING* MENGGUNAKAN *ROBUST JACKKNIFE RIDGE REGRESSION* DENGAN *ESTIMATOR GENERALIZED-M* UNTUK PERAMALAN CURAH HUJAN DI KABUPATEN PANGKEP

RUSLINDA
H051201045

Skripsi

sebagai salah satu syarat memperoleh gelar Sarjana Statistika

Program Studi Statistika

pada

PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
S MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

SKRIPSI

PEMODELAN *STATISTICAL DOWNSCALING* MENGGUNAKAN *ROBUST JACKKNIFE RIDGE REGRESSION* DENGAN *ESTIMATOR GENERALIZED-M* UNTUK PERAMALAN CURAH HUJAN DI KABUPATEN PANGKEP

yang disusun dan diajukan oleh

RUSLINDA
H051201045

Skripsi,

Telah dipertahankan di depan Panitia Ujian Sarjana Statistika pada Tanggal 15 Agustus 2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

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

Mengesahkan:
Pembimbing Tugas Akhir,

Sitti Sahriman, S.Si., M.Si.
NIP. 19881018 201504 2 002

Mengetahui:
Ketua Program Studi,

Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19770808 200501 2 002

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Pemodelan *Statistical Downscaling* Menggunakan *Robust Jackknife Ridge Regression* dengan *Estimator Generalized-M* untuk Peramalan Curah Hujan di Kabupaten Pangkep" adalah benar karya saya dengan arahan dari pembimbing Sitti Sahriman, 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, 15 Agustus 2024

10000 METERAI TEMPEL 401C9ALX325661947

Ruslinda
NIM H051201045

v

UCAPAN TERIMA KASIH

Assalamu'alaikum warahmatullahi wabarakatuh

Bismillahirrahmanirrahim, puji syukur penulis panjatkan kepada Allah *Subhanahu Wa ta'ala* yang telah melimpahkan segala ramhat, berkat, dan hidayah-Nya, serta memberikan ilmu pengetahuan, pengalaman, kekuatan, kesabaran, dan kesempatan sehingga penulis dapat menyelesaikan skripsi ini, ***Alhamdulillahrobbil'alamin***, shalawat dan salam senantiasa tercurah kepada junjungan besar kita, Nabi Muhammad *Shallallahu 'Alaihi Wa sallam*, pembawa cahaya kebenaran beserta keluarga dan para sahabatnya.

Penulis menyampaikan ucapan terima kasih yang mendalam kepada **Ibu Sitti Sahriman, S.Si., M.Si.** selaku Dosen Pembimbing Utama, yang dengan penuh kesabaran dan ketulusan meluangkan waktu di tengah kesibukannya untuk membimbing, memberi arahan dan masukan yang membangun, serta motivasi dalam penulisan skripsi ini. Terima kasih yang mendalam kepada **Ibu Dr. Anna Islamiyati, S.Si., M.Si.,** dan **Bapak Dr. Nirwan, M.Si.** selaku Tim Dosen Penguji atas saran dan kritikan yang membangun dalam penyempurnaan penulisan skripsi ini. Serta, ucapan terima kasih yang tulus kepada **Pimpinan Universitas Hasanuddin**, **Jajaran Dosen,** serta **Staf lingkup Departemen Statistika** dan **Fakultas MIPA** yang telah memberikan ilmu dan kemudahan selama masa perkuliahan.

Dengan segenap kerendahan hati dan penuh cinta, penulis mengucapkan terima kasih yang tak terhingga kepada orang tua tercinta, Ibunda **Haswati** dan Ayahanda **Abd. Jalil** atas segala doa, pengorbanan, cinta, kasih sayang, dan dukungan yang tiada henti. Keduanya telah menjadi sumber kekuatan dan semangat yang telah mengiringi setiap langkah penulis. Terima kasih juga kepada kakak saya tercinta **Arni** atas segala bantuan, dukungan dan hiburan yang tak ternilai. Terima kasih juga untuk kedua saudara tercinta kakak **Ratmi** dan adik **Rifay** atas dorongan semangat dan perhatian yang tulus. Terima kasih kepada sahabat saya tercinta sejak kecil **Amar Ma'rub, Anastasya L,** dan **Efi Rahayu** yang telah menjadi sahabat setia dalam setiap momen, menjadi tempat curahan hati dan berbagi segala kesulitan serta suka cita dalam perjuangan di perantauan ini, terima kasih telah hadir membawa tawa dan bahagia.

Terima kasih yang sebesar-besarnya kepada **STATISTIKA 2020** atas kebersamaan, bantuan, dan motivasi selama menjalani masa studi. Juga, terima kasih kepada keluarga besar **Himastat FMIPA Unhas**, khususnya **POIS20N** atas setiap kenangan indah, canda tawa, dukungan luar biasa, dan bantuan tak terhingga dalam setiap langkah perjalanan akademis ini. Bersama-sama, kita telah melewati setiap semester dengan penuh perjuangan dan kebersamaan. Tetaplah **SALING MENGINGAT** di manapun kita berada teruntuk sahabat-sahabat penulis, **Liza**, **Nahda**, **Ayu**, **Najla**, **Alisha**, **Uci**, **Irma**, **Ara**, **Azal**, **Rani**, **Izzul**, **Kur**, **Ngkal**, **Razy**, **Ryan**, **Fadlan**, **Ridwan**, **Mukhlis**, **Divia**, **Daya**, **Heri**, dan seluruh teman yang tak sempat saya sebutkan namanya. Kepada teman-teman teman-teman **KKN GEL.110**, khususnya **Nagita, Regita, Ana, Dinda,** dan **Dwi** atas kenangan indah. Dan terima kasih kepada diri sendiri yang telah berhasil menyelesaikan skripsi ini dengan tekad dan usaha yang tiada henti. Terima kasih telah menghadapi tantangan dengan keteguhan, tidak menyerah, dan terus berusaha keras.

Penulis menyadari bahwa masih banyak kekurangan dan kesalahan dalam penyusunan skripsi ini. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini dapat memberikan manfaat untuk segala pihak. ***Wasalamu'alaikum warahmatullahi wabarakatuh***

Penulis

Ruslinda

ABSTRAK

RUSLINDA. **Pemodelan *Statistical Downscaling* Menggunakan *Robust Jackknife Ridge Regression* dengan *Estimator Generalized-M* untuk Peramalan Curah Hujan di Kabupaten Pangkep** (dibimbing oleh Sitti Sahriman S.Si., M.Si.).

Latar Belakang. *Statistical Downscalling* (SD) adalah model statistik yang menghubungkan variabel prediktor berskala besar dari *Global Circulation Model* (GCM) dengan variabel respon berskala kecil berupa curah hujan lokal. SD digunakan untuk meramalkan curah hujan lokal berdasarkan data presipitasi dari GCM. Namun, data GCM memiliki dimensi besar yang menyebabkan masalah multikolinearitas dan pencilan, sehingga analisis menjadi tidak akurat. Untuk mengatasi masalah ini, menggunakan metode *Robust Jackknife Ridge Regression* (RJRR) estimasi gabungan dari *Jackknife Ridge Regression* (JRR) dan estimator *Generalized-M*. JRR mengatasi multikolinearitas melalui prosedur *Jackknife* untuk mengurangi bias, sedangkan estimator GM untuk mengatasi pencilan, sehingga meningkatkan akurasi peramalan curah hujan lokal. **Tujuan**. Penelitian ini bertujuan untuk mendapatkan model SD menggunakan RJRR dan mendapatkan hasil peramalan curah hujan di Kabupaten Pangkep periode Januari -Desember 2023. **Metode**. Metode penelitian ini terdiri dari 3 tahap, yakni:1) penentuan variabel *dummy* berdasarkan metode *k-means*; 2) pendugaan model RJRR dengan estimator GM pada data curah hujan; 3) peramalan curah hujan di Kabupaten Pangkep. **Hasil**. Hasil analisis menunjukkan peramalan curah hujan menggunakan SD dengan RJRR Estimator GM mampu menjelaskan keragaman data curah hujan berdasarkan nilai koefisien determinasi (R^2) sebesar (93%). Dengan adanya variabel *dummy*, nilai koefisien determinasi (R^2) meningkat sebesar 25%. Selain itu, nilai RMSEP menurun dari 161.44 menjadi 70.17 menunjukkan hasil ramalan lebih mendekati data aktual. Nilai korelasi yang tinggi sebesar 0.965 menunjukkan kemampuan model dalam menangkap pola curah hujan aktual dengan baik. **Kesimpulan**. Hasil peramalan model RJRR GM *dummy* periode Januari – Desember Tahun 2023 memiliki akurasi lebih tinggi dan mengikuti pola curah hujan aktual dengan baik.

Kata Kunci: *Statistical Downscaling*, *Robust Jackknife Ridge Regression*, *Generalized-M*, *Dummy*, Multikolinearitas, Pencilan.

ABSTRACT

RUSLINDA. **Statistical Downscaling Modeling Using *Robust* Jackknife Ridge Regression with Generalized-M Estimator for Rainfall Forecasting in Kabupaten Pangkep** (*Supervised by Sitti Sahriman, S.Si., M.Si*.).

Introduction. Statistical Downscaling (SD) is a statistical model that links large-scale predictor variables from the Global Circulation Model (GCM) with small-scale response variables, such as local precipitation. SD is used to forecast local rainfall based on precipitation data from GCM. However, GCM data have large dimensions that can cause issues with multicollinearity and outliers, leading to inaccurate analyses. To address these issues, the *Robust* Jackknife Ridge Regression (RJRR) method, a combination of Jackknife Ridge Regression (JRR) and Generalized-M estimator, is used. JRR addresses multicollinearity through the Jackknife procedure, reducing bias, while the GM estimator deals with outliers, thereby enhancing the accuracy of local rainfall forecasts. **Purpose**. This study aims to develop an SD model using RJRR and to forecast rainfall in Kabupaten Pangkep for the period of January to December 2023. **Method**. This study consisted of three stages: 1) adding dummy variables using k-means; 2) estimating the *Robust* Jackknife Ridge Regression model with the Generalized-M estimator; 3) forecasting rainfall. **Results**. The analysis shows that rainfall forecasting using SD with the RJRR Estimator GM can explain the variability in rainfall data, as indicated by a coefficient of determination (R^2) of 93%. With the inclusion of a dummy variable, the coefficient of determination(R^2) increased by 25%. Additionally, the RMSEP value decreased from 161.44 to 70.17, indicating that the forecasted results are closer to the actual data. A high correlation value of 0.965 demonstrates the model's ability to accurately capture the actual rainfall patterns. **Conclusion**. The RJRR GM Dummy model had higher accuracy, particularly in predicting extreme rainfall, which could enhance salt production in Kabupaten Pangkep.

Keywords: Statistical Downscaling, *Robust* Jackknife Ridge Regression, Generalized-M, Dummy, Multicollinearity, Outlier.

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Statistical Downscalling	Metode untuk memperkirakan data iklim atau cuaca berskala regional atau lokal.
Global Circulation Model	Data simulasi kondisi atmosfer tentang perubahan iklim
Presipitasi	Mencakup curah hujan serta bentuk-bentuk lain seperti salju dan hujan es.
Robust	Metode analisis yang andal dan akurat dengan memberikan pembobot pada data yang mengandung pencilan.
Ridge	Metode statistik yang digunakan untuk mengatasi data yang mengalami multikolinearitas
Jackknife	Adanya modifikasi dari *ridge* yang akan menghasilkan estimasi bias yang lebih kecil
Variabel *Dummy*	Variabel numerik yang digunakan dalam analisis regresi untuk mewakili kelompok
Estimasi	Proses memperkirakan nilai dari parameter populasi berdasarkan data sampel
Parameter	Nilai yang menggambarkan sifat atau karakteristik suatu populasi
Generalized-M	Estimasi regresi *robust* untuk yang dapat mengatasi pencilan pada variabel prediktor dan variabel respon.
Multikolinearitas	Kondisi dengan hubungan (korelasi) tinggi antara dua atau lebih variabel prediktor.
Pencilan	Suatu pengamatan yang nilainya jauh dari pusat data dan berpengaruh terhadap koefisien regresi
Ortogonal	Kondisi dua atau lebih variabel prediktor yang memiliki nilai korelasi sama dengan nol.
Leverage Value	Nilai yang mengukur pengaruh observasi pada model. Nilai tinggi menunjukkan potensi besar mempengaruhi hasil analisis.
Root Mean Square Error Prediction	Digunakan untuk mengukur akurasi model prediksi dengan menghitung akar kuadrat rata-rata kesalahan prediksi.
Pemusatan	Perbedaan antara masing-masing pengamatan dan rata-rata semua pengamatan untuk setiap variabel.
	Gambaran pengamatan pada kesatuan (unit) standar deviasi dari pengamatan untuk variabel.
	Pengelompokan data pada suatu pengamatan
	Proses pengulangan langkah-langkah dalam model hingga mencapai hasil yang baik atau konvergen.

DAFTAR LAMBANG DAN SINGKATAN

Lambang/Singkatan	Arti dan Penjelasan
y_i	Variabel respon pada pengamatan ke-i
x_i	Variabel prediktor pada pengamatan ke-i
ε_i	Nilai residual pada pengamatan ke-i
$\hat{y}_i$	Nilai dugaan data i ke-i , $i = 1,2, \ldots , n$
β_j	Koefisien regresi dari variabel prediktor ke-j
β_0	Koefisien konstanta
p	Jumlah variabel bebas
n	Jumlah Pengamatan
x_{ij}	Nilai pengamatan ke-i dari variabel prediktor ke-j
H	Matriks *hat*
h_{ii}	Nilai elemen diagonal dari matriks *hat*
t_i	Residual standar pada pengamatan ke-i
$\bar{y}$	Rata-rata dari variabel Y
S_y	Standar Deviasi dari variabel Y
S_{x_j}	standar deviasi X dari prediktor ke-j
$\bar{X}_j$	rata rata dari nilai x_{ij} pada prediktor ke-j
Y^*	Nilai Standarisasi dari variabel respon
X^*	Nilai Standarisasi dari variabel prediktor
P	Vektor Eigen
Λ	Elemen diagonal dari nilai eigen berkuran $p \times p$
Z	Matriks prediktor yang telah ortogonal
W	Pembobot Huber *Generalized-M*
$\hat{\sigma}_{GM}$	Skala estimasi *Generalized-M*
π_i	Pembobot *Schweppe*
ψ	Fungsi pengaruh
$\boldsymbol{\hat{\beta}_{GM}}$	Koefisien regresi dari model GM
$\boldsymbol{\hat{\alpha}_{R-i}}$	Estimator *ridge regression* tanpa pengamatan ke-i
k	Tetapan bias k
$\hat{\sigma}_{GM}{}^2$	Varians *robust* GM
R_j^2	Koefisien determinasi dari peubah prediktor ke-j
$\boldsymbol{\hat{\alpha}_{RJRR}}$	Estimator *robust jackknife ridge regression*
VIF	*Variance Inflantion Factor*
RMSEP	*Root Mean Square Error Prediction*
	Difference In Fit Standardized
	Within Cluster Sum of Squares
	Robust Jackknife Ridge Regression
	Generalized-M

DAFTAR ISI

Halaman

DAFTAR TABEL

Tabel **Halaman**

DAFTAR GAMBAR

Gambar **Halaman**

DAFTAR LAMPIRAN

Lampiran **Halaman**

BAB I
PENDAHULUAN

1.1 Latar Belakang

Perubahan iklim merupakan fenomena global yang memiliki dampak yang signifikan terhadap aspek kehidupan manusia dan lingkungan. Perubahan iklim bisa dilihat dengan perubahan ataupun pergerakan cuaca yang terjadi dalam jangka panjang. Salah satu dampak yang paling terasa adalah perubahan dalam pola dan intensitas curah hujan ekstrim, yang berdampak langsung pada berbagai sektor termasuk pertanian, perikanan, kesehatan dan industri (Sirajuddin dkk, 2023). Curah hujan merupakan jumlah air hujan yang diukur saat jatuh di permukaan sebelum mengalami penguapan, atau terserap ke dalam tanah. Sedangkan, Intensitas curah hujan merupakan volume air hujan yang terkumpul di permukaan datar dalam rentang waktu tertentu, diukur dalam satuan ketinggian milimeter (mm). Satuan milimeter artinya tempat yang datar dengan luas satu meter persegi, dapat menampung air hujan setinggi satu milimeter atau sebanyak satu liter (Ruswanti, 2020).

Sulawesi Selatan mempunyai iklim tropis musim hujan dan musim kemarau pada periode tertentu. Pada tahun 2024 tingkat curah hujan cukup tinggi di daerah Sulawesi Selatan bagian barat salah satunya adalah Kabupaten Pangkep (BMKG, 2024). Memiliki musim hujan yang berlangsung lama dan musim kemarau yang singkat karena ketidakstabilan iklim, dikarenakan oleh keadaan geografi daerah tersebut, dengan rata-rata curah hujan 3.200 mm per tahun (BPS, 2024). Kabupaten Pangkep salah satu wilayah penghasil produksi garam, jika musim hujan berlangsung lama maka hal ini berpengaruh pada penurunan produksi garam lokal oleh petambak garam, serta tidak stabilnya harga garam yang bergantung pada matahari (Sahriman dkk, 2024).

Data dari Badan Pusat Statistik (BPS) dan Kementerian Kelautan dan Perikanan (KKP), pada tahun 2019 hingga 2022 produksi garam menurun secara signifikan dari 140.338 ton, 1.283,5 ton dan 45 ton jauh dari kebutuhan garam nasional 4.5 juta ton. Oleh karena itu, diperlukan upaya strategis untuk mendukung peningkatan produksi dan kualitas garam yaitu dengan peramalan curah hujan. Kurangnya informasi dan peramalan curah hujan yang tidak akurat mempengaruhi mata pencaharian penduduk. Dengan peramalan curah hujan memungkinkan produsen garam di Kabupaten Pangkep dapat mengantisipasi periode curah hujan tinggi dan merencanakan produksi yang lebih efektif. Sehingga, meningkatkan produksi garam dan kesejahteraan ekonomi masyarakat di Kabupaten Pangkep.

...erupakan kejadian kompleks karena melibatkan topografi, ...lan atmosfer sehingga mempersulit peramalan curah hujan di ...h karena itu, dibutuhkan model peramalan curah hujan lokal ...angkan sirkulasi atmosfer global yang diperoleh dari luaran ...*odel* (Wilby dan Wigley, 1997). GCM merupakan alat penduga ...aca yang berfungsi sebagai sumber informasi untuk menilai ... iklim. Namun, informasi GCM masih berskala global dan belum dapat mempertimbangkan wilayah yang bersifat lokal dengan resolusi lebih tinggi dari data luaran GCM (Sahriman dan Yulianti, 2021). Oleh karena itu, untuk dapat menggunakan data GCM dalam peramalan curah hujan lokal dapat menggunakan teknik *Statistical Downscalling* (Wigena, 2011).

Statistical Downscalling adalah model statistika yang dapat menggambarkan hubungan antara variabel prediktor berskala besar berupa iklim luaran GCM dengan variabel respon berskala kecil berupa curah hujan lokal dalam periode waktu tertentu (Wigena, 2015). Namun, data luaran GCM tidak dapat digunakan secara langsung sebagai variabel prediktor. Data luaran GCM merupakan data spasial dan temporal, dengan kemungkinan terdapat korelasi spasial dan temporal antar data grid yang berbeda dalam satu wilayah (Wigena, 2006). Sehingga permasalahan yang sering terjadi pada data luaran GCM yaitu terjadinya multikolinearitas antar variabel prediktor, serta adanya kasus pencilan pada data (Upa dkk, 2021).

Multikolinearitas adalah model regresi dengan korelasi tinggi yang terjadi di antara beberapa atau semua variabel prediktor dalam data. Adanya multikolinearitas menyebabkan kesalahan eror pada pendugaan parameter menjadi lebih besar sehingga hasil pendugaan tidak akurat (Shresta, 2020). Salah satu metode yang digunakan untuk mengatasi multikolinearitas yaitu *Ridge Regression*. Metode RR diperkenalkan oleh Hoerl dan Kennard yang merupakan modifikasi dari metode kuadrat terkecil dengan menambahkan konstanta bias pada matriks diagonal sehingga memiliki variansi yang minimum (Lestari dkk, 2022). Kelemahan metode RR yaitu nilai bias yang dihasilkan masih besar, sehingga Singht pada tahun 1986 memperkenalkan metode untuk mengurangi bias tersebut dengan menggunakan metode *Jackknife Ridge Regression* (Malau, 2021).

Pencilan adalah suatu pengamatan yang nilainya jauh dari pusat data dan berpengaruh terhadap koefisien regresi (Daniel, 2019). Terdapat pencilan pada variabel prediktor yang akan mengakibatkan pengujian statistik kurang tepat. Ada dua cara untuk mengatasi pencilan, cara pertama yaitu dengan menghapus data, namun dikhawatirkan terdapat informasi penting pada pencilan tersebut. Cara kedua yaitu dengan mempertahankan seluruh data, untuk mengatasi masalah pencilan menggunakan analisis regresi *robust* (Nurbaroqah, 2022). Dalam regresi *robust* terdapat beberapa metode estimasi, salah satunya adalah Estimasi GM. GM-estimator merupakan pengembangan dari estimasi M ketika estimasi M kurang resisten terhadap pencilan pada variabel prediktor. Hal yang mendasari GM estimator adalah untuk membatasi pengaruh pencilan dan menambahkan fungsi pembobot baru pada variabel prediktor (Junus, 2021).

Abqorunnisa (2022) melakukan penelitian yang membahas mengenai metode *Jackniffe Ridge Regression* untuk mengatasi multikolinearitas persamaan. Selanjutnya, Upa dan Sahriman (2021) melakukan mbahas mengenai metode pemodelan SD dengan regresi e *Ridge Dummy* berbasis *K-means* untuk mengatasi nelitian lainnya yaitu Aristiarto dkk (2023) mengatasi pencilan si *robust* estimasi GM.

Berdasarkan uraian sebelumnya, metode yang digunakan oleh peneliti sebelumnya belum mempertimbangkan efek pencilan dan multikolinearitas secara bersamaan. Oleh karena itu, pada penelitian ini akan menggunakan pemodelan SD untuk mengatasi masalah multikolinearitas dan pencilan secara bersamaan yaitu menggunakan metode *Robust Jackknife Ridge Regression* (RJRR) pada pemodelan SD.

1.2 Batasan Masalah

Pada penelitian ini, penulis memberikan batasan masalah sebagai berikut:

1. Data curah hujan yang digunakan adalah data curah hujan Kabupaten Pangkep Periode Januari 1999 - Desember 2023.
2. Metode estimasi parameter model menggunakan Generalized M-*estimator* dengan fungsi pembobot Huber.
3. Menambahkan variabel *dummy k-means* pada data.

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Memperoleh model *Statistical Downscaling* menggunakan *Robust Jackknife Ridge Regression* pada data curah hujan di Kabupaten Pangkep.
2. Mendapatkan hasil peramalan curah hujan pada data curah hujan di Kabupaten Pangkep menggunakan model *Statistical Downscaling* dengan metode *Robust Jackknife Ridge Regression*.

1.4 Manfaat Penelitian

Penelitian ini bermanfaat menambah wawasan keilmuan dan sebagai salah satu sumber referensi kepada pembaca tentang cara memodelkan SD dengan *Robust Jackknife Ridge Regression* dalam peramalan curah hujan yang dapat dijadikan sebagai acuan dan literatur tambahan untuk penelitian selanjutnya yang memiliki relevansi dengan penelitian ini. Selain itu, dapat bermanfaat sebagai sumber informasi tentang curah hujan di Kabupaten Pangkep serta memberikan kontribusi terhadap pengembangan ilmu pengetahuan dan teknologi kedepannya.

1.5 Teori

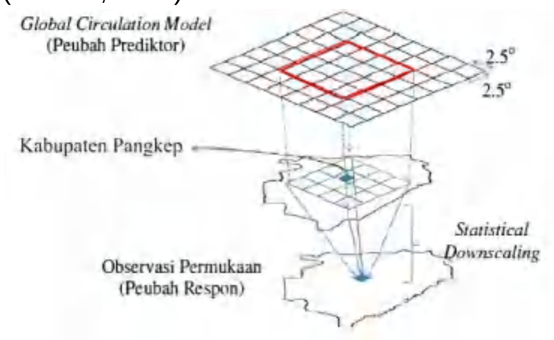
1.5.1 *Global Circulation Model* dan *Statistical Downscalling*

General Circulation Model (GCM) adalah salah satu alat yang penting dalam studi keragaman iklim dan perubahan iklim (Zorita dan Storch, 1999). GCM merupakan gambaran hubungan matematik dari berbagai interaksi fisika, kimia dan dinamika atmosfer bumi. Model ini mempunyai luaran grid yang berukuran 100-500 km
g dan bujurnya (Stroch dkk 1993) serta model ini mampu
ola iklim dari masa lampau, sekarang dan yang akan datang
M dapat mempresentasikan keadaan iklim berskala global
ngan baik, namun tidak dalam berskala lokal/regional (resolusi
tu, diperlukan suatu teknik untuk menduga variabel iklim skala
akurasi tinggi yaitu menggunakan teknik *Statistical Downscaling*

Model *Statistical Downscaling* (SD) didasarkan pada hubungan antara grid global (sebagai prediktor) dengan grid lokal (sebagai respon) yang dijelaskan melalui model statistik dalam periode waktu tertentu (Zorita & Storch, 1999) dan berguna untuk mengonversi anomali-anomali skala global menjadi anomali dari berbagai variabel iklim lokal. Model SD akan memberikan hasil yang baik dengan syarat yaitu hubungan antara variabel respon dengan prediktor harus berkorelasi tinggi untuk menjelaskan keragaman iklim lokal dengan baik, variabel prediktor harus disimulasikan dengan baik oleh GCM, dan hubungan antara variabel respon dengan prediktor tidak berubah dengan adanya perubahan waktu serta tetap sama meskipun ada perubahan iklim di masa depan (Wigena dkk 2015). Bentuk umum model SD dapat disajikan pada Persamaan (1) berikut (Wigena, 2006):

$$y = f(\boldsymbol{X}) \tag{1}$$

dengan $y_{(nx1)}$ adalah variabel-variabel iklim lokal (misalnya: curah hujan), f dapat berbentuk persamaan linier maupun non-linier, $\boldsymbol{X}_{nxp}$ adalah variabel-variabel luaran GCM (misalnya: presipitasi), *n* adalah banyaknya waktu (misalnya: harian atau bulanan), *p* adalah banyaknya grid domain GCM. Ilustrasi proses SD dapat dilihat pada Gambar 2.1 (Sutikno, 2008).



Gambar 1. Ilustrasi Proses *Statistical Downscaling* (Sumber: Sutikno, 2008)

1.5.2 Multikolinearitas

Istilah multikolinearitas awalnya ditemukan oleh Ragnar Frisch pada tahun 1943. Multikolinearitas artinya terdapat hubungan linear atau korelasi yang tinggi antar variabel prediktor dalam model regresi. Salah satu cara untuk mengetahui keberadaan multikolinearitas dalam suatu model regresi adalah dengan menggunakan nilai *tolerance* atau *Variance Inflantion Factor* (VIF) pada Persamaan (2) berikut (Gujarati, 1995).

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

dan p adalah banyaknya variabel prediktor sedangkan R_j^2 erminasi dari variabel prediktor yang diregresikan terhadap nya. Menurut Wirnancy (2015), nilai $VIF > 10$ mengindikasikan as sehingga dilakukan reduksi dimensi.

Menurut Tinungki (2019) jika terjadi masalah multikolinearitas pada model regresi maka dibutuhkan sebuah metode untuk menanganinya, yaitu dengan menggunakan metode yang dikembangkan dari regresi *Ridge* yaitu *Jackknife Ridge.*

1.5.3 Pencilan

Pencilan merupakan suatu pengamatan yang menyimpang cukup jauh dari pengamatan lainnya sehingga menimbulkan kecurigaan bahwa pengamatan tersebut berasal dari distribusi data yang berbeda (Hawkins dalam Sujatmiko, 2005). Untuk mendeteksi pencilan pada variabel Y menggunakan metode DFFITS dan mengidentifikasi pencilan pada variabel X menggunakan metode *Leverage Value.*

1. Metode DFFITS

   Pengamatan pencilan dapat dengan mudah terlihat dengan menggunakan beberapa plot sederhana, seperti *scatter plot, steam and leaf, boxplot*, dan sebagainya, sedangkan pada data *multivariate* identifikasi pencilan umumnya didasarkan pada metode *Difference In Fit Standardized* (DFFITS). Metode *DFFITS* merupakan gabungan antara metode *lavarage* (h_{ii}) dan t_i. t_i adalah *R-student (Studentized Deleted Residual)* untuk kasus ke-*i.* Rumus untuk menghitung nilai DFFITS yaitu (Mardiana,2019):

   $$DFFITS = t_i\left(\frac{h_{ii}}{1-h_{ii}}\right)^{\frac{1}{2}} \qquad (3)$$

   dengan

   $$t_i = e_i\sqrt{\frac{n-p-1}{SSE(1-h_{ii})-e_i^2}}$$

   dengan e_i adalah residual ke-*i* dan adalah SSE jumlah kuadrat residual. Suatu data akan dikatakan pencilan apabila $|DFFITS| > 2\sqrt{\frac{p}{n}}$ dengan p adalah banyaknya variabel dalam model dan n adalah banyaknya pengamatan.

2. Metode *Leverage Value* (h_{ii})

   Metode lain untuk mendeteksi pencilan pada variabel prediktor adalah metode *leverage value*. Nilai *leverage* h_{ii} yang bernilai diantara $0 < h_{ii} \leq 1$ merupakan elemen-elemen diagonal dari matriks H yang merupakan matriks berukuran $n\ x\ n$, yaitu:

   $$H = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \qquad (4)$$

   Nilai elemen-elemen diagonal h_{ii} untuk model regresi dapat dihitung dengan persamaan:

   $$h_{ii} = \boldsymbol{x_i}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x_i}', \quad i = 1,2,\ldots,n \qquad (5)$$

   ...lin dan Welsch (1978), suatu pengamatan yang diduga sebagai ...*oint* adalah pengamatan yang memiliki nilai *leverage value* ...n m adalah jumlah variabel prediktor ditambah konstan dan n ...ngamatan.

1.5.4 Variabel *Dummy* dengan *K-Means Clustering*

Variabel *dummy* atau variabel boneka adalah variabel buatan yang dibuat untuk mewakili atribut dengan dua kategori atau kategori yang berbeda. Variabel *dummy* hanya memiliki dua nilai, yaitu 0 dan 1 guna menggambarkan keanggotaan pada kategori yang saling eksklusif dan menyeluruh. Setiap data curah hujan yang berada dalam kelompok cluster yang sama akan bernilai 1 pada variabel *dummy* tersebut. Sementara itu, data curah hujan yang berada diluar kelompok tersebut akan diberikan nilai 0 pada variabel *dummy*. Variabel *dummy* menentukan angka 0 dan 1. (Parjiono dkk,2018).

Penambahan variabel *dummy* ke dalam model dapat mengatasi masalah keheterogenan ragam sisaan model. Variabel *dummy* ditambahkan berdasarkan hasil pengelompokan model dengan menggunakan teknik *cluster* (Sahriman dkk., 2014). Penentuan variabel *dummy* dengan k-means merupakan salah satu metode *cluster non-hirarki*. MacQueen menyarankan penggunaan *k-means* untuk menjelaskan algoritma dalam penentuan suatu objek ke dalam *cluster* tertentu berdasarkan rataan terdekat (Mattjik & Sumertajaya, 2011).

Penentuan jumlah *cluster* yang optimal telah dilakukan oleh Sahriman dkk. (2014) pada data curah hujan di Indramayu dan Upa dkk (2021) pada data curah hujan di Kabupaten Pangkep dengan melihat kelompok warna yang dominan pada curah hujan. Hasilnya adalah terbentuk data curah hujan dengan 4 kelompok warna sehingga terdapat 4 *cluster* yang optimal. Selain itu, salah satu metode yang dapat digunakan untuk menentukan jumlah k *cluster* adalah metode *elbow* dilihat dari nilai *Within Cluster Sum of Squares* (WSS) yang rendah. Jumlah *cluster* yang diambil adalah titik yang membentuk siku, biasa disajikan dalam bentuk grafik untuk mengetahui lebih jelas siku yang terbentuk. Maksud dari titik yang membentuk siku adalah pada titik yang terjadi signifikan penurunan antara 2 titik *cluster* dan kemudian diikuti oleh nilai yang relatif konstan (Maori, 2023).

1.5.5 Metode Pemusatan dan Penskalaan

Pemusatan dan penskalaan adalah bagian dari transformasi data dengan membakukan (*standardized)* variabel. Modifikasi sederhana dari pembakuan atau standarisasi variabel ini adalah transformasi korelasi (*correlation transformation*). Pemusatan merupakan perbedaan antara masing-masing pengamatan dengan rata-rata dari semua pengamatan untuk variabel. Sedangkan penskalaan menggambarkan pengamatan pada kesatuan (unit) standar deviasi dari pengamatan untuk variabel (Kutner dkk, 2005). Tujuan utama dari transformasi ini untuk mempermudah analisis data yang memiliki satuan yang berbeda-beda, sehingga

ada dalam skala yang sesuai sehingga model lebih efektif.

orelasi merupakan fungsi sederhana dari pembakuan variabel

erikut (Montgomery (1992):

$$y_i^* = \frac{y_i - \bar{y}}{\sqrt{n-1}\,(S_y)} \quad \text{dengan} \quad S_Y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n-1}} \tag{6}$$

$$X_{ij}^{*} = \frac{X_{ij}-\bar{X}_j}{\sqrt{n-1}\,(S_{xj})} \quad \text{dengan} \quad S_j = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{i_j}-\hat{X}_j\right)^2}{n-1}} \tag{7}$$

Berdasarkan transformasi korelasi untuk variabel respon Y_i^* dan variabel bebas X_{ij}^* yang didefinisikan dengan transformasi korelasi pada Persamaan (6) dan (7) diperoleh model regresi sebagai berikut.

$$Y_i^* = \beta_1^* X_{i1}^* + \beta_2^* X_{i2}^* + \cdots + \beta_p^* X_{ip}^* + \varepsilon_i^* \tag{8}$$

Persamaan (8) disebut sebagai model regresi yang baku (*standardized regression model).* Diantara penduga parameter $\beta_1^*, \beta_2^*, \cdots, \beta_p^*$ pada model regresi baku dengan penduga parameter awal dari $\beta_1, \beta_2, \cdots, \beta_p$ pada model regresi linier berganda terdapat hubungan linear (Kutner 2005).

1.5.6 Regresi *Ridge*

Regresi *Ridge* atau *Ridge Regression* adalah suatu metode yang diperkenalkan oleh Hoerl dan Kennard pada tahun 1962. Metode ini digunakan untuk mengatasi masalah multikolinearitas. Disamping itu, regresi *ridge* merupakan penduga parameter yang bias dan memiliki varians yang minimum, sehingga parameter yang dihasilkan lebih stabil dan memiliki nilai presisi yang lebih baik dibandingkan parameter *unbiased* pada metode *Ordinary Least Square* (Putri & Anggorowati, 2017).

Persamaan umum regresi *ridge* diperoleh dengan mereduksi matriks $X^{*\prime}X^*$, dimana matriks tersebut simetris, sehingga terdapat matriks ortogonal $\boldsymbol{P}$. Sehingga persamaan berubah menjadi

$$\begin{aligned} \boldsymbol{P}'\left(\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\right)\boldsymbol{P} &= \boldsymbol{\Lambda} \\ \boldsymbol{P}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\boldsymbol{P} &= \boldsymbol{\Lambda} \\ (\boldsymbol{X}^*\boldsymbol{P})'\boldsymbol{X}^*\boldsymbol{P} &= \boldsymbol{\Lambda} \\ \boldsymbol{Z}'\boldsymbol{Z} &= \boldsymbol{\Lambda} \end{aligned} \tag{9}$$

dengan $\boldsymbol{\Lambda}$ adalah matriks diagonal berukuran $p \times p$ di mana elemen diagonalnya adalah nilai eigen dari matriks $X^{*\prime}X^*$ dan $\boldsymbol{P}$ adalah matriks ortogonal berukuran $p \times p$ yang elemennya adalah vector eigen dari matriks $X^{*\prime}X^*$, dengan $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}$, sehingga diperoleh persamaan baru linier berganda dapat ditulis:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{y} &= (\boldsymbol{X}^*\boldsymbol{P})(\boldsymbol{P}'\boldsymbol{\beta}) + \boldsymbol{\varepsilon} \\ \boldsymbol{y} &= \boldsymbol{Z}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \end{aligned} \tag{10}$$

dengan $\boldsymbol{Z} = \boldsymbol{X}^*\boldsymbol{P}$, $\boldsymbol{\alpha} = \boldsymbol{P}'\boldsymbol{\beta}$ dan $\boldsymbol{\beta} = \boldsymbol{P}\boldsymbol{\alpha}$. (Montgomery dan Peek, 2012).

D... an MKT, Persamaan (9) estimator $\hat{\boldsymbol{\alpha}}$ berubah menjadi

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_{ols} &= (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{y} \\ &= \boldsymbol{\Lambda}^{-1}\boldsymbol{Z}'\boldsymbol{y} \end{aligned} \tag{11}$$

...at dibentuk menjadi

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_{ols} &= [(\boldsymbol{X}^*\boldsymbol{P})'(\boldsymbol{X}^*\boldsymbol{P})^{-1}](\boldsymbol{X}^*\boldsymbol{P})'\boldsymbol{y} \\ &= (\boldsymbol{P}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\boldsymbol{P})^{-1}\,\boldsymbol{P}'\boldsymbol{X}^{*\prime}\boldsymbol{y} \\ &= (\boldsymbol{P}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\boldsymbol{P})^{-1}\,\boldsymbol{P}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\hat{\boldsymbol{\beta}} \end{aligned}$$

$$= \quad \boldsymbol{P}'(\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})^{-1}(\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})\widehat{\boldsymbol{\beta}}$$
$$= \quad \boldsymbol{P}'\widehat{\boldsymbol{\beta}} \qquad (12)$$

Berdasarkan persamaan kanonik diperoleh parameter $\widehat{\boldsymbol{\beta}}$ pada persamaan berikut.

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{P}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{ols}} \qquad (13)$$

Selanjutnya, untuk mendapatkan estimator *ridge regression* menggunakan pengali *Lagrange*, dimana nilai $\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}}$ yang minimum dengan fungsi tujuan dengan syarat $\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}}'\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} - c^2 = 0$.

$$\begin{aligned} \boldsymbol{F} = &\quad (\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\alpha}_{\boldsymbol{R}})'(\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\alpha}_{\boldsymbol{R}}) + k(\boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{\alpha}_{\boldsymbol{R}} - c^2) \\ = &\quad \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{Z}\boldsymbol{\alpha}_{\boldsymbol{R}} - \boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{Z}'\boldsymbol{y} + \boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{\alpha}_{\boldsymbol{R}} + k\boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{\alpha}_{\boldsymbol{R}} - c^2) \\ = &\quad \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{Z}'\boldsymbol{y} + \boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{\alpha}_{\boldsymbol{R}} + k\boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{\alpha}_{\boldsymbol{R}} - c^2) \qquad (14) \end{aligned}$$

Solusi untuk F akan minimum ketika $\frac{\partial F}{\partial \widehat{\alpha}_R} = 0$, maka

$$\frac{\partial \boldsymbol{F}}{\partial \widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}}} = \left.\frac{\partial(\boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{Z}'\boldsymbol{y} + \boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{\alpha}_{\boldsymbol{R}} + k\boldsymbol{\alpha}_{\boldsymbol{R}}'\boldsymbol{\alpha}_{\boldsymbol{R}} - c^2)}{\partial \widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}}}\right|_{\alpha_R = \widehat{\alpha}_R}$$

$$0 \quad = -2\boldsymbol{Z}'\boldsymbol{y} + 2\boldsymbol{Z}'\boldsymbol{Z}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + 2k\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}}$$
$$0 \quad = -\boldsymbol{Z}'\boldsymbol{y} + \boldsymbol{Z}'\boldsymbol{Z}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + k\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}}$$
$$\boldsymbol{Z}'\boldsymbol{Z}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + \boldsymbol{k}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} = \boldsymbol{Z}'\boldsymbol{y}$$
$$(\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} = \boldsymbol{Z}'\boldsymbol{y}$$
$$\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} = (\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1}\boldsymbol{Z}'\boldsymbol{y}$$

Sehingga estimator *ridge regression* adalah

$$\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} = (\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1}\boldsymbol{Z}'\boldsymbol{y} \qquad (15)$$

dimana Z adalah variabel prediktor X^* yang ortogonal dan, y^* adalah variabel respon Y yang telah ditransformasi, k adalah sebuah bilangan positif atau nilainya lebih dari nol dan I adalah matriks identitas.

Dengan $\boldsymbol{A} = \boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I},$ sehingga Persamaan (12) diubah menjadi

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} &= \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{y} \\ &= \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{Z}\widehat{\boldsymbol{\alpha}} \\ &= \boldsymbol{A}^{-1}(\boldsymbol{A} - k\boldsymbol{I})\widehat{\boldsymbol{\alpha}} \\ &= \boldsymbol{A}^{-1}\boldsymbol{A} - \boldsymbol{A}^{-1}k\boldsymbol{I})\widehat{\boldsymbol{\alpha}} \\ &= (\boldsymbol{I} - \boldsymbol{A}^{-1}k\boldsymbol{I})\widehat{\boldsymbol{\alpha}} \qquad (16) \end{aligned}$$

1.5.7 Estimasi *Generalized-M*

Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari residual tidak normal atau adanya pencilan yang mempengaruhi model. Metode ini digunakan untuk mendeteksi *outlier* atau pencilan dan memberikan hasil yang resisten terhadap adanya pencilan (Hidayatulloh dkk, 2015). Efisiensi menjelaskan seberapa baiknya suatu teknik *robust* sebanding dengan *Least Square* ada data ... makin tinggi efisiensi dan *breakdown point* dari suatu estimator ... en terhadap pencilan.

... metode estimasi dalam mengatasi regresi *robust* salah satu ... Estimasi M yang diperkenalkan oleh Huber. Menurut Huber ... ensitive terhadap pencilan baik dari variabel X dan variabel Y. ... 975) memperkenalkan pengembangan dari estimator M yaitu

Generalized-M. Metode estimasi GM adalah metode yang mampu menyelesaikan *high-breakdown point* sampai 50% dari data. Ide dasar yang melatarbelakangi estimasi GM adalah untuk membatasi pengaruh dari pencilan dengan pembobot w_i (Junus,2021).

Estimasi M dapat diperoleh dengan meminimumkan jumlahan skala residual. Persamaan dari Estimasi M adalah sebagai berikut:

$$\sum_{i=1}^{n} \rho(e_i) = \sum_{i=1}^{n} \rho\left(y_i - \sum_{j=1}^{p} Z_{ij}\beta_j\right) \tag{17}$$

Model Persamaan (17) minimum ketika turunan pertama dari pembobot ρ secara parsial terhadap β_j sama dengan nol. Residual pada fungsi tersebut harus distandarisasi dengan skala *robust* σ. Sehingga persamaan menjadi:

$$\frac{\partial \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma}\right)}{\partial \beta_j} = 0 \tag{18}$$

karena $e_i = y_i - \sum_{j=1}^{p} \beta_j Z_{ij}$, maka $u_i = \frac{e_i}{\sigma}$ sehingga Persamaan (18) dapat dituliskan sebagai berikut:

$$\frac{\partial \sum_{i=1}^{n} \rho(u_i)}{\partial \beta_j} = 0 \tag{19}$$

dengan menggunakan aturan berantai maka diperoleh persamaan:

$$\frac{\partial \sum_{i=1}^{n} \rho(u_i)}{\partial u_i} \times \frac{\partial u_i}{\partial \beta_j} = 0$$

$$\frac{\sum_{i=1}^{n} \rho'(u_i)}{\partial u_i} \times \frac{\partial \left[\frac{\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} Z_{ij}\beta_j\right)}{\sigma}\right]}{\partial \beta_j} = 0$$

$$\sum_{i=1}^{n} \rho'(u_i) \times \left(-\frac{\sum_{i=1}^{n}\sum_{j=1}^{p} Z_{ij}}{\sigma}\right) = 0$$

$$\sum_{i=1}^{n} \rho'\left(y_i - \sum_{j=1}^{p} Z_{ij}\beta_j\right) \times \left(-\frac{\sum_{i=1}^{n}\sum_{j=1}^{p} Z_{ij}}{\sigma}\right) = 0$$

masing-masing kedua ruas dikali dengan $(-\sigma)$ sehingga diperoleh:

$$\sum_{i=1}^{n} Z_{ij}\rho'\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma}\right) = 0 \tag{20}$$

Dengan ψ merupakan fungsi pengaruh turunan pertama ρ dimana $\rho' = \psi$, sehingga persamaan (20) dapat berubah sebagai:

$$\sum_{i=1}^{n} \sum_{j=1}^{p} Z_{ij}\,\psi\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma}\right) = 0 \tag{21}$$

...samaan yang diperoleh pada Persamaan (21) menghasilkan ...*d*-M. Menurut Wilcox (2005), estimator GM merupakan solusi ...al:

$$\sum_{i=1}^{n} Z_{ij}\pi_i\psi\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma\pi_i}\right) = 0 \tag{22}$$

Dimana nilai $j = 1,2, \ldots. k.$ Dengan σ merupakan nilai skala estimasi GM yang digunakan adalah

$$\hat{\sigma}_{GM} = 1.4825\, MAD = 1.4825\, median\{|e_i - median(e_m)|\}$$

Dapat dilihat pada Persamaan (21) dan (22) memiliki persamaan, hanya saja pada Persamaan (22) terdapat π_i yang berfungsi untuk menurunkan bobot pada pengamatan yang memiliki nilai *leverage point* yang tinggi. Untuk estimator GM pembobot yang digunakan adalah pembobot Schweppe:

$$\pi_i = \sqrt{1 - h_{ii}} \tag{23}$$

Fungsi pada Persamaan (23) akan digunakan untuk memperoleh pembobot estimasi $w_i = \frac{\psi\,(u_i/\pi_i)}{u_i/\pi_i}$ yang memiliki nilai di antara 0 dan 1, sehingga diperoleh persamaan:

$$w_i = \frac{\psi\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma\pi_i}\right)}{\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma\pi_i}}$$

$$w_i\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma}\right) = \pi_i\psi\left(\frac{y_i - \sum_{j=1}^{p} Z_{ij}\beta_j}{\sigma\pi_i}\right) \tag{24}$$

sehingga Persamaan (24) berubah menjadi:

$$\sum_{i=1}^{n} Z_{ij}w_i\left(y_i - \sum_{j=1}^{p} Z_{ij}\beta_j\right) \tag{25}$$

Persamaan (25) diselesaikan dengan metode *iteratively reweighted least square* (IRLS), estimasi awal $\hat{\beta}^1$ dan residual e_i^1 diambil dari hasil estimasi metode kuadrat terkecil dari data hasil transformasi, sehingga mendapatkan persamaan:

$$\sum_{i=1}^{n} Z_{ij}w_iy_i - \sum_{j=1}^{p} Z_{ij}^2w_i\beta_j = 0$$

$$\sum_{j=1}^{p} Z_{ij}^2w_i\beta_j = \sum_{i=1}^{n} Z_{ij}w_iy_i \tag{26}$$

dengan $w_i = \frac{\pi_i}{u_i}\psi\left(\frac{u_i}{\pi_i}\right)$, Persamaaan (26) diubah ke dalam bentuk matriks maka akan menjadi :

$$\boldsymbol{Z'WZ\beta = Z'Wy}$$

triks diagonal berukuran $n \times n$ dengan elemen diagonal adalah n adalah banyaknya pengamatan. Untuk setiap iterasi i berikut:

$$\hat{\boldsymbol{\beta}}_{\mathbf{GM}} = (\boldsymbol{Z'WZ})^{-1}\boldsymbol{Z'Wy} \tag{27}$$

Selanjutnya, untuk mendapatkan nilai $\boldsymbol{W}$ dihitung menggunakan $\widehat{\boldsymbol{\beta}}_{\mathbf{GM}}$, skala *robust* $\hat{\sigma}$, dan nilai dari pembobot Schweppe. Iterasi akan berhenti Ketika estimator $\widehat{\boldsymbol{\beta}}_{\mathbf{GM}}{}^{(t)} = \left(\boldsymbol{Z}'\boldsymbol{W}^{(t-1)}\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{W}^{(t-1)}\boldsymbol{y}$ sampai $\sum_{j=1}^{p}|e_i{}^{(t)}|$yang diperoleh konvergen, yaitu $\widehat{\boldsymbol{\beta}}_{GM}^{t+1}$ dengan $\widehat{\boldsymbol{\beta}}_{GM}^{t}$ mendekati 0, dengan t adalah banyaknya iterasi. Selanjutnya, diperoleh estimator dari $\widehat{\boldsymbol{\alpha}}_{GM} = \boldsymbol{P}'\widehat{\boldsymbol{\beta}}_{\mathbf{GM}}$.

Pembobot nilai W yang digunakan adalah pembobot Huber. Fungsi objektif adalah fungsi yang digunakan untuk mencari fungsi pembobot pada regresi *robust.*

$$w(u_i) = \begin{cases} -\dfrac{c}{u_i} & ,jika\; u_i < c \\ 1 & ,jika\; |u_i| \leq c \\ \dfrac{c}{u_i} & ,jika\; ,jika\; u_i > \; c \end{cases}$$

Dimana nilai c pada estimator GM adalah fungsi huber dengan rumus $c = 2\sqrt{p + 1/n}$ dengan p merupakan banyaknya variabel prediktor dengan konstan dan n merupakan banyaknya pengamatan (Bagheri, 2010).

1.5.8 Metode *Jackknife Ridge Regression*

Metode *Jacknife Ridge* pertama kali diperkenalkan oleh Hinkley pada tahun 1977. Metode ini merupakan pengembangan dari metode *Ridge Regression,* yang didapatkan dengan melakukan *resampling jackknife* pada parameter RR. Adanya modifikasi tersebut akan menghasilkan estimasi bias yang lebih kecil (Arrasyid dkk, 2021). Prosedur *Jackknife* yang digunakan adalah penghapusan satu sampel atau sekelompok sampel dari sampel awal dan mengulanginya sebanyak sejumlah sampel data yang ada (Iskandar, 2013). Singh dkk (1986) menyarankan untuk menggunakan teknik JRR seperti pada persamaan

$$\boldsymbol{y}_{-i} = \boldsymbol{Z}_{-i}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + \boldsymbol{\varepsilon}$$

dengan y_{-i} adalah matriks $\boldsymbol{y}$ tanpa pengamatan baris ke$-i$ dan Z_{-i} adalah matriks $\boldsymbol{Z}$ tanpa pengamatan baris ke$-i$. Persamaan berubah menjadi:

$$\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}-\boldsymbol{i}} = (\boldsymbol{Z}'_{-i}\boldsymbol{Z}_{-i} + k\boldsymbol{I})^{-1}\boldsymbol{Z}'_{-i}\boldsymbol{y}_{-i} \tag{28}$$

Persamaan (28) adalah estimator *ridge regression* tanpa pengamatan ke$-i$, dimana $\boldsymbol{Z}'_{-i}\boldsymbol{Z}_{-i} = \boldsymbol{Z}'\boldsymbol{Z} - \boldsymbol{z}_i\boldsymbol{z}_i'$ dan $\boldsymbol{Z}'_{-i}\boldsymbol{y}_{-i} = \boldsymbol{Z}'\boldsymbol{y} - \boldsymbol{z}_i\boldsymbol{y}_i$. Dengan $\boldsymbol{z}_i$ adalah matriks kolom yang elemen di dalamnya merupakan pengamatan ke-i berukuran $p \times 1$. Kemudian melakukan subtitusi ke Persamaan (28).

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}-\boldsymbol{i}} &= (\boldsymbol{Z}'_{-i}\boldsymbol{Z}_{-i} + k\boldsymbol{I})^{-1}\boldsymbol{Z}'_{-i}\boldsymbol{y}_{-i} \\ &= (\boldsymbol{Z}'\boldsymbol{Z} - \boldsymbol{z}_i\boldsymbol{z}_i' + k\boldsymbol{I})^{-1}\boldsymbol{Z}'\boldsymbol{y} - \boldsymbol{z}_i\boldsymbol{y}_i \end{aligned} \tag{29}$$

Dengan menggunakan invers jumlahan matriks seperti yang diperkenalkan oleh ...nifah (2020).

$$(\boldsymbol{A} - \boldsymbol{c}\boldsymbol{c}')^{-1} = \boldsymbol{A}^{-1} + \frac{\boldsymbol{A}^{-1}\boldsymbol{c}\boldsymbol{c}'\boldsymbol{A}^{-1}}{\mathbf{1} - \boldsymbol{c}'\boldsymbol{A}^{-1}\boldsymbol{c}}$$

... dan $\boldsymbol{c} = \boldsymbol{z_i}$, diperoleh:

$$\ldots \boldsymbol{z}_i\boldsymbol{z}_i')^{-1} = (\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1} + \frac{(\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\,(\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1}}{\mathbf{1} - \boldsymbol{z}_i'\,(\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1}\boldsymbol{z}_i}$$

...n (28) berubah menjadi:

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}}_{R-i} &= (\boldsymbol{Z}'_{-i}\boldsymbol{Z}_{-i} + k\boldsymbol{I})^{-1}\boldsymbol{Z}'_{-i}\boldsymbol{y}_i \\
&= (\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I} - \boldsymbol{z}_i\boldsymbol{z}_i')^{-1}\boldsymbol{Z}'\boldsymbol{y} - \boldsymbol{z}_i\boldsymbol{y}_i \\
&= \left[\boldsymbol{A}^{-1} + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\right](\boldsymbol{Z}'\boldsymbol{y} - \boldsymbol{z}_i\boldsymbol{y}_i) \\
&= \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\boldsymbol{Z}'\boldsymbol{y} - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\boldsymbol{z}_i\boldsymbol{y}_i \\
&= \hat{\boldsymbol{\alpha}}_R - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i(\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i)}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}(\boldsymbol{Z}'\boldsymbol{y})}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\boldsymbol{z}_i\boldsymbol{y}_i \\
&= \hat{\boldsymbol{\alpha}}_R - \left(\frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i - \boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\right) + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}(\boldsymbol{Z}'\boldsymbol{y})}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}\,(\boldsymbol{z}_i\boldsymbol{y}_i)}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} \\
&= \hat{\boldsymbol{\alpha}}_R - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}(\boldsymbol{Z}'\boldsymbol{y})}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\boldsymbol{A}^{-1}\,(\boldsymbol{z}_i\boldsymbol{y}_i)}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} \\
&= \hat{\boldsymbol{\alpha}}_R - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} + \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\hat{\boldsymbol{\alpha}}_R}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} \\
&= \hat{\boldsymbol{\alpha}}_R - \left(\frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{y}_i}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i} - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{z}_i'\hat{\boldsymbol{\alpha}}_R}{\boldsymbol{1} - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\right) \\
&= \hat{\boldsymbol{\alpha}}_R - \left(\frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i(\boldsymbol{y}_i - \boldsymbol{z}_i'\hat{\boldsymbol{\alpha}}_R)}{1 - \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i}\right)
\end{aligned}
$$

dengan $e_i = (y_i - z_i'\hat{\boldsymbol{\alpha}}_R)$, dan $v_i = \boldsymbol{z}_i'\boldsymbol{A}^{-1}\boldsymbol{z}_i$, maka uraian di atas dapat disederhanakan sebagai berikut.

$$\hat{\boldsymbol{\alpha}}_{R-i} = \hat{\boldsymbol{\alpha}}_R - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i e_i}{1 - v_i} \tag{30}$$

Menurut Hinkley (1977) mendefinisikan nilai *weighted pseudo-value* sebagai berikut.

$$Q_i = \hat{\boldsymbol{\alpha}}_R + n(1 - \boldsymbol{v}_i)(\hat{\boldsymbol{\alpha}}_R - \hat{\boldsymbol{\alpha}}_{R-1}) \tag{31}$$

Persamaan (30) di substitusi ke Persamaan (31):

$$
\begin{aligned}
Q_i &= \hat{\boldsymbol{\alpha}}_R + n(1 - \boldsymbol{v}_i)(\hat{\boldsymbol{\alpha}}_R - \left(\hat{\boldsymbol{\alpha}}_R - \frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{e}_i}{1 - v_i}\right) \\
&= \hat{\boldsymbol{\alpha}}_R + n(1 - \boldsymbol{v}_i)\frac{\boldsymbol{A}^{-1}\boldsymbol{z}_i e_i}{1 - v_i} \\
&= \hat{\boldsymbol{\alpha}}_R + n\boldsymbol{A}^{-1}\boldsymbol{z}_i\boldsymbol{e}_i
\end{aligned}
$$

Selanjutnya, estimasi parameter dengan estimasi *jackknife ridge* dilakukan dengan mengambil rata-rata *weighted pseudo-value.*

$$\hat{\boldsymbol{\alpha}}_{JRR} = \bar{\boldsymbol{Q}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Q}_i$$

$$\boldsymbol{A}^{-1}\boldsymbol{z}_i e_i)$$

$$\boldsymbol{z}_i(\boldsymbol{y}_i - \boldsymbol{z}_i'\hat{\boldsymbol{\alpha}}_R)$$

$$\boldsymbol{z}_i\boldsymbol{y}_i - \boldsymbol{z}_i\boldsymbol{z}_i'\hat{\boldsymbol{\alpha}}_R)$$

$$\boldsymbol{z}_i\boldsymbol{y}_i - \boldsymbol{A}^{-1}\sum_{i=1}^{n} \boldsymbol{z}_i\boldsymbol{z}_i'\hat{\boldsymbol{\alpha}}_R$$

$$\begin{aligned}\hat{\alpha}_{JRR} &= \hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{Z}\hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} \\ &= \hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + \hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} - \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{Z}\hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} \\ &= \hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + (\boldsymbol{I} - \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{Z})\,\hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} \\ &= \hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} + (\boldsymbol{A}^{-1}k\mathbf{I})\,\hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} \\ &= (\boldsymbol{I} + \boldsymbol{A}^{-1}k\mathbf{I})\,\hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} \end{aligned} \tag{32}$$

Berdasarkan Persamaan (32), $\hat{\boldsymbol{\alpha}}_{\boldsymbol{R}}$ dapat dituliskan sebagai transformasi linier dari $\hat{\boldsymbol{\alpha}}$ pada Persamaan (16). Sehingga estimator *jackknife ridge regression* dapat dituliskan sebagai berikut.

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_{\boldsymbol{JRR}} &= (\boldsymbol{I} + \boldsymbol{A}^{-1}k\boldsymbol{I})(\boldsymbol{I} - \boldsymbol{A}^{-1}k\boldsymbol{I})\hat{\boldsymbol{\alpha}} \\ &= [\boldsymbol{I} - (\boldsymbol{A}^{-1}k\boldsymbol{I})^{2}]\hat{\boldsymbol{\alpha}} \\ &= [\boldsymbol{I} - k\boldsymbol{I}(\boldsymbol{Z}'\boldsymbol{Z} + k\boldsymbol{I})^{-1}]^{2}\hat{\boldsymbol{\alpha}} \end{aligned} \tag{33}$$

dan estimator *jackknife ridge regression* untuk $\hat{\boldsymbol{\beta}}$ adalah:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{JRR}} = \boldsymbol{P}\,\hat{\boldsymbol{\alpha}}_{\boldsymbol{JRR}} \tag{34}$$

1.5.9 *Robust Jackknife Ridge Regression*

Metode *Robust Jackknife Ridge* yang merupakan gabungan dari regresi *ridge* dan regresi *robust* untuk mngatasi multikolinearitas dan pencilan secara bersamaan. Estimator untuk regresi *Robust Jackknife Ridge* dengan estimator GM akan diperoleh dengan mengganti nilai $\hat{\boldsymbol{\alpha}}$ pada regresi *jackknife ridge* menjadi $\hat{\boldsymbol{\alpha}}_{\boldsymbol{GM}}$ dan dapat didefinisikan kembali menjadi:

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{RJRR}} = (\boldsymbol{I} - k_{GM}\boldsymbol{I}(\boldsymbol{Z}'\boldsymbol{Z} + k_{GM}\boldsymbol{I})^{-1})^{2}\hat{\boldsymbol{\alpha}}_{\boldsymbol{GM}} \tag{35}$$

dengan k_{GM} adalah parameter regresi *robust ridge* maupun regresi *robust jackknife ridge* (Hanifah, 2020).
dan estimator *jackknife ridge regression* untuk $\hat{\boldsymbol{\beta}}$ adalah:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{RJRR}} = \boldsymbol{P}\,\hat{\boldsymbol{\alpha}}_{\boldsymbol{RJRR}} \tag{36}$$

1.5.10 Parameter *Ridge* K

Menentukan nilai k untuk mendapatkan koefisien regresi yang stabil. Metode yang digunakan yaitu memilih konstanta bias (k), metode ini ditemukan oleh Hoerl, Kenard dan Balwin (1975). Untuk menghitung nilai k dengan menggunakan rumus dengan persamaan:

$$k = \frac{p\hat{\sigma}^{2}}{\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}}} \tag{37}$$

Dimana p adalah banyaknya variabel prediktor dan $\hat{\sigma}$ merupakan *mean square error* hasil estimasi MKT. Namun, untuk regresi *robust jackknife ridge* menggunakan estimator GM. Mengganti nilai k menjadi nilai k_{GM} untuk menentukan ... mana nilai $\hat{\sigma}$ dan $\hat{\beta}$ diganti dengan estimasi *robust.* Sehingga ... persamaan berikut:

$$k_{GM} = \frac{p\hat{\sigma}_{GM}{}^{2}}{\hat{\boldsymbol{\alpha}}_{\boldsymbol{GM}}{}'\hat{\boldsymbol{\alpha}}_{\boldsymbol{GM}}} \tag{38}$$

$$\hat{\sigma}_{GM}{}^2 = \frac{(\boldsymbol{y} - \boldsymbol{Z}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{GM}})'(\boldsymbol{y} - \boldsymbol{Z}\widehat{\boldsymbol{\alpha}}_{\boldsymbol{GM}})}{n - p}$$

Nilai k akan digunakan untuk perhitungan pada *ridge regression* dan *jackknife ridge regression*, sedangkan k_{GM} digunakan pada regresi *ridge robust* dan *jackknife ridge robust* (Hanifah, 2020).

1.5.11 Validasi Model

Root Mean Square Error Prediction (RMSEP) dapat digunakan untuk mengetahui kriteria kebaikan model dalam validasi. RMSEP adalah nilai varians dari residual yang dapat menunjukkan keakuratan suatu model. RMSEP mempunyai nilai minimal 0, semakin kecil nilai RMSEP menunjukkan bahwa perbedaan antara nilai dugaan hasil pemodelan dengan data aktualnya semakin kecil pula. Sehingga model yang terbaik adalah model dengan nilai RMSEP terkecil (Law & Kelton, 1991) dengan menggunakan Persamaan (31).

$$RMSEP = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{39}$$

1.5.12 Curah Hujan

Menurut Badan Meteorologi Klimatologi dan Geofisika, curah hujan adalah jumlah ketinggian air hujan yang terkumpul di tempat datar yang tidak meresap atau tidak mengalir di permukaan tanah (BMKG). Curah hujan adalah volume air yang jatuh di atas permukaan bidang datar dalam jangka waktu tertentu dalam ketinggian milimiter (Laia dan Setyawan, 2022).

Di Indonesia khususnya daerah Sulawesi Selatan curah hujan tergolong ekstrim karena dampak iklim tropis basah dengan rata-rata curah hujan 3.200 mm per tahun (BPS, 2024). Kondisi ini dipengaruhi oleh cuaca yang tidak stabil dan lokasi yaitu lautan dan sungai yang luas (Hartono dan Sutikno, 2020). Intensitas curah hujan adalah jumlah air hujan yang terkumpul di permukaan datar dalam waktu tertentu dalam ketinggian milimeter (mm). Curah hujan menurut BMKG dibagi menjadi empat kelompok, yaitu:

1. Curah hujan rendah: 0-20 mm, 21-50 mm, 51-100 mm.
2. Curah hujan menengah: 101-150 mm, 151-200 mm, 201-300 mm.
3. Curah hujan tinggi: 301- 500 mm.
4. Curah hujan sangat tinggi: > 500 mm.

Pada tahun 2023, tingkat curah hujan di daerah Sulawesi Selatan bagian barat, termasuk Kabupaten Pangkep mencapai tingkat yang cukup tinggi. Kabupaten Pangkep, yang terletak di bagian barat Sulawesi Selatan, memiliki musim hujan yang ... an musim kemarau yang singkat, dikarenakan oleh keadaan ... an perputaran arus udara yang mengakibatkan ketidakstabilan ...

... adalah salah satu komponen yang dapat mempengaruhi ... bupaten Pangkep. Laju penguapan air di lokasi produksi garam ... ep akan dipengaruhi oleh durasi musim kemarau yang relatif singkat (Sahriman dan Yulianti, 2023). Jumlah hujan yang relatif tinggi akan mempengaruhi jumlah produksi garam di wilayah tersebut serta dapat mempengaruhi ekonomi penduduk.