

BAB I PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara kepulauan yang memiliki garis pantai terpanjang kedua di dunia setelah Kanada (Hanmina dkk., 2024). Garis pantai yang membentang sepanjang 108.000 km pada sekitar 17.504 pulau memberikan Indonesia potensi besar dalam pengembangan berbagai sumber daya kelautan, termasuk budidaya garam pesisir (Hidayat dkk., 2023). Kementerian Kelautan dan Perikanan (KKP) pada tahun 2024 mencatatkan keberhasilan produksi garam sebesar 147 persen dari target capaian yang ditetapkan tahun 2023 sebesar 1,7 juta ton. Salah satu wilayah yang memberikan kontribusi signifikan terhadap produksi garam nasional adalah Kabupaten Pangkep (Thahir, 2024). Kabupaten Pangkep merupakan salah satu wilayah di Sulawesi Selatan yang selama ini memproduksi garam. Kabupaten Pangkep memiliki rata-rata 80 orang petani garam dalam satu desa dan terdapat tiga kecamatan dengan luas lahan tambak garam mencapai 853 ha (Mukhlis, 2024).

Potensi besar yang dimiliki Kabupaten Pangkep sebagai daerah penghasil garam, juga berbanding lurus dengan potensi tantangan yang dihadapi sebagai daerah beriklim tropis. Proses produksi garam yang umumnya masih menggunakan cara tradisional bergantung pada kondisi curah hujan daerah setempat. Fenomena perubahan iklim dan cuaca ekstrem seperti hujan di luar musim dapat mengganggu siklus produksi garam, mengakibatkan penurunan hasil panen dan pendapatan petani (Akbar dkk., 2023). Hal ini menunjukkan pentingnya pengembangan metode peramalan curah hujan untuk meningkatkan ketahanan produksi garam. Metode peramalan curah hujan yang efektif akan memberikan informasi bagi para petambak garam terkait periode optimal untuk melakukan produksi garam terkhusus di Kabupaten Pangkep. Akan tetapi, kondisi curah hujan yang dipengaruhi oleh berbagai faktor membuat ketersediaan data curah hujan memiliki resolusi yang rendah dan membuat proses peramalan menjadi tidak akurat.

Salah satu model data yang dapat dimanfaatkan untuk peramalan curah hujan adalah *Global Circulation Model* (GCM). Kajian klimatologi jangka panjang menjelaskan bahwa GCM merupakan suatu model global prediksi utama iklim secara numerik dan sebagai sumber informasi primer untuk menilai pengaruh perubahan iklim (Estiningtyas & Wigena, 2011). Meskipun GCM menyediakan data primer, penerapannya pada skala lokal sering kali menghadapi keterbatasan akurasi. Luaran GCM tidak dapat mempertimbangkan wilayah berskala lokal dan memiliki resolusi yang lebih tinggi dari luaran GCM (Pitri dkk., 2022). Penanganan masalah tersebut dapat diatasi dengan implementasi model *Statistical Downscaling* (SD) untuk mengkonversi data GCM yang berskala global menjadi informasi yang relevan dengan kondisi spesifik di suatu daerah



SD efektif dalam menghasilkan data iklim skala lokal yang lebih akurat dari skala global sehingga memungkinkan analisis yang lebih relevan untuk kebutuhan spesifik daerah (Sahrman dkk., 2014). proses analisis model SD sering kali dihadapkan dengan kondisi memiliki dimensi besar dan korelasi yang tinggi antara hubungan data spasialnya (Soleh dkk., 2015). Kondisi tersebut dapat

menyebabkan adanya multikolinearitas pada data yang mengakibatkan besarnya variansi pendugaan (Kyriazos & Poga, 2023).

Ridge Regression (RR) merupakan salah satu metode untuk mengatasi multikolinearitas dengan menambahkan parameter *shrinkage* pada matriks diagonal yang membuat modelnya lebih stabil dan mengurangi keragaman koefisien regresi (Hoerl & Kennard, 1970). Akan tetapi, RR menggunakan parameter *shrinkage* yang melakukan penyusutan secara seragam, baik pada variabel dengan informasi lemah, maupun variabel dengan informasi yang kuat. Hal tersebut membuat nilai bias yang dihasilkan RR tidak selalu kecil (Andini dkk., 2025). Selanjutnya, Liu (1993) memperkenalkan *Liu Regression* (LR) yang menggunakan konstanta Liu (d) untuk mereduksi multikolinearitas pada data. Efisiensi penaksir LR bergantung pada nilai d yang didapatkan. Penggunaan nilai d yang kecil dapat mengakibatkan *overfitting*, sedangkan nilai d yang besar menyebabkan *trade-off* bias pada variansi estimator. Kedua metode tersebut baik RR, maupun LR memiliki kelemahan berupa parameter penyusutan yang cenderung tidak adaptif untuk mengatasi multikolinearitas pada data.

Kibria & Lukman (2020) merumuskan *Kibria-Lukman Regression* (KLR) sebagai metode alternatif untuk mengatasi masalah tersebut. Metode KLR yang menggabungkan keunggulan RR dalam meminimalkan variansi dengan konsep pengoreksian bias pada LR untuk mengatasi multikolinearitas. Metode KLR bekerja dengan cara menambahkan elemen bias adaptif pada model yang menyesuaikan tingkat penyusutan berdasarkan informasi masing-masing variabel. Sehingga memberikan *trade-off* antara bias dan variansi yang lebih optimal.

Penelitian terdahulu yang dilakukan oleh Kibria & Lukman (2020) menggunakan *Kibria-Lukman Regression* untuk mengatasi multikolinearitas dalam pemodelan data kandungan yang terdapat pada semen *Portland* dan data ekonomi Prancis. Hasil penelitian tersebut menunjukkan bahwa model KLR merupakan model terbaik pada kedua dataset tersebut dengan memberikan hasil *mean square error* terendah dibandingkan dengan metode lain. Penelitian lain dilakukan oleh Dawoud dkk. (2022) dengan melakukan perbandingan performa prediksi KLR dengan metode pereduksi multikolinearitas lain. Hasil penelitian tersebut menunjukkan bahwa model KLR mampu memberikan hasil prediksi yang lebih baik. Lukman dkk. (2023) juga melakukan penelitian terkait KLR dengan metode regresi logistik menggunakan data kategori gejala kanker yang memberikan hasil performa terbaik dibandingkan metode estimasi lain

Berdasarkan penelitian terdahulu, KLR mampu memberikan hasil yang baik dalam mengatasi multikolinearitas dan memberikan prediksi yang baik. Oleh karena itu, penelitian ini akan menggunakan model *Statistical Downscaling* dengan metode *Kibria-Lukman Regression* untuk peramalan curah hujan Kabupaten Pangkep. Penelitian ini diharapkan dapat memberikan hasil yang lebih optimal untuk mengatasi multikolinearitas

asil peramalan akurat pada curah hujan Kabupaten Pangkep.



lah

alahan yang telah diuraikan sebelumnya, batasan masalah pada sebagai berikut.

unakan pada penelitian ini adalah data curah hujan di Kabupaten 1999-2023.

2. Penambahan variabel *dummy* menggunakan teknik *k-means clustering*.

1.3 Tujuan dan Manfaat Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Memperoleh model *Statistical Downscaling* dengan metode *Kibria-Lukman Regression* untuk data curah hujan di Kabupaten Pangkep.
2. Memperoleh hasil peramalan curah hujan di Kabupaten Pangkep tahun 2023 menggunakan model *Statistical Downscaling* dengan metode *Kibria-Lukman Regression*.

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai referensi untuk menambah wawasan pembaca tentang cara menganalisis model *Statistical Downscaling* dengan *Kibria-Lukman Regression* sebagai salah satu metode untuk mengatasi multikolinearitas dalam melakukan peramalan curah hujan. Selain itu, dapat bermanfaat sebagai sumber informasi terkait curah hujan Kabupaten Pangkep.

1.4 Landasan Teori

1.4.1 Global Circulation Model dan Statistical Downscaling

Global Circulation Model (GCM) adalah bentuk dan pola sistem iklim, serta interaksi antar atmosfer, lautan dan komponen-komponen terkait lainnya yang direpresentasikan dalam model numerik (Wigena dkk., 2015). Data yang dihasilkan GCM berbentuk grid atau petak wilayah dengan resolusi $2,5^\circ$ atau $\pm 300 \text{ km}^2$ yang memberikan informasi iklim secara global (Notodiputro dkk., 2012). Model GCM merupakan salah satu alat terpenting dalam mempelajari sistem iklim, khususnya curah hujan (Upa dkk., 2021). Akan tetapi, GCM hanya memberikan informasi variabel iklim global yang memiliki resolusi rendah. Oleh karena itu, dibutuhkan *Statistical Downscaling* yang dapat menghubungkan variabel iklim dalam skala global dengan skala lokal yang memiliki resolusi tinggi untuk melakukan prediksi dengan tingkat akurasi tinggi (Osses dkk., 2020).

Statistical Downscaling (SD) merupakan pendekatan statistik yang didasarkan pada hubungan antara skala global (prediktor) dan skala lokal (respon) untuk menerjemahkan anomali global menjadi anomali dalam variabel iklim lokal (Sahrman dkk., 2024). Model SD akan menghasilkan akurasi yang baik dengan syarat, antara lain (Wigena dkk., 2015):

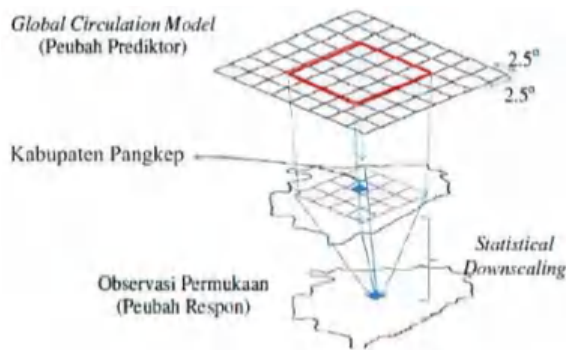
1. Korelasi yang tinggi antara variabel prediktor dan variabel respon.
2. Variabel prediktor disimulasikan dengan baik oleh GCM.
3. Hubungan antara variabel respon dan variabel prediktor tidak berubah seiring dengan perkembangan waktu dan tetap sama meskipun terjadi perubahan iklim di masa depan.

Bentuk umum model SD dapat disajikan pada Persamaan (1) berikut:



$$y = f(x) \quad (1)$$

akan variabel iklim lokal (misalnya: curah hujan), f dapat berbentuk apapun non-linier, $x_{(n \times p)}$ adalah variabel-variabel luaran GCM in), n adalah banyaknya observasi dalam satuan waktu (misalnya: dan p adalah banyaknya grid domain GCM.



Gambar 1. Proses *Statistical Downscaling*

Gambar 1 menunjukkan proses *Statistical Downscaling* GCM di suatu daerah. Jumlah grid GCM yang berada pada suatu daerah disesuaikan dengan luas daerahnya. Keseluruhan grid yang mengandung informasi iklim pada suatu daerah kemudian dihubungkan dengan kondisi iklim lokal untuk mendapatkan hasil yang optimal dalam melakukan pemodelan dan prediksi (Sutikno, 2008).

1.4.2 Analisis Regresi Linier Berganda

Regresi linier berganda adalah salah analisis statistika yang digunakan untuk menjelaskan hubungan model beberapa variabel prediktor (x) terhadap variabel respon (y) (Prahutama & Rahmawati, 2024). Variabel prediktor merupakan variabel yang mempengaruhi variabel respon secara linier, sehingga variabel respon dapat diduga dari variabel prediktornya. Bentuk umum persamaan regresi linier berganda dapat ditulis pada Persamaan (2).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j x_{ij} + \dots + \beta_p x_{ip} + e_i \quad (2)$$

dengan,

y_i : variabel respon pengamatan ke- i ($i = 1, 2, \dots, n$)

β_0 : konstanta parameter regresi

$\beta_1, \beta_2, \dots, \beta_p$: parameter regresi ke- j ($j = 1, 2, \dots, p$)

x_{ij} : variabel prediktor ke- j untuk observasi ke- i

e_i : variabel galat ke- i

dalam notasi matriks, model pada Persamaan (2) yang dapat ditulis pada Persamaan (3).

$$Y = X\beta + e \quad (3)$$

dimana,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$



variabel respon yang berukuran ($n \times 1$)

variabel prediktor yang berukuran ($n \times (p + 1)$)

parameter yang berukuran ($(p + 1) \times 1$)

galat yang berukuran ($n \times 1$)

1.4.3 Metode Ordinary Least Square

Ordinary Least Square (OLS) merupakan metode regresi yang meminimumkan jumlah galat (*error*) kuadrat (Makarti & Karim, 2017). Metode estimasi parameter yang digunakan adalah menduga koefisien regresi (β) dengan meminimumkan jumlah kuadrat galat. Adapun penaksir parameternya pada Persamaan (4)

$$\hat{\beta}_{OLS}^* = (X'X)^{-1}X'Y \quad (4)$$

dengan $\hat{\beta}_{OLS}^*$ merupakan parameter regresi OLS tanpa intersep dan X merupakan matriks pada variabel prediktor dengan ukuran $n \times p$. Persamaan (3) juga dapat dituliskan dalam bentuk ortogonal dengan melakukan dekomposisi spektral pada matriks $X'X$ yang merupakan matriks simetris, kemudian dapat digunakan untuk membentuk matriks ortogonal P . Dekomposisi spektral matriks $X'X$ dituliskan pada Persamaan (5)

$$\begin{aligned} X'X &= P'\Lambda P \\ P(X'X)P' &= P(P'\Lambda P)P' \\ P(X'X)P' &= (P'P)\Lambda(P'P) \\ P(X'X)P' &= \Lambda \end{aligned} \quad (5)$$

dengan Λ merupakan matriks diagonal berukuran $p \times p$ yang elemennya adalah nilai eigen dari matriks $X'X$ dan P adalah matriks ortogonal berukuran $p \times p$ yang elemennya adalah vektor *eigen* dari matriks $X'X$. Karena Λ merupakan matriks diagonal dan simetris, maka $\Lambda' = \Lambda$. Sehingga Persamaan (5) dapat ditulis menjadi Persamaan (6) sebagai berikut.

$$\begin{aligned} (P(X'X)P')' &= \Lambda' \\ P'(X'X)P &= \Lambda \\ P'X'XP &= \Lambda \\ (XP)'XP &= \Lambda \\ Z'Z &= \Lambda \end{aligned} \quad (6)$$

Kemudian, Persamaan (6) dapat digunakan untuk mengubah Persamaan (3) menjadi bentuk ortogonal pada Persamaan (7)

$$\begin{aligned} Y &= X\beta^* + \varepsilon \\ &= (XP)(P'\beta^*) + \varepsilon \\ &= Z\alpha + \varepsilon \end{aligned} \quad (7)$$

dengan β^* merupakan parameter regresi tanpa intersep, $Z = XP$, $\alpha = P'\beta^*$, dan $\beta^* = P\alpha$. Selanjutnya, dengan menggunakan persamaan (4), Persamaan (8) estimator $\hat{\alpha}$ ditulis sebagai berikut.

$$\begin{aligned} \hat{\alpha}_{OLS} &= (Z'Z)^{-1}Z'Y \\ &= \Lambda^{-1}Z'Y \end{aligned} \quad (8)$$

Persamaan (8) kemudian dapat dibentuk kembali menjadi koefisien β menggunakan

$$\begin{aligned} \hat{\alpha}_{OLS} &= [(XP)'(XP)^{-1}](XP)'Y \\ &= (P'X'XP)^{-1}P'X'Y \\ &= (P'X'XP)^{-1}P'X'X\hat{\beta}_{OLS}^* \\ &= P'(X'X)^{-1}(X'X)\hat{\beta}_{OLS}^* \\ &= P'\hat{\beta}_{OLS}^* \end{aligned} \quad (9)$$



atau dapat ditulis menjadi Persamaan (10).

$$\hat{\beta}_{OLS}^* = P\hat{\alpha}_{OLS} \quad (10)$$

Dengan menggunakan Persamaan (8), nilai harapan dari $\hat{\alpha}_{OLS}$ dapat ditulis pada Persamaan (11).

$$\begin{aligned} E(\hat{\alpha}_{OLS}) &= E(\Lambda^{-1}Z'Y) \\ &= E(\Lambda^{-1}Z'(Z\alpha + \varepsilon)) \\ &= E(\Lambda^{-1}Z'Z\alpha + \Lambda^{-1}Z'\varepsilon) \\ &= E(\Lambda^{-1}\Lambda\alpha + \Lambda^{-1}Z'\varepsilon) \\ &= \Lambda^{-1}\Lambda E(\alpha) + \Lambda^{-1}Z'E(\varepsilon) \\ &= I\alpha + 0 \\ &= \alpha \end{aligned} \quad (11)$$

Selanjutnya, Persamaan (7) dapat digunakan untuk membentuk Persamaan (12).

$$\begin{aligned} Z'Y &= Z'Z\alpha + Z'\varepsilon \\ &= \Lambda\alpha + Z'\varepsilon \end{aligned} \quad (12)$$

kemudian, Persamaan (12) dapat disubstitusi pada Persamaan (7) untuk mendapatkan Persamaan (13).

$$\hat{\alpha}_{OLS} = \Lambda^{-1}(\Lambda\alpha + Z'\varepsilon) = \alpha + \Lambda^{-1}Z'\varepsilon \quad (13)$$

dengan menggunakan aturan variansi linier $Var(A\varepsilon) = A Var(\varepsilon)A'$ dan $Var(\varepsilon) = \sigma^2 I$, $Var(\hat{\alpha}_{OLS})$ dituliskan pada Persamaan (14).

$$\begin{aligned} Var(\hat{\alpha}_{OLS}) &= Var(\Lambda^{-1}Z'\varepsilon) \\ &= \Lambda^{-1}Z'(\sigma^2 I)(\Lambda^{-1}Z')' \\ &= \Lambda^{-1}Z'(\sigma^2 I)Z\Lambda^{-1} \\ &= \sigma^2 \Lambda^{-1}Z'Z\Lambda^{-1} \\ &= \sigma^2 \Lambda^{-1}\Lambda\Lambda^{-1} \\ &= \sigma^2 \Lambda^{-1} \end{aligned} \quad (14)$$

1.4.4 Multikolinearitas

Multikolinearitas merupakan masalah yang sering muncul dalam analisis regresi linier berganda dengan keadaan adanya keterkaitan yang kuat antara variable prediktor (Sriningsih dkk., 2018). Menurut Arisandi dkk. (2021) data yang terdampak oleh multikolinearitas dapat mengakibatkan penaksiran yang berbias, tidak stabil, dan mungkin jauh dari nilai sasaran. Salah satu cara untuk mendeteksi multikolinearitas adalah dengan melakukan perhitungan nilai *Variance Inflation Factor* (VIF).

Persamaan untuk menghitung nilai VIF dapat dilihat pada Persamaan (15) berikut:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (15)$$



$j, \dots, p$ dan p merupakan banyaknya variabel prediktor, serta R_j^2 rminasi dari variabel prediktor yang diregresikan terhadap variabel inya. Variabel akan terdeteksi mengalami multikolinearitas apabila dkk., 2021).

1.4.5 Metode Kibria-Lukman Regression

Multikolinearitas pada data dapat diatasi dengan menambahkan konstanta *shrinkage* (penyusutan) pada proses pemodelannya. Penggunaan konstanta *shrinkage* dilakukan oleh (Hoerl & Kennard, 1970) dengan metode *Ridge Regression* (RR) yang menggunakan konstanta bias non-negatif k . Metode RR dalam mengatasi multikolinearitas memiliki bentuk persamaan umum ortogonal yang ditulis pada Persamaan (16).

$$\hat{\alpha}_{RR} = (\Lambda + kI)^{-1}Z'Y \quad (16)$$

Konstanta bias k melakukan penyusutan terhadap variansi seluruh variabel dalam model. Akan tetapi, penyusutan seragamnya dapat menimbulkan bias yang signifikan, terutama pada koefisien variabel dengan informasi kuat. Selanjutnya, Liu (1993) merumuskan *Liu Regression* (LR) dengan konstanta *shrinkage* d dengan efek yang lebih halus untuk mengatasi bias pada pembentukan model. Metode LR memiliki bentuk umum persamaan yang ditulis pada Persamaan (17).

$$\hat{\alpha}_{LR} = (\Lambda + I)^{-1}(\Lambda + dI)\hat{\alpha}_{OLS} \quad (17)$$

Konstanta d dalam mengatasi multikolinearitas pada data bersifat sangat sensitif. Pemilihan yang bersifat subjektif dan tidak tepat dapat menghasilkan pendugaan dengan *trade-off* yang tinggi antara bias dan variansi. Kibria & Lukman (2020) kemudian mengusulkan sebuah parameter regresi yang disebut dengan *Kibria-Lukman Regression* (KLR) pada model linear yang bertujuan untuk memperkenalkan parameter baru k yang membantu mengatasi multikolinearitas lebih efektif dengan menyeimbangkan bias dan variansi. Metode KLR bekerja dengan cara menambahkan elemen bias ke dalam perhitungan untuk mengurangi variansi dan multikolinearitas. Metode KLR merupakan modifikasi dari fungsi objektif LR dan konstanta bias k pada RR yang memiliki karakteristik pereduksi multikolinearitas pada parameter biasnya. Kombinasi karakteristik kedua metode di atas dapat menjadi alternatif dalam mengatasi multikolinearitas dengan Persamaan (18)

$$\hat{\alpha}_{KL} = (\Lambda + k_{KL}I)^{-1}(\Lambda - k_{KL}I)\hat{\alpha}_{OLS} \quad (18)$$

dengan k_{KL} merupakan nilai konstanta *Kibria-Lukman* untuk KLR. Matriks $(\Lambda + k_{KL}I)^{-1}$ pada Persamaan (18) berfungsi seperti pada RR untuk melakukan penyusutan terhadap variansi pada model. Sedangkan, matriks $(\Lambda - k_{KL}I)\hat{\alpha}_{OLS}$ berperan seperti LR dengan mengembalikan nilai eigen yang disusutkan secara berlebihan untuk mengoptimalkan bias pada model. Bentuk koefisien regresi $\hat{\beta}_{KL}$ dapat diperoleh menggunakan Persamaan (10) menjadi Persamaan (19).

$$\hat{\beta}_{KL} = P\hat{\alpha}_{KL} \quad (19)$$

Nilai k_{KL} dapat diperoleh dengan mencari nilai optimal menggunakan Persamaan (20).

$$k_{j\ KL} = \frac{\hat{\sigma}^2}{2\hat{\alpha}_{j\ OLS}^2 + \left(\frac{\hat{\sigma}^2}{\lambda_j}\right)} \quad (20)$$



ii optimal k_{KL} dapat menggunakan beberapa persamaan sebagai satu persamaan yang dapat digunakan adalah dengan maan yang dirumuskan oleh Özkale & Kaçiranlar (2007) dan maan (21).

$$k_{KLmin} = \min \left(\frac{\hat{\sigma}^2}{2\hat{\alpha}_{jOLS}^2 + \left(\frac{\hat{\sigma}^2}{\lambda_j} \right)} \right) \quad (21)$$

Selain itu, Hoerl dkk. (1975) juga merumuskan persamaan untuk menentukan nilai k_{KL} optimal yang dituliskan pada Persamaan (22)

$$k_{KLHMN} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^p \left(2\hat{\alpha}_{jOLS}^2 + \left(\frac{\hat{\sigma}^2}{\lambda_j} \right) \right)} \quad (22)$$

dengan $\hat{\sigma}^2$ merupakan variansi yang dapat dihitung menggunakan Persamaan (23)

$$\hat{\sigma}^2 = \left(\frac{(Y - Z\hat{\alpha}_{OLS})'(Y - Z\hat{\alpha}_{OLS})}{(n - p)} \right) \quad (23)$$

1.4.6 Validasi Model

Root Mean Square Error Prediction (RMSEP) adalah salah satu teknik untuk melakukan validasi dalam kebaikan model. RMSEP adalah nilai variansi dari residual yang dapat menunjukkan keakuratan suatu model. Semakin kecil nilai RMSEP, maka semakin akurat model yang digunakan (Hendarwati & Lepong, 2023). Sehingga model terbaik adalah model yang memiliki RMSEP terkecil. Berikut persamaan untuk menghitung nilai RMSEP pada Persamaan (24).

$$RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (24)$$

dengan

Y_i = nilai variabel respon pada data validasi ke- i , $i = 1, 2, \dots, n$

$\hat{Y}_i$ = nilai dugaan pada data validasi ke- i , $i = 1, 2, \dots, n$

n = banyaknya obsevasi data

1.4.7 Curah Hujan

Curah Hujan adalah akumulasi air hujan yang jatuh pada periode waktu tertentu yang pengukurannya menggunakan satuan tinggi di atas permukaan tanah horizontal yang diasumsikan tidak terjadi infiltrasi, *run off*, maupun evaporasi (Ruswanti, 2020). Curah hujan merupakan salah satu unsur cuaca yang datanya diperoleh dengan cara mengukurnya dengan menggunakan alat penakar hujan, sehingga dapat diketahui jumlahnya dalam satuan mm (Chandra & Suprpto, 2016). Menurut Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), curah hujan dibagi menjadi empat kelompok, yaitu:

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



di Indonesia khususnya Sulawesi Selatan memiliki curah hujan yang basah dengan rata-rata curah hujan 3.200 mm per tahun. Pola Provinsi Sulawesi Selatan terdapat tiga tipe pola curah hujan yaitu monsun, tipe equatorial dan tipe lokal. Pada Sulawesi Selatan kabupaten Pangkep memiliki intensitas hujan yang tinggi (> 300 mm/bulan)

pada bulan November hingga Maret, dan periode curah hujan rendah (<50 mm/bulan) pada bulan Mei hingga Oktober (Wicaksono, 2022). Curah hujan yang tinggi dengan periode yang relatif lama akan berpengaruh terhadap produksi garam Kabupaten Pangkep yang sebagian besar masih menggunakan sistem *solar evaporation* (Sahriman dkk., 2024).



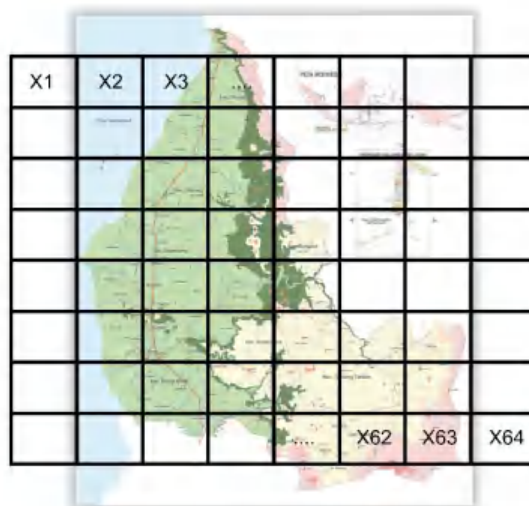
BAB II METODE PENELITIAN

2.1 Sumber Data

Penelitian ini menggunakan data presipitasi/curah hujan luaran GCM *Climate Model Intercomparison Project* (CMIP5) dalam satuan mm/bulan yang diperoleh dari situs web KNMI belanda di <https://www.climatexp.knmi.nl/>. Domain grid data GCM yang digunakan terdiri atas 8×8 grid pada $14.75 \text{ LS} - 5.25 \text{ LU}$ dan $109.54 \text{ BT} - 129.54 \text{ BT}$ sehingga terdapat 64 variabel. Data curah hujan yang digunakan dalam penelitian ini merupakan rata-rata curah hujan bulanan dari 3 pos hujan di Kabupaten Pangkajene Kepulauan (Pangkep) yaitu Bungoro, Ma'rang dan Labakkang yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) wilayah IV Makassar.

2.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah variabel respon (Y) yang menyatakan data intensitas curah hujan (mm/bulan) di Kabupaten Pangkep tahun 1999-2023. Variabel prediktor (X) terdiri dari 67 variabel yang menyatakan data presipitasi luaran GCM sebanyak 64 variabel dan variabel *dummy* sebanyak 3 variabel. Adapun Ilustrasi luaran GCM pada Gambar 2.



Gambar 2. Ilustrasi Variabel Prediktor pada Grid Luaran GCM



Optimized using
trial version
www.balesio.com

Gambar 2 menunjukkan ilustrasi grid pada Kabupaten Pangkep sebesar 8×8 (64) nya berisi beberapa variabel iklim yang salah satunya adalah curah hujan mm/bulan sehingga variabel prediktor hasil luaran GCM terdapat 64 variabel. Selain luaran GCM, variabel *dummy* yang digunakan sebagai variabel hasil pengelompokkan variabel respon dengan *k-means clustering* yang terdiri dari 4 *cluster* yang optimal sehingga terdapat 67 variabel *dummy* sebagai batas (pemisah) tiap *cluster* yang terbentuk. Variabel *dummy* tersebut dapat dilihat pada Tabel 1 berikut dan tabulasi data dituliskan pada

Lampiran 1. Adapun nilai variabel *dummy* dan keterangannya dapat dilihat pada Lampiran 2.

Tabel 1. Struktur data

Observasi	Y	X_1	X_2	...	X_{64}	D_1	D_2	D_3
1	Y_1	$x_{1(1)}$	$x_{1(2)}$	...	$x_{1(64)}$	$D_{1(1)}$	$D_{1(2)}$	$D_{1(3)}$
2	Y_2	$x_{2(1)}$	$x_{2(2)}$	...	$x_{2(64)}$	$D_{2(1)}$	$D_{2(2)}$	$D_{2(3)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
298	Y_{298}	$x_{298(1)}$	$x_{298(2)}$	...	$x_{298(64)}$	$D_{298(1)}$	$D_{298(2)}$	$D_{298(3)}$
299	Y_{299}	$x_{299(1)}$	$x_{299(2)}$	...	$x_{299(64)}$	$D_{299(1)}$	$D_{299(2)}$	$D_{299(3)}$
300	Y_{300}	$x_{300(1)}$	$x_{300(2)}$	...	$x_{300(64)}$	$D_{300(1)}$	$D_{300(2)}$	$D_{300(3)}$

2.3 Metode Analisis

Berikut adalah langkah-langkah yang digunakan dalam penelitian ini.

1. Melakukan analisis deskriptif data curah hujan sebagai informasi awal keragaman data amatan.
2. Membagi data menjadi dua, yaitu data *training* untuk pemodelan tahun 1999-2022 dan data *test* untuk validasi tahun 2023.
3. Mengidentifikasi multikolinearitas pada data presipitasi menggunakan VIF dengan Persamaan (15).
4. Melakukan pemodelan *Statistical Downscaling* pada data curah hujan skala global dan curah hujan di Kabupaten Pangkep menggunakan metode *Kibria-Lukman Regression*. Adapun langkah – langkahnya adalah sebagai berikut:

- a. Menghitung nilai eigen dan vektor eigen dari matriks $X'X$, serta membentuk matriks P yaitu matriks vektor eigen dan membentuk matriks $Z = XP$. Sehingga model regresi menjadi $y = Z\alpha + \varepsilon$.

- b. Menghitung nilai estimasi $\hat{\alpha}_{OLS}$ dengan metode kuadrat terkecil

$$\hat{\alpha}_{OLS} = (Z'Z)^{-1}Z'y$$

- c. Menghitung nilai konstanta untuk *Kibria-Lukman Regression* menggunakan Persamaan (21) dan Persamaan (22).

- d. Menghitung estimasi koefisien parameter dengan menggunakan *Kibria-Lukman Regression* ($\hat{\alpha}_{KL}$) dengan menggunakan Persamaan (18).

$$\hat{\alpha}_{KL} = (\Lambda + kI)^{-1}(\Lambda - kI) \hat{\alpha}_{OLS}$$

- e. Mengubah model persamaan regresi menjadi

$$\hat{\beta}_{KL} = P\hat{\alpha}_{KL}$$

- f. Menghitung nilai intersep persamaan regresi menggunakan persamaan

$$\beta_{KL0} = \bar{Y} - \beta_{KL1}\bar{X}_1 - \beta_{KL2}\bar{X}_2 - \dots - \beta_{KL D_3}\bar{X}_{D_3}$$

- g. Membentuk model persamaan regresi.

- h. Pemilihan model terbaik dan validasi model menggunakan t).

- i. ah hujan Kabupaten Pangkep menggunakan data *test* tahun 2023.

