JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 22, No. 3, May 2026, pp. 627-654
DOI: 10.20956/j.v22i3.49179

Statistical Downscaling Using the Kibria-Lukman Regression with Dummy Variables for Rainfall Forecasting

Sitti Sahriman[1*], Syamsuddin Toaha[2], Anisa Kalondeng[3], M. Zaky Hisyam Gozhi[4], Aidul Fitri Mustamin[5]

[1,3,4,5]*Departemen of Statistics, Hasanuddin University*
[2]*Departemen of Mathematics, Hasanuddin University*
Email: [1]*sittisahrimansalam@gmail.com*, [2]*syamsuddint@gmail.com*, [3]*nkalondeng@gmail.com*, [4]*zakyhg25@gmail.com*, [5]*aidulfitrim@gmail.com*
**Corresponding author*
Received: 15 December 2026, revised: 8 April 2026, accepted: 13 April 2026

Abstract

Global Circulation Models (GCM) are widely used to project climate variables at the global scale. However, the relatively coarse spatial resolution of GCM outputs makes direct GCM-based climate forecasting generally less accurate at the local scale. To bridge this scale mismatch, this study applies statistical downscaling (SD). The main challenge in GCM-based SD is the large number of predictor (grid) variables that are strongly correlated, which induces multicollinearity and can reduce the stability of coefficient estimation. To address this issue, Kibria-Lukman Regression (KLR) is used, which is a shrinkage method that combines the variance-reduction property of Ridge Regression (RR) with the bias-control concept of Liu Regression (LR). This study aims to obtain the best SD model and to produce local rainfall forecasting in Pangkep Regency during the 2023–2024 testing period. The research stages included: (1) developing an SD model using KLR with the addition of dummy variables as additional predictors; and (2) evaluating the forecasts using testing data. The results showed that the KLR model with dummy variables provided the best performance, with an RMSE of 73.002 and a coefficient of determination (R^2) of 94.069%. At the validation stage, the model also produced a high pattern agreement (correlation of 0.936) and a relatively low forecasting error (RMSEP of 94.614), and it outperformed other multicollinearity-handling approaches. Thus, the proposed model has the potential to serve as a tool for local-scale rainfall forecasting to support salt production planning in Pangkep Regency.

Keywords: dummy variable, global circulation model, kibria-lukman regression, multicollinearity, rainfall, statistical downscaling

1. INTRODUCTION

Indonesia is an archipelagic country with the second-longest coastline in the world after Canada (Hanmina et al., 2024). With a coastline of approximately 108,000 km and around 17,504 islands, Indonesia has great potential for the development of marine resources, including coastal salt farming (Hidayat et al., 2023). In line with this potential, the Ministry of Marine Affairs and Fisheries (KKP) in 2024 reported salt production achievement of 147% of the 2023 target of 1.7 million tons. One of the regions that contributed significantly to national salt production is Pangkep Regency (Thahir, 2024). Pangkep Regency in South Sulawesi is known as a salt-producing area with production characteristics dominated by small-scale farmers. On average, there are about 80

salt farmers in one village, and production activities are concentrated in three sub-districts with a total salt pond area of 853 ha (Mukhlis, 2024).

On the other hand, salt production in tropical regions is strongly influenced by weather conditions, especially rainfall. Production that is still largely carried out using traditional methods depends on sufficiently long dry periods, so off-season rainfall can disrupt the production process, reduce harvest yields, and affect farmers' income (Akbar et al., 2023). Therefore, more reliable rainfall forecasting information is needed to help salt farmers determine more appropriate production timing. However, rainfall forecasting at the local scale is not straightforward because rainfall is influenced by many factors, while the available data are often not detailed enough to capture local variability well.

One widely used source of climate data is the Global Circulation Model (GCM). GCM is a global-scale numerical model that serves as a reference in long-term climate studies and provides primary information for assessing the impacts of climate change (Estiningtyas & Wigena, 2011). However, GCM outputs generally have relatively coarse spatial resolution, making them less appropriate for direct use at the local scale (Wigena, 2006). To bridge this scale difference, the Statistical Downscaling (SD) approach can be used to convert global-scale climate information from GCM into information that is more suitable for local conditions. SD is considered effective because it can produce more accurate local-scale climate data than using GCM outputs directly (Sahriman et al., 2014). Nevertheless, SD implementation often faces a large number of GCM predictors and high inter-grid correlations due to spatial proximity and temporal linkage (Soleh et al., 2015). This condition leads to multicollinearity, which can make coefficient estimation unstable and increase its variance (Kyriazos & Poga, 2023).

To handle multicollinearity, several shrinkage methods have been developed. Ridge Regression (RR) adds a shrinkage parameter to make the model more stable and reduce the variability of regression coefficients (Hoerl & Kennard, 1970). However, shrinkage in RR is uniform for all coefficients, so the resulting bias is not always small, especially when some predictors actually carry strong information. Liu Regression (LR) was then introduced with a constant (d) as an alternative to reduce multicollinearity (Liu, 1993), but its performance depends heavily on the choice of the constant d. If d is not chosen properly, the results can exhibit an unfavorable bias–variance trade-off. In other words, both RR and LR still have limitations when applied to high-dimensional and strongly correlated predictors such as GCM data. (Kibria & Lukman, 2020) formulated Kibria–Lukman Regression (KLR) as an alternative method to address this problem. KLR combines the advantage of RR in minimizing variance with the bias-correction concept of LR to handle multicollinearity. KLR works by adding an adaptive bias element to the model that adjusts the degree of shrinkage based on the information contained in each variable, thereby providing a more optimal trade-off between bias and variance.

Previous studies by (Kibria & Lukman, 2020) applied KLR to address multicollinearity in modeling the composition data of Portland cement and French economic data. The results showed that KLR provided the best performance for both datasets, indicated by the lowest mean square error compared to several competing methods. Furthermore, (Dawoud et al., 2022) compared the predictive performance of KLR with various other multicollinearity-reduction approaches and reported that KLR produced better predictions. In addition, (Lukman et al., 2023) developed the application of KLR in logistic regression for categorical data related to cancer symptom categories and obtained superior performance compared to other estimation methods.

Although KLR has shown good results in various contexts, its application in SD modeling for rainfall forecasting based on high-dimensional GCM predictors remains relatively limited, especially at the local scale that is directly related to salt production activities. In addition, local-scale rainfall often fluctuates from very dry periods to very wet periods, while continuous predictors from GCM do not always fully capture these differences, so additional predictors are needed to enrich local information in the model.

Based on the above description, this study applies an SD model with the KLR approach for rainfall forecasting in Pangkep Regency. This study aims to forecast rainfall for the 2023–2024 period based on the KLR model with the addition of dummy variables. It is expected that this approach can produce a more stable model under multicollinearity conditions and provide more accurate rainfall forecasting, thereby being more useful for salt production planning in Pangkep Regency.

2. METHODOLOGY

2.1. Data Sources

This study used precipitation data from GCM outputs of the Climate Model Intercomparison Project Phase 5 (CMIP5), in units of mm/month, obtained through the KNMI Climate Explorer (the Netherlands) at https://climexp.knmi.nl/. The GCM grid domain used covered an 8×8 grid within the range 14.75°S–5.25°N and 109.54°E–129.54°E, thereby producing 64 predictor variables. The observed rainfall data was the average monthly rainfall from three rainfall stations in Pangkajene and Islands (Pangkep) Regency, namely Bungoro, Ma'rang, and Labakkang, for the period January 1999 to December 2024, obtained from BMKG Region IV Makassar.

The response variable (Y) in this study was monthly rainfall intensity (mm/month) in Pangkep Regency during January 1999–December 2024. The predictor variables (X) consisted of 67 variables, namely 64 precipitation variables for the same period and 3 dummy variables. These dummy variables are derived from the optimal clustering of rainfall data in Pangkep Regency (Sahriman et al., 2019). Furthermore, the data were divided into two parts, namely training data for model development for the period January 1999–December 2022 and testing data for validation for the period January 2023–December 2024. An illustration of the distribution of the GCM output grids over Pangkep Regency is presented in Figure 1.

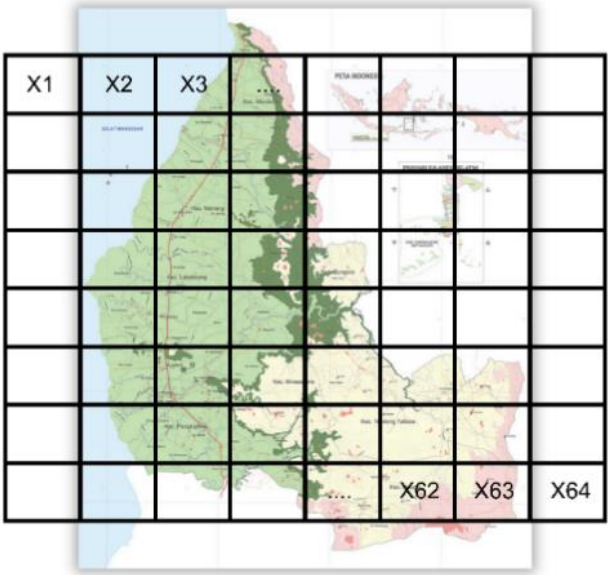

Figure 1. GCM Grid Domain (8×8) over Pangkep Regency

Figure 1 shows an 8×8 GCM grid domain covering Pangkep Regency and the surrounding area. Each grid cell represents one GCM grid labeled X_1 to X_{64}, so the GCM outputs were used as 64 predictor variables. In this study, the predictor taken from each grid was precipitation in units of mm/month, so the value X_j represents GCM output precipitation at the j −th grid. In addition to

predictors from GCM outputs, this study also added dummy variables as predictors. The dummy variables were constructed by clustering the response variable Y using the K-means clustering method with an optimal number of 4 clusters, so 3 dummy variables $(k_m - 1)$ were required to represent cluster membership in the regression model. The determination of the optimal number of clusters was based on the Within Sum of Squares (WSS) values using the Elbow method (Figure 2).

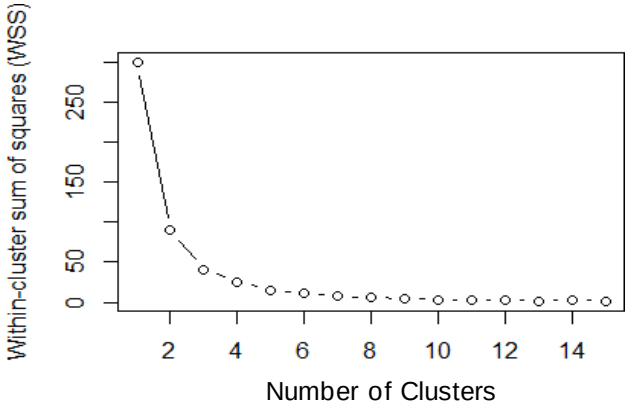

Figure 2 Elbow plot of WSS for optimal cluster determination

Figure 2 showed that the WSS value decreased significantly as the number of clusters increased from $k = 1$ to $k = 3$, then the decrease began to slow at $k = 4$ and tended to level off for subsequent values of k. This pattern indicated the presence of an elbow point at $k = 4$, suggesting that adding more clusters beyond this point did not result in a meaningful reduction in variability. Therefore, the optimal number of clusters selected in this study was 4.

2.2. Methods

2.2.1. Statistical Downscaling

GCM **is** one of the main tools for studying climate systems. However, GCM outputs are generally global-scale climate information with relatively low spatial resolution, so they are less adequate when used directly at the local scale. Therefore, statistical downscaling is used to link global-scale climate variables with higher-resolution local-scale climate variables so that predictions at the local scale can be made with better accuracy (Araya-Osses et al., 2020). Statistical downscaling (SD) is a statistical approach that models the relationship between global-scale predictors (GCM outputs) and local-scale responses (observations) to transform information/anomalies at the global scale into estimates of local climate variables (Sahriman et al., 2024). The SD performance is generally good when several assumptions are satisfied, namely: there is a strong correlation between predictors and responses, predictors are adequately simulated by the GCM, and the relationship between predictors and responses is stable over time, including under future changing climate conditions (Busuioc et al., 2001; Wigena et al., 2015). The general form of the SD model can be expressed in Equation (1) as follows:

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

where $\mathbf{y}_{(n\times 1)}$ denotes a local climate variable (e.g., rainfall), $f(\cdot)$ is a function that can be linear or nonlinear, $\mathbf{X}_{(n\times p)}$ is the predictor matrix derived from GCM outputs, n denotes the number of observations in time units (e.g., daily or monthly), and p is the number of predictors (e.g., the number of grids in the GCM domain).

2.2.2. Kibria-Lukman Regression

GCM-output precipitation data used as predictors in the statistical downscaling model have a large dimension, so the predictor variables tend to be strongly correlated and induce multicollinearity in the regression model. This condition needs to be addressed because it can make coefficient estimates unstable, increase the variance of the estimators, and reduce rainfall forecasting accuracy. One commonly used approach is shrinkage, namely adding a shrinkage constant in the estimation process to stabilize coefficient estimation. Hoerl & Kennard (1970) introduced Ridge Regression (RR), which uses a non-negative bias parameter k. In orthogonal form, the RR estimator is stated as in Equation (2) as follows:

$$\hat{\boldsymbol{\alpha}}_{RR} = (\boldsymbol{\Lambda} + k\boldsymbol{I})^{-1}\boldsymbol{Z}'\boldsymbol{y} \tag{2}$$

where $\hat{\boldsymbol{\alpha}}_{RR}$ is the vector of ridge coefficient estimates in orthogonal form with size $p \times 1$; $\mathbf{Z} = \mathbf{XP}$ is the predictor matrix transformed into the orthogonal space with size $n \times p$; $\mathbf{y}$ is the response vector with size $n \times 1$; $\boldsymbol{\Lambda}$ is a diagonal matrix with size $p \times p$ whose diagonal elements $\lambda_j (j = 1,2, \ldots, p)$ are the eigenvalues of the matrix $\mathbf{X}'\mathbf{X}$; and $k \geq 0$ is the ridge parameter (shrinkage constant) that controls the degree of coefficient shrinkage. The ridge coefficients in the original space can be recovered through $\hat{\boldsymbol{\beta}}^*_{RR} = \mathbf{P}\hat{\boldsymbol{\alpha}}_{RR}$, where $\hat{\boldsymbol{\beta}}^*_{RR}$ denotes the coefficients without an intercept.

The parameter k shrinks the coefficient variances across all components. However, uniform shrinkage can produce relatively large bias, particularly for coefficients that contain strong information. To reduce this bias, Liu (1993) proposed Liu Regression (LR) with a shrinkage constant d that provides a smoother effect. The general LR estimator is written in Equation (3) as follows:

$$\hat{\boldsymbol{\alpha}}_{LR} = (\boldsymbol{\Lambda} + \boldsymbol{I})^{-1}(\boldsymbol{\Lambda} + d\boldsymbol{I})\hat{\boldsymbol{\alpha}}_{OLS} \tag{3}$$

where d is the Liu shrinkage parameter (generally $0 < d < 1$) and $\hat{\boldsymbol{\alpha}}_{OLS}$ is the Ordinary Least Squares (OLS) estimator in orthogonal form, i.e., $\hat{\boldsymbol{\alpha}}_{OLS} = \boldsymbol{\Lambda}^{-1}\mathbf{Z}'\mathbf{y}$. The parameter d controls the degree of shrinkage so that bias can be better controlled than under uniform shrinkage in RR. Nevertheless, the LR estimator is sensitive to the choice of d. An inappropriate choice of d can increase the bias–variance trade-off.

Kibria & Lukman (2020) then proposed Kibria–Lukman Regression (KLR) in the linear model by introducing a new parameter k_{KL} to balance bias and variance more effectively. The KLR estimator is a development that combines shrinkage characteristics as in RR and adjustment as in LR. In orthogonal form, the KLR estimator is stated in Equation (4) as follows:

$$\hat{\boldsymbol{\alpha}}_{KL} = (\boldsymbol{\Lambda} + k_{KL}\mathbf{I})^{-1}(\boldsymbol{\Lambda} - k_{KL}\boldsymbol{I})\hat{\boldsymbol{\alpha}}_{OLS} \tag{4}$$

where k_{KL} is the Kibria–Lukman constant. In Equation (4), the matrix $(\boldsymbol{\Lambda}+k_{KL}\mathbf{I})^{-1}$ performs variance shrinkage as in RR, while the component $(\boldsymbol{\Lambda} - k_{KL}\mathbf{I})\hat{\boldsymbol{\alpha}}_{OLS}$ acts as in LR, namely readjusting excessive shrinkage so that model bias becomes more optimal. The regression coefficients in the original space are obtained through the transformation in Equation (5) as follows:

$$\hat{\boldsymbol{\beta}}_{KL} = \boldsymbol{P}\hat{\boldsymbol{\alpha}}_{KL} \tag{5}$$

The value of k_{KL} can be determined by computing the optimal value for each component j using Equation (6) as follows:

$$k_{j\,KL} = \frac{\hat{\sigma}^2}{2\hat{\alpha}^2_{j\,OLS} + \left(\frac{\hat{\sigma}^2}{\lambda_j}\right)} \tag{6}$$

The determination of k_{KL} can also use several comparative formulas. One option is to select the minimum value as written in Equation (7) (Özkale & Kaçiranlar, 2007):

$$k_{KL_{min}} = min\left(\frac{\hat{\sigma}^2}{2\hat{\alpha}_{j\,OLS}^2 + \left(\frac{\hat{\sigma}^2}{\lambda_j}\right)}\right) \tag{7}$$

Another alternative is the formula of Hoerl et al. (1975) that uses a weighted average in Equation (8) as follows:

$$k_{KL_{HMN}} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p}\left(2\hat{\alpha}_{j\,OLS}^2 + \left(\frac{\hat{\sigma}^2}{\lambda_j}\right)\right)} \tag{8}$$

where $\hat{\sigma}^2$ is the error variance computed using Equation (9) as follows:

$$\hat{\sigma}^2 = \left(\frac{(\boldsymbol{y}-\boldsymbol{Z}\hat{\boldsymbol{\alpha}}_{OLS})'(\boldsymbol{y}-\boldsymbol{Z}\hat{\boldsymbol{\alpha}}_{OLS})}{(n-p)}\right) \tag{9}$$

3. RESULT AND DISCUSSION

3.1 Descriptive Statistics

The descriptive analysis of rainfall in Pangkep Regency for the 1999–2024 period is conducted using the average monthly rainfall data from three rain gauges, namely Bungoro, Ma'rang, and Labakkang. To illustrate the seasonal pattern and data distribution, the surface monthly rainfall is presented in Figure 3, while the statistical summary in the form of the mean, standard deviation, minimum, and maximum for each month is provided in Table 1. This presentation is used to provide an overview of the rainfall tendency and its variability level for each month.

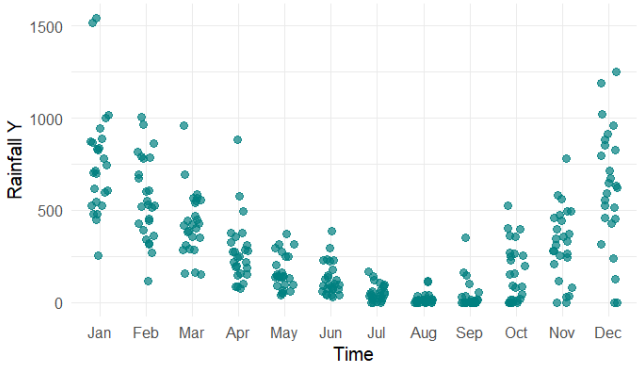

Figure 3. Monthly Rainfall in Pangkep Regency for the 1999–2024 Period

Based on Figure 3, surface rainfall in Pangkep tended to be high in January–March and increased again in November–December, while relatively low rainfall occurred in June–October. This pattern formed a strong seasonal character resembling a U-shape, which reflected a clear difference between the wet period (late to early in the year) and the dry period (mid-year). In addition to changes in monthly means, the scatter of points in Figure 3 also showed that interannual variation was larger in the wet months, indicated by a wider spread and the presence of several extreme values, whereas in the dry period the values tended to cluster in a low range and in some years reached 0 mm/month.

Table 1. Descriptive Statistics of Monthly Rainfall in Pangkep Regency, 1999–2023

Month	Mean	Standard Deviation	Minimum	Maximum
Jan	764.132	292.742	255.500	1540.500
Feb	564.241	225.168	116.000	1006.500
Mar	429.910	175.527	150.000	958.000
Apr	274.653	173.919	75.500	884.000
May	163.373	93.471	42.500	374.300
Jun	128.612	89.567	31.500	385.500
Jul	52.887	45.794	0.000	169.500
Aug	18.617	31.077	0.000	117.000
Sep	36.997	77.844	0.000	350.000
Oct	160.827	156.662	0.000	524.300
Nov	315.245	196.643	0.000	779.500
Dec	623.306	326.188	0.000	1252.500

The summary in Table 1 reinforced these findings. The highest maximum monthly rainfall occurred in January, namely 1540.500 mm/month, while the minimum value of 0 mm/month appeared in several months, indicating the presence of months without rainfall in certain years. Overall, the average monthly rainfall was 296.490 mm/month, with the highest averages in January (764.132 mm/month) and December (623.306 mm/month), while the lowest average occurred in August (18.617 mm/month). In terms of variability, the largest standard deviation was found in December (326.188 mm/month) and the smallest was in August (31.077 mm/month), so rainfall fluctuations in the wet months tended to be higher than those in the dry months. After the surface rainfall characteristics were described, the seasonal pattern was then compared with the GCM output precipitation as a predictor as shown in Figure 4.

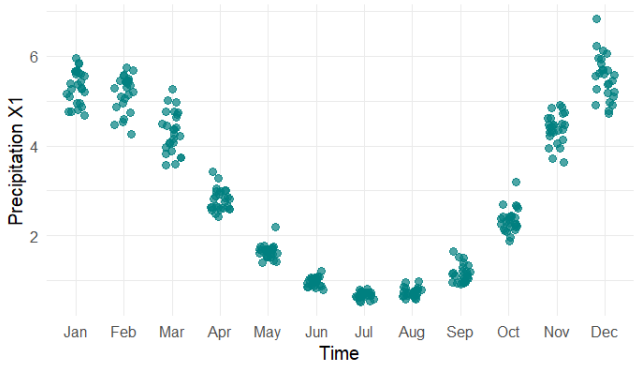

Figure 4. Monthly GCM Output Precipitation Plot (X_1), 1999–2024

Figure 4 shows GCM output precipitation at grid 1 (X_1), which showed a seasonal pattern consistent with rainfall in Pangkep Regency in Figure 3, namely higher at the end to the beginning of the year and lower in mid-year. A similar pattern was also observed in other grids (X_2 to X_{64}), so all GCM output precipitation variables were considered able to represent the main seasonal rainfall pattern in Pangkep and were suitable for use in statistical downscaling modeling. Furthermore, because GCM precipitation predictors come from many adjacent grids and tend to move following similar patterns, correlations among variables are potentially high. Therefore, before the regression modeling stage is carried out, the predictor variables need to be evaluated through a multicollinearity test; however, first, dummy variables are constructed based on the clustering results of rainfall data in Pangkep Regency using the K-means clustering technique.

3.2 Dummy Variables as Additional Predictors

In addition to the GCM output precipitation predictors, this study adds dummy variables as additional predictors to capture differences in rainfall characteristics that are not fully reflected by the continuous GCM variables. The dummy variables are formed based on the clustering results of observed rainfall data in Pangkep Regency using the K-means technique, so that variations in rainfall patterns can be represented into several relatively homogeneous groups. The results of dummy variable construction are presented in Table 2.

Table 2. K-Means Clustering and Dummy Variables

No	Time	Y	Cluster	D_1	D_2	D_3
1	Jan-1999	1017.5	4	0	0	1
2	Feb-1999	427	2	0	1	0
3	Mar-1999	444	2	0	1	0
4	Apr-1999	577	2	0	1	0
5	May-1999	202.5	3	1	0	0
6	Jun-1999	61.5	3	1	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
60	Dec-2003	1252.5	1	0	0	0
61	Jan-2004	526.5	2	0	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
311	Nov-2024	242.67	2	0	1	0
312	Dec-2024	914.67	4	0	0	1

Table 2 summarized the clustering results of monthly rainfall in Pangkep Regency for the January 1999–December 2024 period using the K-means method with an optimal number of four clusters, which was selected based on cluster validity indices. The four clusters represented different levels of rainfall intensity, namely cluster 1 for very high rainfall (1189.5–1540.5 mm) that occurred rarely and generally appeared at the peak of the rainy season (December–February), cluster 4 for high rainfall (around 619–1019 mm) that occurred relatively often in the early to mid-rainy season (December–March), cluster 2 for moderate intensity (240.5–607 mm) that often occurred in transition seasons, and cluster 3 for low rainfall (0–235.7 mm) that dominated the dry season, especially July–October. For statistical downscaling modeling purposes, the four clusters were then encoded into three dummy variables (D_1, D_2, D_3) with cluster 1 as the reference category (all dummy variables equal to 0), while the other clusters used a one-hot scheme. From 312 observations, the number of members in each cluster sequentially was 4, 105, 163, and 40, which indicated the dominance of low–moderate rainfall conditions, while high rainfall events occurred seasonally. The next stage is to conduct a multicollinearity test by evaluating VIF values for the GCM predictor variables.

3.3 Multicollinearity Test of GCM Precipitation Data

Multicollinearity occurs when predictor variables have a strong linear relationship, so that the information carried by each variable overlaps. In this study, multicollinearity is evaluated using the Variance Inflation Factor (VIF) for 64 GCM output precipitation predictor variables (X_1–X_{64}) for the 1999–2024 period. The VIF results are presented in Table 3 and indicate very strong multicollinearity in the GCM predictors.

Table 3. VIF Values of 64 GCM Precipitation Variables

Predictor	VIF	Predictor	VIF	Predictor	VIF
X_1	**41.88**	X_{23}	8.30	X_{45}	**51.40**
X_2	**469.35**	X_{24}	6.93	X_{46}	**25.33**

Table 3. VIF Values of 64 GCM Precipitation Variables

Predictor	VIF	Predictor	VIF	Predictor	VIF
X_3	**525.46**	X_{25}	**1619.46**	X_{47}	**49.81**
X_4	**412.33**	X_{26}	**2069.97**	X_{48}	**89.42**
X_5	**26.77**	X_{27}	**1359.37**	X_{49}	**618.48**
X_6	**16.96**	X_{28}	**851.28**	X_{50}	**2291.27**
X_7	**40.64**	X_{29}	**155.94**	X_{51}	**1182.32**
X_8	**20.68**	X_{30}	**43.28**	X_{52}	**958.21**
X_9	**439.31**	X_{31}	4.08	X_{53}	**46.07**
X_{10}	**1141.16**	X_{32}	**28.68**	X_{54}	**21.42**
X_{11}	**1030.30**	X_{33}	**1833.81**	X_{55}	**36.55**
X_{12}	**1012.19**	X_{34}	**2329.46**	X_{56}	**107.22**
X_{13}	**117.36**	X_{35}	**1216.22**	X_{57}	**65.93**
X_{14}	**34.53**	X_{36}	**554.47**	X_{58}	**949.37**
X_{15}	**22.22**	X_{37}	**38.30**	X_{59}	**587.17**
X_{16}	**10.53**	X_{38}	**24.43**	X_{60}	**423.35**
X_{17}	**1122.69**	X_{39}	**48.83**	X_{61}	**69.41**
X_{18}	**1343.42**	X_{40}	**108.10**	X_{62}	**29.53**
X_{19}	**829.68**	X_{41}	**1534.64**	X_{63}	**50.12**
X_{20}	**898.57**	X_{42}	**2565.04**	X_{64}	**80.39**
X_{21}	**103.57**	X_{43}	**1248.96**		
X_{22}	7.85	X_{44}	**401.72**		

Based on Table 3, the VIF values of the GCM predictors ranged from 4.08 to 2291.27, with most predictors having VIF > 10. Several predictors even showed very large VIF values, for example X_{50}(2291.27) and X_{42}(2565.04), which indicated very high linear dependence among certain grids. This pattern was consistent with the characteristics of GCM data, because adjacent grids generally had relatively similar atmospheric dynamics, which increased correlations among predictors. If used directly in regression, this condition could cause coefficients to become unstable, standard errors to increase, and predictive ability to decrease. Therefore, a more robust estimation method for multicollinearity was required, and in this study Kibria–Lukman Regression (KLR) was used as the approach to handle it.

3.4 Statistical Downscaling Modeling with Kibria–Lukman Regression

3.4.1 Transformation of Predictors and Dummy Variables into an Orthogonal Space

The initial stage of modeling was conducted by forming an orthogonal transformation of the predictor matrix consisting of 64 GCM output precipitation variables and 3 dummy variables. This transformation was constructed through spectral decomposition of the matrix $\boldsymbol{X'X}$, thereby obtaining eigenvalues and eigenvectors that became the basis for forming an orthogonal matrix. The eigenvalues for all predictors were presented in Table 4.

Table 4. Eigenvalues of the $\mathbf{X'X}$ Matrix

Symbol	Eigenvalue	Symbol	Eigenvalue	Symbol	Eigenvalue
λ_1	873483.664	λ_{24}	31.258	λ_{47}	5.014
λ_2	94033.296	λ_{25}	30.753	λ_{48}	4.539
λ_3	4143.694	λ_{26}	26.857	λ_{49}	4.077
λ_4	1600.034	λ_{27}	26.205	λ_{50}	3.848
λ_5	1148.945	λ_{28}	24.193	λ_{51}	3.583
λ_6	854.888	λ_{29}	21.096	λ_{52}	3.357
λ_7	392.694	λ_{30}	19.333	λ_{53}	2.926
λ_8	210.761	λ_{31}	18.371	λ_{54}	2.758
λ_9	172.994	λ_{32}	17.108	λ_{55}	2.471
λ_{10}	126.594	λ_{33}	15.012	λ_{56}	2.290

Table 4. Eigenvalues of the $\mathbf{X}'\mathbf{X}$ Matrix

Symbol	Eigenvalue	Symbol	Eigenvalue	Symbol	Eigenvalue
λ_{11}	111.812	λ_{34}	14.371	λ_{57}	2.245
λ_{12}	108.839	λ_{35}	13.247	λ_{58}	2.016
λ_{13}	87.343	λ_{36}	12.483	λ_{59}	1.878
λ_{14}	72.208	λ_{37}	11.365	λ_{60}	1.782
λ_{15}	69.049	λ_{38}	10.150	λ_{61}	1.621
λ_{16}	67.919	λ_{39}	9.890	λ_{62}	1.317
λ_{17}	59.332	λ_{40}	9.762	λ_{63}	0.922
λ_{18}	54.792	λ_{41}	9.020	λ_{64}	0.751
λ_{19}	49.177	λ_{42}	7.832	λ_{D1}	0.622
λ_{20}	46.594	λ_{43}	7.130	λ_{D2}	0.445
λ_{21}	38.228	λ_{44}	6.729	λ_{D3}	0.370
λ_{22}	37.207	λ_{45}	6.277		
λ_{23}	33.783	λ_{46}	5.216		

The eigenvalues in Table 4 were then used to construct the eigenvector matrix $\boldsymbol{P}$, whose elements represented the directions of the principal components in the predictor space. A summary of the eigenvector matrix was presented in Table 5.

Table 5. Eigenvector Matrix $\boldsymbol{P}$

No	P_1	P_2	⋯	P_{64}	P_{D1}	P_{D2}	P_{D3}
1	-0.056	-0.082	⋯	0.015	0.000	-0.003	0.013
2	-0.078	-0.098	⋯	-0.130	-0.155	-0.039	0.073
3	-0.093	-0.108	⋯	0.073	0.155	0.072	-0.071
4	-0.129	-0.112	⋯	-0.086	-0.145	-0.002	0.038
5	-0.135	0.145	⋯	0.017	0.018	0.025	0.013
6	-0.159	0.125	⋯	0.011	0.002	0.008	-0.002
7	-0.168	0.135	⋯	0.013	-0.010	-0.002	-0.009
⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮
64	-0.155	-0.039	⋯	0.039	0.003	0.013	-0.005
65	0.000	-0.001	⋯	-0.073	0.006	-0.034	0.001
66	-0.006	-0.007	⋯	0.003	-0.001	0.026	-0.026
67	-0.009	0.023	⋯	-0.010	0.015	0.031	-0.025

Based on the eigenvalues (Table 4) and eigenvectors (Table 5), the predictor matrix was then transformed into an orthogonal matrix $\boldsymbol{Z} = \boldsymbol{XP}$. This transformation produced orthogonal variables $(Z_1, \dots, Z_{64}, Z_{D1}, Z_{D2}, Z_{D3})$ that were uncorrelated with each other, so it was expected to reduce the impact of multicollinearity in the parameter estimation stage. The resulting orthogonal matrix values were presented in Table 6.

Table 6. Orthogonal Matrix Z

No	Z_1	Z_2	⋯	Z_{64}	Z_{D_1}	Z_{D_2}	Z_{D_3}
1	-67.688	-28.798	⋯	0.116	0.039	0.065	-0.076
2	-61.012	-20.636	⋯	0.031	0.053	-0.039	-0.048
3	-57.585	-11.056	⋯	0.050	0.093	0.063	-0.008
4	-58.122	1.204	⋯	0.008	0.007	0.091	0.011
5	-54.678	13.153	⋯	-0.078	-0.046	-0.052	0.040
6	-51.578	18.527	⋯	-0.016	-0.004	-0.039	0.019
⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮
283	-48.515	21.176	⋯	0.005	0.017	-0.002	-0.020

Table 6. Orthogonal Matrix Z

No	Z_1	Z_2	⋯	Z_{64}	Z_{D_1}	Z_{D_2}	Z_{D_3}
284	-46.094	21.218	⋯	-0.022	-0.012	-0.020	-0.004
285	-47.111	22.074	⋯	-0.026	0.011	-0.005	0.008
286	-49.012	12.059	⋯	-0.065	0.037	0.010	0.035
287	-54.105	-5.393	⋯	0.052	-0.010	0.011	-0.003
288	-64.321	-24.478	⋯	-0.022	-0.034	-0.058	0.039

The orthogonal matrix values in Table 6 could then be used to compute the OLS parameter estimator in the orthogonal space.

3.4.2 OLS Parameter Estimation in the Orthogonal Space

After the orthogonal matrix $\boldsymbol{Z}$ was obtained, the next step was to compute the OLS parameter estimator in the orthogonal space, namely $\hat{\boldsymbol{\alpha}}_{OLS}$, as the basis for calculating the Kibria–Lukman constant. The estimation results of $\hat{\boldsymbol{\alpha}}_{OLS}$ were presented in Table 7.

Table 7. OLS Parameter Estimates in the Orthogonal Space $(\hat{\boldsymbol{\alpha}}_{OLS})$

Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
α_{OLS_1}	-5.849	$\alpha_{OLS_{24}}$	15.649	$\alpha_{OLS_{47}}$	-14.649
α_{OLS_2}	-11.319	$\alpha_{OLS_{25}}$	29.371	$\alpha_{OLS_{48}}$	-370.342
α_{OLS_3}	4.634	$\alpha_{OLS_{26}}$	-47.913	$\alpha_{OLS_{49}}$	9.878
α_{OLS_4}	-1.830	$\alpha_{OLS_{27}}$	-53.648	$\alpha_{OLS_{50}}$	74.299
α_{OLS_5}	8.894	$\alpha_{OLS_{28}}$	77.439	$\alpha_{OLS_{51}}$	-61.569
α_{OLS_6}	2.356	$\alpha_{OLS_{29}}$	92.484	$\alpha_{OLS_{52}}$	196.493
α_{OLS_7}	-13.216	$\alpha_{OLS_{30}}$	118.124	$\alpha_{OLS_{53}}$	-195.187
α_{OLS_8}	-6.425	$\alpha_{OLS_{31}}$	-125.166	$\alpha_{OLS_{54}}$	165.932
α_{OLS_9}	-3.349	$\alpha_{OLS_{32}}$	14.530	$\alpha_{OLS_{55}}$	-99.942
$\alpha_{OLS_{10}}$	-21.213	$\alpha_{OLS_{33}}$	122.186	$\alpha_{OLS_{56}}$	-40.520
$\alpha_{OLS_{11}}$	-18.038	$\alpha_{OLS_{34}}$	39.515	$\alpha_{OLS_{57}}$	-46.498
$\alpha_{OLS_{12}}$	-6.663	$\alpha_{OLS_{35}}$	-104.097	$\alpha_{OLS_{58}}$	-86.377
$\alpha_{OLS_{13}}$	17.776	$\alpha_{OLS_{36}}$	212.446	$\alpha_{OLS_{59}}$	-165.408
$\alpha_{OLS_{14}}$	52.328	$\alpha_{OLS_{37}}$	-70.195	$\alpha_{OLS_{60}}$	61.142
$\alpha_{OLS_{15}}$	-2.462	$\alpha_{OLS_{38}}$	-298.918	$\alpha_{OLS_{61}}$	-202.034
$\alpha_{OLS_{16}}$	-67.018	$\alpha_{OLS_{39}}$	-34.327	$\alpha_{OLS_{62}}$	-110.622
$\alpha_{OLS_{17}}$	-11.505	$\alpha_{OLS_{40}}$	166.621	$\alpha_{OLS_{63}}$	47.634
$\alpha_{OLS_{18}}$	3.156	$\alpha_{OLS_{41}}$	296.960	$\alpha_{OLS_{64}}$	-54.932
$\alpha_{OLS_{19}}$	19.219	$\alpha_{OLS_{42}}$	-209.193	$\alpha_{OLS_{D1}}$	-99.732
$\alpha_{OLS_{20}}$	1.457	$\alpha_{OLS_{43}}$	-98.455	$\alpha_{OLS_{D2}}$	380.104
$\alpha_{OLS_{21}}$	-53.196	$\alpha_{OLS_{44}}$	-314.853	$\alpha_{OLS_{D3}}$	-252.896
$\alpha_{OLS_{22}}$	-56.730	$\alpha_{OLS_{45}}$	-136.028		
$\alpha_{OLS_{23}}$	67.857	$\alpha_{OLS_{46}}$	-166.094		

The estimation results in Table 7 were then used as the basis for calculating the Kibria–Lukman constant (k_{KL}).

3.4.3 Determination of the Kibria–Lukman Constant

The values of $\hat{\boldsymbol{\alpha}}_{OLS}$ in Table 7 were then used to compute the error variance $\hat{\sigma}^2$, which was required to determine the Kibria–Lukman shrinkage parameter. In this study, $\hat{\sigma}^2$ was computed as follows:

$$\hat{\sigma}^2 = \left(\frac{(\boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\alpha}}_{OLS})'(\boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\alpha}}_{OLS})}{(288 - 67)} \right) = 6938.869$$

The value of $\hat{\sigma}^2$ was then used to compute candidate k_{KL} for each component ($j = 1,2, \ldots, 64, D_1, D_2, D_3$), and then the minimum value $k_{(KL\_min)}$ was selected using Equation (7). Based on the calculation results, the minimum value was obtained for component D_2, namely $k_{(KL\_min)} = 0.023$.

3.4.4 Estimation of Kibria–Lukman Regression Parameters and Model Coefficients

After the constant k_{KL} was obtained, the Kibria–Lukman Regression parameters in the orthogonal space $\hat{\boldsymbol{\alpha}}_{KL}$ were computed to produce estimates that were more stable against multicollinearity. A summary of $\hat{\boldsymbol{\alpha}}_{KL}$ values was presented in Table 8.

Table 8. KLR Parameter Estimates in the Orthogonal Space $(\hat{\alpha}_{KL})$

Parameter	Estimate
α_{KL_1}	-5.849
α_{KL_2}	-11.319
α_{KL_3}	4.634
α_{KL_4}	-1.830
α_{KL_5}	8.893
⋮	⋮
$\alpha_{KL_{64}}$	-51.696
$\alpha_{KL_{D1}}$	-92.687
$\alpha_{KL_{D2}}$	343.077
$\alpha_{KL_{D3}}$	-223.553

Next, the parameters in the orthogonal space were converted back to the regression coefficients in the original space, thereby obtaining $\hat{\boldsymbol{\beta}}_{KL}$ as shown in Table 9.

Table 9. KLR Regression Coefficients in the Original Space $(\hat{\beta}_{KL})$

Parameter	Estimate
β_{KL_1}	-25.904
β_{KL_2}	-17.899
β_{KL_3}	99.927
β_{KL_4}	-15.150
β_{KL_5}	-33.474
⋮	⋮
$\beta_{KL_{64}}$	22.014
$\beta_{KL_{D1}}$	549.139
$\beta_{KL_{D2}}$	-340.020
$\beta_{KL_{D3}}$	-613.628

After the regression coefficients were obtained, the intercept $\beta_{(KL,0)}$ was computed using the mean of the response variable and the mean of the predictors. Based on substituting the mean values into the intercept equation, $\hat{\beta}_{KL,0} = 0.316$ was obtained. Thus, the KLR-based statistical downscaling model could be written as:

$$\hat{y}_{KL} = 0.316 - 25.904X_1 - 17.899X_2 + \cdots + 22.014X_{64} + 549.139D_1 - 340.020D_2 - 613.628D_3$$

The magnitude of the coefficients on the dummy variables indicated differences in effects across the rainfall groups formed by clustering. This confirmed the role of the dummy variables as additional predictors to represent local structures that were not fully captured by the GCM predictors.

3.5 Model Performance Evaluation

Based on the model that had been built, the next stage was to evaluate performance and determine the appropriate model. The evaluation was conducted using Root Mean Square Error (RMSE) to measure the magnitude of prediction errors on the training data and the coefficient of determination (R^2) to assess the proportion of rainfall variability that could be explained by the model. The comparison results of two model specifications, namely Kibria–Lukman without dummy (KL) and Kibria–Lukman with dummy (KLD), were presented in Table 10.

Table 10. Performance of the Kibria–Lukman Regression Models

Model	Variable	RMSE	R^2
Kibria-Lukman (KL)	$X_1, X_2, \ldots, X_{62}, X_{63}, X_{64}$	150.691	74.730%
Kibria-Lukman Dummy (KLD)	$X_1, \ldots, X_{64}, X_{D1}, X_{D2}, X_{D3}$	73.076	94.023%

Table 10 showed that adding dummy variables produced a significant improvement in performance. The KLD model had a smaller RMSE (73.076) than the KL model (150.691), which indicated that KLD predictions had lower errors. In line with that, the R^2 value for the KLD model was also higher (94.023%) than KL (74.730%), so the model with dummy variables was able to explain local rainfall variation better. With this consideration, the KLD model was selected as the superior model for the validation stage.

3.6 Model Validation on Testing Data

After the best model was established, validation was conducted to assess the model's generalization ability on data that were not used in modeling. The testing data used was monthly rainfall for the period January 2023 to December 2024. Validation performance was measured using Root Mean Square Error of Prediction (RMSEP) as a measure of prediction error and the correlation coefficient to evaluate the agreement of prediction patterns with actual data. The validation results were presented in Table 11.

Table 11. Performance on Testing Data

Model	Correlation	RMSEP
Ridge Dummy	0.934	96.129
Liu Dummy	0.925	111.575
Kibria-Lukman Dummy	0.936	94.614

Based on Table 11, the KLD model provided the best validation results compared with the other dummy models, namely Ridge Dummy (RD) and Liu Dummy (LD). The KLD model produced the highest correlation (0.936), which indicated strong agreement between the predicted pattern and the actual rainfall pattern. In addition, the RMSEP value for KLD (94.614) was also the smallest, so the prediction error was relatively lower than RD (96.129) and LD (111.575). This finding confirmed that the combination of the Kibria–Lukman shrinkage approach and additional information from the dummy variables improved prediction accuracy.

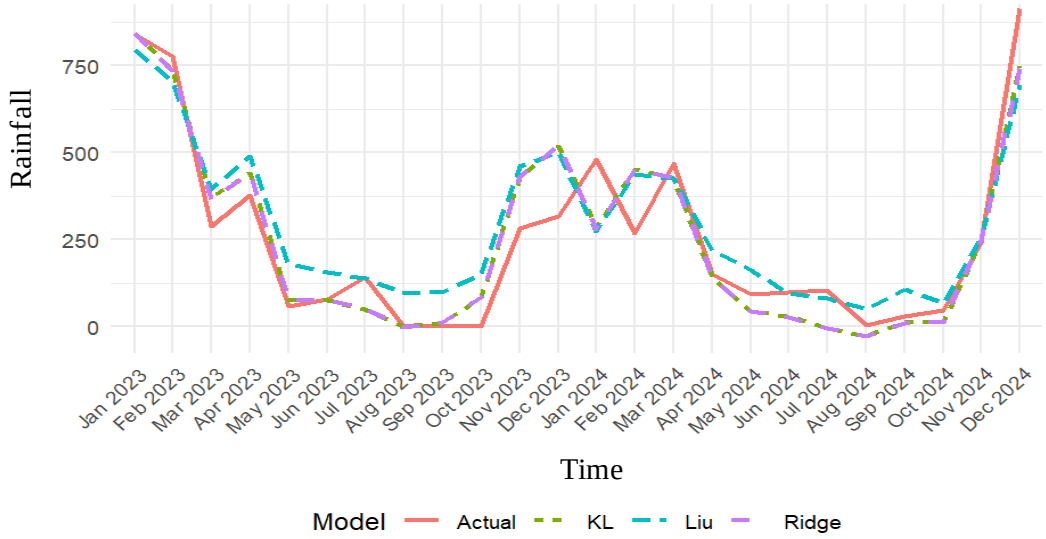

Figure 5. Monthly Rainfall Forecasting Results for Pangkep Regency, 2023–2024

Figure 5 showed the actual rainfall curve and the forecasting results from three models, namely KLD, RD, and LD. Visually, KLD showed the most consistent fit to the actual data and was able to follow seasonal changes more stably, both under high rainfall conditions and under low rainfall conditions. At the beginning of the year (around January to April), when rainfall was relatively high, KLD captured fluctuations better than RD and LD. In contrast, LD in several dry-season months (around June to September) tended to produce forecasts above the actual values (overestimation), whereas RD at several points was below the actual values and even approached zero, which indicated a tendency toward underestimation in certain periods. Toward the end of the year (around October to December), the three models followed the increasing direction of rainfall, but at the extreme peak (for example December 2024) the forecast values still tended to be lower than the actual values. This underestimation may have been caused by an unusually high rainfall event associated with short-term climate variability or anomalies, which were not fully captured by the models. Overall, the more stable proximity pattern of KLD throughout the testing period supported that this model provided the best representation of rainfall dynamics compared with RD and LD.

4. CONCLUSION

The statistical downscaling model in this study was developed using the Kibria–Lukman Regression approach to address multicollinearity that arose due to high correlations among precipitation grids of the Global Circulation Model data. Based on the analysis results, the obtained model utilized 64 Global Circulation Model precipitation predictors as well as three dummy variables as additional predictors. The model with dummy variables showed better performance, indicated by relatively small errors and a higher ability to explain rainfall variability in Pangkep Regency (R^2 around 94%) compared with the model without dummy variables (R^2 around 75%). In the validation stage using testing data for the 2023 to 2024 period, the model also provided

optimal results with a correlation of 0.936 and a root mean square error of prediction of 94.614, as well as better performance than other approaches in handling multicollinearity, so it was suitable for local-scale rainfall forecasting in Pangkep Regency. Thus, the dummy variables contributed as additional information that strengthened the representation of local rainfall variation and improved forecasting accuracy.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to the leadership of Hasanuddin University and the Institute for Research and Community Service (LP2M) for funding this study through the "Penelitian Dosen Pemula Unhas" (PDPA) scheme for the 2025 fiscal year (Contract No. 01260/UN4.22/PT.01.03/2025). We also thank the staff members who provided facilities and administrative support that contributed to the completion of this research.

REFERENCES

[1]. Akbar, M. A., Adrian, F., & Rahmatillah, L. F., 2023. Potensi dan Tantangan Produksi Garam Nasional. *ARMADA: Jurnal Penelitian Multidisiplin*, *1*(12), 1433–1438. https://doi.org/10.55681/armada.v1i12.1085

[2]. Araya-Osses, D., Casanueva, A., Román-Figueroa, C., Uribe, J. M., & Paneque, M., 2020. Climate change projections of temperature and precipitation in Chile based on statistical downscaling. *Climate Dynamics*, *54*(9–10). https://doi.org/10.1007/s00382-020-05231-4

[3]. Busuioc, A., Chen, D., & Hellström, C., 2001. Performance of statistical downscaling models in GCM validation and regional climate change estimates: Application for Swedish precipitation. *International Journal of Climatology*, *21*(5), 557–578. https://doi.org/10.1002/joc.624

[4]. Dawoud, I., Abonazel, M. R., & Eldin, E. T. A. G., 2022. Predictive Performance Evaluation of the Kibria-Lukman Estimator. *WSEAS Transactions on Mathematics*, *21*, 641–649. https://doi.org/10.37394/23206.2022.21.75

[5]. Estiningtyas, W., & Wigena, A. H., 2011. Teknik Statistical Downscaling dengan Regresi Komponen Utama dan Regresi Kuadrat Terkecil Parsial untuk Prediksi Curah Hujan pada Kondisi El Nino, La Nina, dan Normal. *Jurnal Meteorologi Dan Geofisika*, *12*(1), 65–72. https://doi.org/https://doi.org/10.31172/jmg.v12i1.87

[6]. Hanmina, M. M. F., Manesi, D., Ximenes, Y., Am'isa, J. C., Karibere, J. M., Bobo, E. K. R., & Berek, E. A. K., 2024. Sosialisasi Pengaruh Hukum Laut dan Hukum Maritim di Daerah Perbatasan RI-RDTL Desa Harekakae Kabupaten Malaka. *Journal Abdi Masyarakat Vokasi*, (1), 1–5.

[7]. Hidayat, M. F., Manaf, Muh. R., Yanuar, Y., Ibnusina, F., Hutahaean, A. A., Doktoralina, C. M., Caniago, A., Alisafira, S., Rumingkang, N. S., Sa'badini, S. A., Mangkurat, R. S. B., Kholil, & Prasetyo, T., 2023. *Menuju Puncak Pengintegrasian Rencana Tata Ruang Darat dan Laut*. Kementerian Koordinator Bidang Kemaritimandan Investasi. https://maritim.go.id/uploads/magazine/20240106110809-2024-01-06magazine110800.pdf

[8]. Hoerl, A. E., Kannard, R. W., & Baldwin, K. F., 1975. Ridge regression: some simulations. *Communications in Statistics*, *4*(2), 105–123. https://doi.org/https://doi.org/10.1080/03610927508827232

[9]. Hoerl, A. E., & Kennard, R. W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55–67. https://doi.org/https://doi.org/10.1080/00401706.1970.10488634

[10]. Kibria, B. M. G., & Lukman, A. F., 2020. A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*, *2020*. https://doi.org/10.1155/2020/9758378
[11]. Kyriazos, T., & Poga, M., 2023. Dealing with Multicollinearity in Factor Analysis: The Problem, Detections, and Solutions. *Open Journal of Statistics*, *13*(03), 404–424. https://doi.org/10.4236/ojs.2023.133020
[12]. Liu, K., 1993. A new class of blased estimate in linear regression. *Communications in Statistics - Theory and Methods*, *22*(2), 393–402. https://doi.org/https://doi.org/10.1080/03610929308831027
[13]. Lukman, A. F., Kibria, B. M. G., Nziku, C. K., Amin, M., Adewuyi, E. T., & Farghali, R., 2023. K-L Estimator: Dealing with Multicollinearity in the Logistic Regression Model. *Mathematics*, *11*(2). https://doi.org/10.3390/math11020340
[14]. Mukhlis., 2024. September 26). *Kebutuhan Industri Besar, Pj Gubernur Prof Zudan Ajak Lintas Stakeholder Kembangkan Potensi Garam di Sulsel*. https://sulselprov.go.id/post/kebutuhan-industri-besar-pj-gubernur-prof-zudan-ajak-lintas-stakeholder-kembangkan-potensi-garam-di-sulsel
[15]. Özkale, M. R., & Kaçiranlar, S., 2007. The Restricted and Unrestricted Two-Parameter Estimators. *Communications in Statistics - Theory and Methods*, *36*(15), 2707–2725. https://doi.org/https://doi.org/10.1080/03610920701386877
[16]. Sahriman, S., Djuraidah, A., & Wigena, A. H., 2014. Application of Principal Component Regression with Dummy Variable in Statistical Downscaling to Forecast Rainfall. *Open Journal of Statistics*, *04*(09), 678–686. https://doi.org/10.4236/ojs.2014.49063
[17]. Sahriman, S., Kalondeng, A., & Koerniawan, V., 2019. Pemodelan Statistical Downscaling Dengan Peubah Dummy Berdasarkan Teknik Cluster Hierarki Dan Non-Hierarki Untuk Pendugaan Curah Hujan. *Indonesian Journal of Statistics and Its Applications*, *3*(3), 295–309. https://doi.org/10.29244/ijsa.v3i3.471
[18]. Sahriman, S., Randa, E. L., Surianda, S. A., Zaky Gozhi Hisyam, M., Taufik, M. I., & Putra, G. D., 2024. Rainfall Forecasting of Salt Producing Areas in Pangkep Regency Using Statistical Downscaling Model with Linearized Ridge Regression Dummy. *Barekeng*, *18*(1), 483–492. https://doi.org/10.30598/barekengvol18iss1pp0483-0492
[19]. Soleh, A. M., Wigena, A. H., Djuraidah, A., & Saefuddin, A., 2015. Statistical Downscaling to predict monthly rainfall using linear regression with L1 regularization (LASSO). *Applied Mathematical Sciences*, *9*(105), 5361–5369. https://doi.org/10.12988/ams.2015.56434
[20]. Thahir, S., 2024, June 5. *Keberadaan Rumah Produksi Bersama Naikkan Produksi Garam Sulsel*. https://kabarika.id/berita/2024/06/05/keberadaan-rumah-produksi-bersama-naikkan-produksi-garam-sulsel/
[21]. Wigena, A. H., 2006. *Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah Hujan Bulanan: Kasus Curah Hujan Bulanan di Indramayu* [Disertasi, Institut Pertanian Bogor]. http://repository.ipb.ac.id/handle/123456789/40597
[22]. Wigena, A. H., Djuraidah, A., & Sahriman, S., 2015. Statistical Downscaling dengan Pergeseran Waktu Berdasarkan Korelasi Silang. *Jurnal Meteorologi Dan Geofisika*, *16*(1), 19–24. https://doi.org/https://doi.org/10.31172/jmg.v16i1.259