Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 26 (3) :324-330 (September - 2024)
https://doi.org/10.54386/jam.v26i3.2614
https://journal.agrimetassociation.org/index.php/jam

Research Paper

Bias correction and ensemble techniques in statistical downscaling model for rainfall prediction using Tweedie-LASSO in West Java, Indonesia

DHEA DEWANTI[1*], ANIK DJURAIDAH[1], BAGUS SARTONO[1] and ARDHASENA SOPAHELUWAKAN[2]

[1]*Department of Statistics, IPB University, Bogor 16680, West Java, Indonesia*
[2]*Agency for Meteorology Climatology and Geophysics (BMKG), Jakarta 10720, Indonesia*
*Corresponding Author: *dheadewanti29@gmail.com*

ABSTRACT

Rainfall is a climate element with high variations in space and time scales, so it is not easy to predict. One way to predict rainfall is statistical downscaling (SD). SD can predict local rainfall based on Global Circulation Model (GCM) data. The Decadal Climate Prediction Project (DCPP), one of the GCMs, originates from adjacent grids and experiences multicollinearity problems. Rainfall as a response variable is Tweedie Compound Poisson Gamma (TCPG) distribution data because it has a discrete component (rainfall events) and a continuous component (rainfall intensity), so SD modelling will be carried out using Tweedie-LASSO. This research aims to compare the performance of bias correction and ensemble methods in SD in predicting rainfall in West Java, Indonesia. Bias correction uses Empirical Quantile Mapping (EQM) with CHIRPS data, and the ensemble method uses a stacking technique with Random Forest (Stacking-RF) due to the varied characteristics of DCPP model sources. Evaluation results using Root Mean Square Error Prediction (RMSEP) and correlation coefficient show that bias correction improves single-model performance but not ensemble models. Besides that, ensemble models outperform single models both before and after bias correction. The combination of bias correction and ensemble modelling can be recommended when conducting SD to enhance the prediction capability of rainfall at stations and other areas.

***Keyword*:** Bias Correction, Empirical Quantile Mapping, Ensemble, Rainfall, Statistical Downscaling, Tweedie-LASSO

Rainfall is a climatic factor that exhibits significant variations across different locations and times, making it challenging to predict, but it has a vital role in tropical regions such as Indonesia (Swarinoto *et al*., 2012). The fundamental role of rainfall cannot be understated, as it significantly impacts a wide range of domains, including agriculture, forestry, plantations, irrigation, marine activities, infrastructure, and beyond. Information related to rainfall can be obtained through the Global Circulation Model (GCM), which results from numerical simulation forecasts. These GCMs were designed by various world climate institutions, which have variations in spatial resolution and the equations used to produce atmospheric parameters. Despite the usefulness of GCM output data, it presents challenges when linking it with local-scale rainfall data due to its global scope and large dimensions. Therefore, specific techniques are required to facilitate this process. According to Dar *et al*. (2018), Statistical Downscaling (SD) is capable of predicting local phenomenon like rainfall using output data from GCM. The SD technique will create a function that transfers information from the GCM output to local variables.

According to Dar *et al*. (2018), the use of GCM data for climate projections is still hampered because there is bias in the observation data, necessitating bias correction. Biases stem from imperfect model conceptualization, short data records, low-quality reference datasets, and poor spatial resolution. Various methods like Empirical Quantile Mapping (EQM) have been developed to address this, which Gudmundsson *et al*. (2012) note for their effectiveness in correcting rainfall prediction biases without assuming data distribution types. Bias correction in GCM output can be done using Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) data, as has been done by Nur *et al*. (2021) on rainfall in Sumatra. CHIRPS data is grid-based rainfall data combining rain station and satellite data with high spatial resolution.

***Article info - DOI*:** *https://doi.org/10.54386/jam.v26i3.2614*
Received: 3 June 2024; Accepted: 4 July 2024 ; Published online : 01 September 2024

Rainfall in SD can be categorized into discrete (no rain) and continuous (variable rain intensity) distributions. Dzupire *et al*. (2018) explain these categories, while Dunn (2004) introduces a joint distribution method between the two using the Tweedie family distribution, known as Tweedie Compound Poisson-Gamma (TCPG). Most SD research usually focuses on one distribution category. Rakhmalia *et al*. (2020) and Hayati *et al*. (2021) used the Tweedie regression model to predict rainfall in West Java. Hayati *et al*. (2021) also integrate LASSO to address multicollinearity in GCM output. However, these studies did not apply bias correction to the GCM data and relied on a single GCM model. Sa'adi *et al*. (2020) used Random Forest (RF) to improve predictions in Borneo using four GCM outputs, demonstrating the superiority of ensemble models over single GCM outputs. Stacking as an ensemble method has also been used by several researchers (Gu *et al*., 2022; Lu *et al*., 2023) in predicting climate elements. It shows better prediction robustness than a single model.

To enhance rainfall prediction accuracy in SD using four CMIP6 GCM outputs, we propose the use of the Tweedie-LASSO model with initial bias correction using CHIRPS data at six rainfall stations in West Java, Indonesia. Tweedie-LASSO is used as a model because Tweedie regression can simultaneously model discrete and continuous components of rainfall to improve future rainfall predictions. LASSO regularization is added to the model to handle violations of the multicollinearity assumption between explanatory variables in GCM data, which can cause the regression coefficient estimates to become unstable (Hayati *et al*., 2021). The Decadal Climate Prediction Project (DCPP) GCM was selected due to its focus on precise climate forecasts for the next decade. West Java, a flood and extreme weather-prone area (BNPB 2024) was chosen for this study. Predictions from each GCM will undergo an ensemble stacking process with RF as the meta-model. RF was selected as the meta-model due to its superior performance in a study by Fernandez-Delgado *et al*. (2014). It evaluated 179 classifiers across 17 families on 121 datasets from multiple disciplines and found RF the best model. We will compare the bias-corrected ensemble model against those without correction and individual GCM outputs. We will evaluate the best model based on its proximity to actual rainfall data, using RMSEP and correlation coefficient metrics.

MATERIALS AND METHODS

Study location

The location of this research is West Java, a province in Indonesia. West Java is geographically located between 5º50'-7º50' S and 104º48'-108º48' E, with a land area of 35377.76 km^2, a coastline of 724.85 km and a sea area of 155128.90 Ha. West Java features varied topography, including coastal and lowland areas in the north, highlands in the centre, and mountains in the south, leading to diverse climate conditions. Specifically, temperatures range from 16-34°C, and annual rainfall varies between 1,000-4,000 mm, with significant differences between the dry and rainy seasons. The study utilizes data from 6 rainfall stations across West Java to represent its topographic diversity: Krangkeng and Cibukamanah (lowlands), Kawali and Katulampa (midlands), and Cibeureum and Gunung Mas (highlands), as detailed in Table 1 and Fig. 1.

Table 1: Description of rainfall stations

Station	Statistical characteristics of rainfall (mm/month)			
	Mean	Sd	Min	Max
Cibukamanah	221.9	181.15	0	836.0
Krangkeng	110.3	104.06	0	639.0
Kawali	253.5	208.01	0	966.0
Katulampa	339.6	179.76	0	914.0
Cibeureum	175.6	219.19	0	706.2
Gunung Mas	303.0	134.23	0	1091.5

Sd: Standard Deviation; Min: Minimum value; Max: Maximum Value

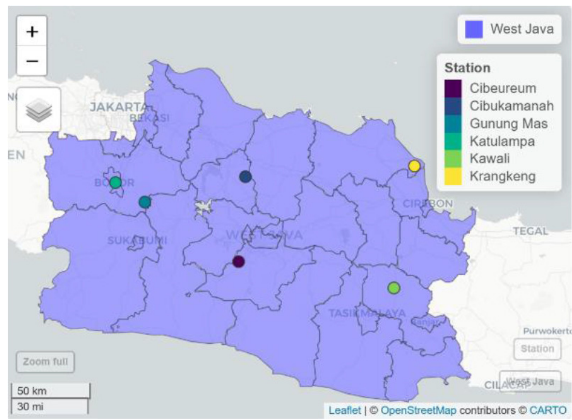


Fig. 1: Map of West Java and location of rainfall stations

Data

This research uses three types of secondary data from January 1991 to December 2020. The first is GCM CMIP6 monthly rainfall data, an explanatory variable in a 5×8 grid. The type of GCM used is DCPP, obtained from the page https://esgf-node.ipsl.upmc.fr/search/cmip6-ipsl/. The four DCPP models used are CNRM-ESM2, IPSL-CM6A-LR, MIROC6 and MPI-ESM1-2-LR. The second is CHIRPS monthly rainfall data obtained from the page https://iridl.ldeo.columbia.edu/SOURCES/.UCSB/.CHIRPS/. CHIRPS locations correspond to six rain stations with a grid size of 20 x 20 and a resolution of 0.05°x 0.05° each. This data will be used to correct bias in the GCM output data. The last is monthly rainfall intensity data from the Centre for Research and Development of the Agency for Meteorology Climatology and Geophysics (BMKG). The location of the rain stations is in West Java Province, with six stations, as in Table 1 and Fig. 1.

Analysis procedure

The data analysis procedures carried out in this research are as follows: (1) Conducting data exploration to determine the characteristics of rainfall data; (2) Splitting the data into training sets (1991–2017) and test sets (2018-2020); (3) Performing DCPP data bias correction on CHIRPS data using EQM; (4) Checking the distribution of rainfall data and estimating index parameters; (5) Doing SD rainfall modelling using two scenarios (uncorrected DCPP and corrected DCPP as the explanatory variable). Modelling was carried out using Tweedie-LASSO regression; (6) Combining rainfall predictions from each scenario using Stacking with Random Forest as a meta-model (k-fold = 3); (7) Evaluating the model by calculating RMSEP and correlation coefficient between actual data and predicted data; (8) Comparing RMSEP and correlation coefficient before and after bias correction as well as single models and ensemble models using the Paired t-test or Wilcoxon Paired test. If the p-value obtained is less than 5%, then at α=5%, it can be concluded that (a) bias correction successfully improves model performance and (b) the ensemble model successfully improves the performance of the single model.

Empirical quantile mapping (EQM)

EQM is a non-parametric technique for bias correction that aligns quantiles between predicted and observed data's CDF (Cumulative Distribution Function) without assuming any specific data distribution. Bias correction via EQM involves (1) Calculating empirical percentiles for both predicted and observed data; (2) Obtain the cumulative distribution function for each predicted and observed data from empirical percentiles. The values that fall between the given percentiles are computed through linear interpolation; (3) Perform bias correction with the following equation (Gudmundsson *et al.* 2012):

$$P_{cor} = F_o^{-1}\left(F_m(P_m)\right) \qquad (1)$$

where P_m is prediction data, F_m is CDF of P_m, F_o^{-1} CDF inverse of P_o , is observation data and is corrected prediction data.

Tweedie-LASSO regression model

The Tweedie distribution belongs to the exponential distribution family. It is included in the Generalized Linear Model (GLM), which uses a link function to relate the expected value to the linear model's systematic components (McCullagh and Nelder, 1989). For Tweedie regression, the link function is specified in equation (2) by Bonat and Kokonendji (2017).

$$\eta_i = g(\mu_i) = log(\mu_i) = \boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta} \qquad (2)$$

where xi and β are vectors of size $k\times 1$, which are explanatory variables and unknown regression coefficients.

LASSO reduces the regression coefficient of highly correlated explanatory variables to almost zero or exactly zero so that it can overcome multicollinearity. The coefficient estimation using the Tweedie-LASSO method can be obtained from the following equation (Qian *et al*. 2016):

$$\hat{\beta}_0, \widehat{\boldsymbol{\beta}} = \underset{\beta_0, \boldsymbol{\beta}}{\arg min} \left\{-\frac{log\{L(\beta_0,\boldsymbol{\beta})\}}{n} + \lambda \textstyle\sum_{j=1}^{k}|\beta_j|\right\} \qquad (3)$$

where $L(\beta_0,\boldsymbol{\beta})$ is the likelihood function of the observed data, n is the number of observations, λ is the tuning parameter (shrinkage parameter that controls the LASSO coefficient) with $\lambda \geq 0$, βj is the regression coefficient parameter and $-\frac{log\{L(\beta_0,\boldsymbol{\beta})\}}{n} = l(\beta_0,\boldsymbol{\beta})$ is the negative function of the log-likelihood of the Tweedie distribution. Estimating the parameters $L(\beta_0,\boldsymbol{\beta})$ cannot be done deductively using calculus but instead uses an optimization method called the IRLS-BMD algorithm, which is an algorithm that incorporates the blockwise majorization descent method into iteratively re-weighted least squares for parameter estimation (Qian *et al*. 2016).

Ensemble method

The ensemble method combines multiple models to decrease prediction errors. It involves two steps: first, generating ensemble members, in this case, using Tweedie-LASSO modelling with four DCPP models; second, merging their predictions using Stacking-Random Forest. Stacking employs various base models for parallel learning and then integrates their outcomes via a meta-model algorithm (Lu *et al*., 2023). Random forest, an extension of the CART algorithm, aggregates numerous trees for classification and regression, enhancing predictive accuracy (Breiman, 2001).

RESULTS AND DISCUSSION

Exploration of rainfall data and bias correction process

Data exploration was carried out on rainfall at the six selected stations mentioned in Table 1. In Fig. 2a, rainfall at all stations has a similar pattern, which resembles the letter U with a peak of the rainy season from the end to the beginning of the year. This pattern is by the BMKG (2021), which states that West Java is a province with a monsoon rainfall pattern. This pattern is unimodal, featuring a single peak season. Rainfall is lowest during June-September and highest in November-February. Altitude influences rainfall intensity. Lowland stations like Cibukamanah and Krangkeng experience less rainfall than midland and highland areas, with Gunung Mas Station in the highland area recording the highest rainfall. The histogram and density plot of rainfall at the six locations from 1991 to 2020 are shown in Fig. 2b. All rainfall at the six stations has a density plot that skews to the right and is positive. Almost all stations have a value of 0 that is greater than other observations except Katulampa Station.

The p index parameter value and 95% confidence interval estimated using maximum likelihood profile in Table 2 are also in the interval $1 < p < 2$. It can be confirmed that rainfall at all stations has a Tweedie distribution. The p index parameters in Table 2 will be used in the next part of SD modelling.

DCPP data bias correction will be done using CHIRPS data consisting of 400 grids. One DCPP data model consisting of 40 grids will be corrected repeatedly by 400 CHIRPS data grids according to the location of rainfall stations. The correction results by 400 grids are then averaged to obtain 40 corrected DCPP data grids. The correction factors obtained from the training data will be applied to the test data. The corrected training and test data will be used as explanatory variables to predict rainfall at each station.

Comparison of Tweedie-LASSO performance between terrain types

Table 3 shows that Tweedie-LASSO has different predictive abilities when viewed based on terrain type. The midlands have the highest average RMSEP value compared to other terrain types because they have diverse rainfall characteristics, as was found when estimating the index on the Tweedie distribution. The

Table 2: Parameter index *p* for all stations

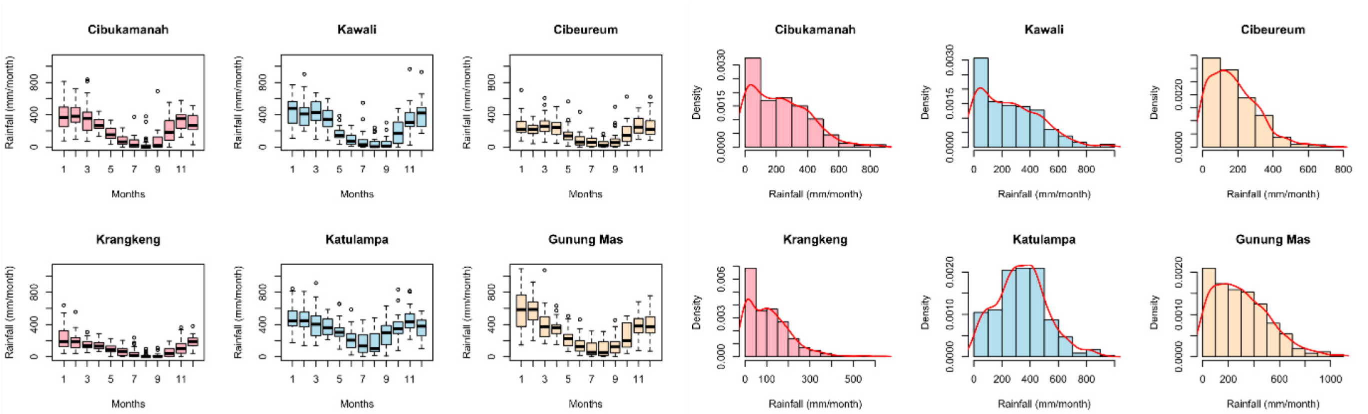
Station	*p*	95% CI
Cibukamanah[a]	1.47	(1.33; 1.49)
Krangkeng[a]	1.45	(1.41; 1.55)
Kawali[b]	1.57	(1.52; 1.63)
Katulampa[b]	1.39	(1.30; 1.50)
Cibeureum[c]	1.53	(1.45; 1.61)
Gunung Mas[c]	1.57	(1.56; 1.65)

a) Lowland; b) Midland; c) Highland: CI: Confidence Interval

Table 3: RMSEP and *r* by terrain type

Terrain Type	RMSEP (Mean ± Sd)	*r* (Mean ± Sd)
Lowland	90.02 ± 34.71	0.69 ± 0.17
Midland	161.30 ± 18.43	0.67 ± 0.14
Highland	119.30 ± 39.23	0.65 ± 0.14
p-value*	1.61 ± 10^{-14}	0.2385

*testing using Kruskal Wallis statistics; Sd: Standard Deviation



(a) (b)

Fig 2: (a) Boxplot, (b) Histogram and Density Plot of Rainfall at Six Stations

difference in the predictive ability of the model based on the terrain types is also supported by the results of statistical testing using the Kruskal Wallis test, with a p-value of 1.61×10^{-14}, which means that there is a difference in the average RMSEP value between the terrain at a significance level of α=0.05 . If we look at the correlation value, the three types of terrain have similar values, supported by a p-value of more than 0.05. This statistic means that the Tweedie-LASSO model has the same ability to predict rainfall patterns in all terrain types.

Comparison of model performance before and after bias correction

Overall, bias correction of DCPP data with CHIRPS data using EQM has various effects on the DCPP model and rainfall station locations. Fig. 3 is a boxplot of the RMSEP values and correlations measured on test data. Each point connected by a line is a pair of observations. The red line means that the bias correction decreases the RMSEP, and the blue line means that the bias correction increases correlation between predictions and actual data. The average decrease in RMSEP value and increase in correlation value after bias correction, along with the p-value of statistical testing, can also be seen in Fig. 3.

Based on Fig. 3a and 3b, it can be seen that the red and blue lines dominate the plot. This plot means bias correction reduces RMSEP and increases single-model correlation before bias correction. Statistical tests also support this statement, so it can be concluded that bias correction has significantly improved the performance of the single model both overall and based on terrain type. On ensemble models, bias correction has different effects. If seen based on the red and blue lines in Fig. 3c and 3d, only a few pairs of observations experience a decrease in RMSEP and an increase in correlation. The statistical test results also showed that bias correction does not improve the performance of the ensemble model, either overall or based on terrain type.

Comparison of the performance of a single model and an ensemble model

The boxplot in Fig. 4 shows that the ensemble model has a smaller RMSEP than the single model RMSEP and is dominated by the red line. The correlation boxplot also clearly shows that the ensemble model has a greater correlation with a more homogeneous distribution than the single model correlation and is dominated by the blue line. Statistical tests also support this statement, so it can be concluded that the ensemble model has better performance compared to the single model both overall and based on the bias correction method.

Ensemble models with DCPP models originating from various sources have been proven to have better abilities in predicting rainfall compared to using just one DCPP model. Using the ensemble model also reduced each model's bias so that there was no need to go through the bias correction stage first if using the ensemble model. Bias correction is required if only one DCPP

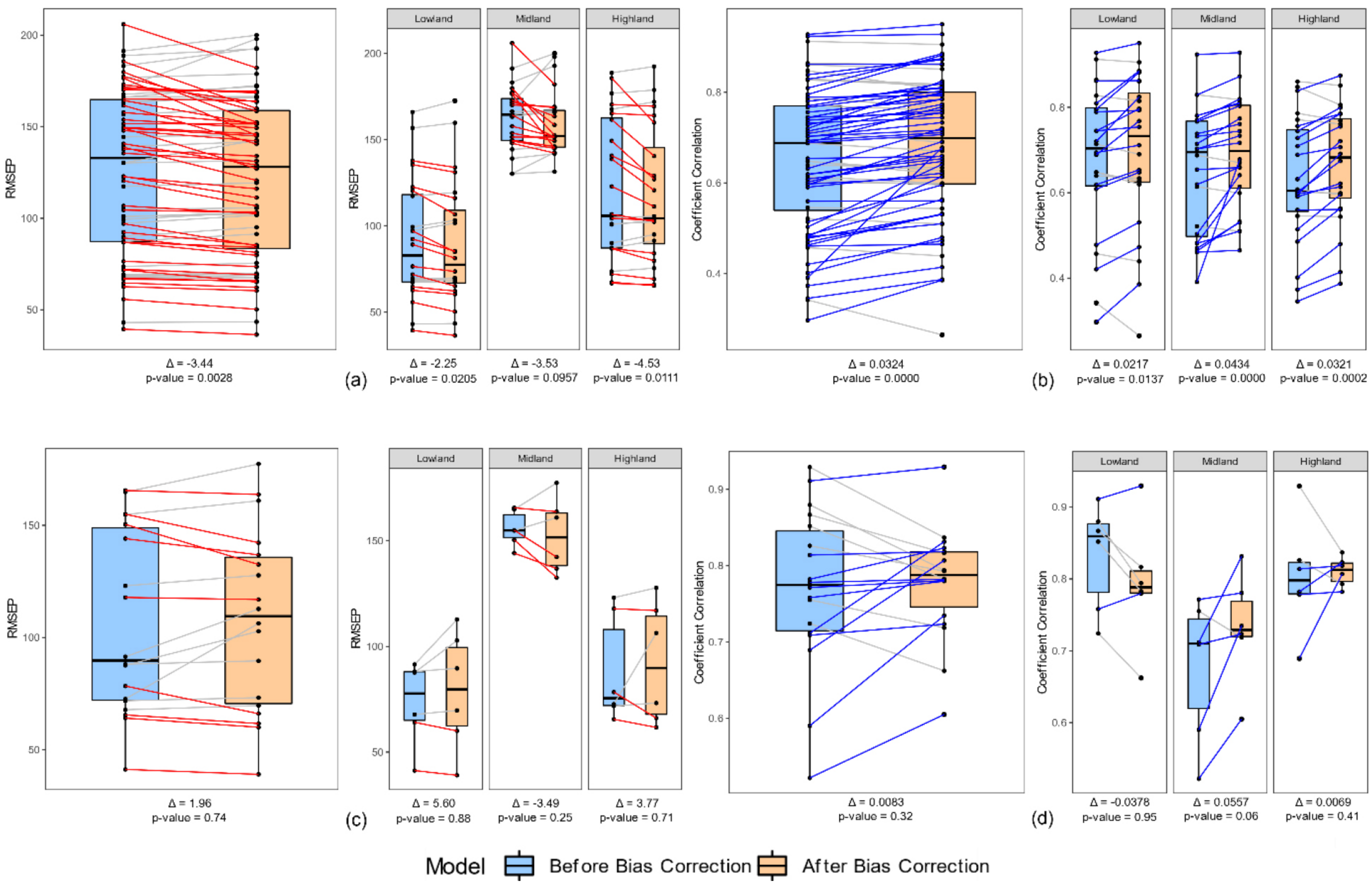


Fig. 3: Boxplot of (a) RMSEP of single model; (b) correlation of single model; (c) RMSEP of ensemble model; (d) correlation of ensemble model before and after bias correction (Δ: The differences between RMSEP or correlation after and before bias correction)

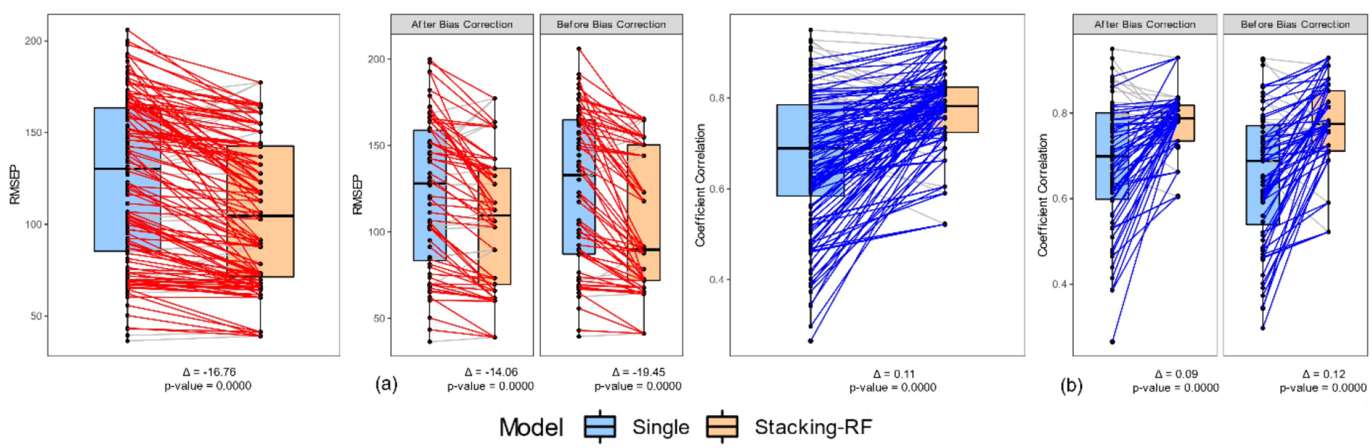
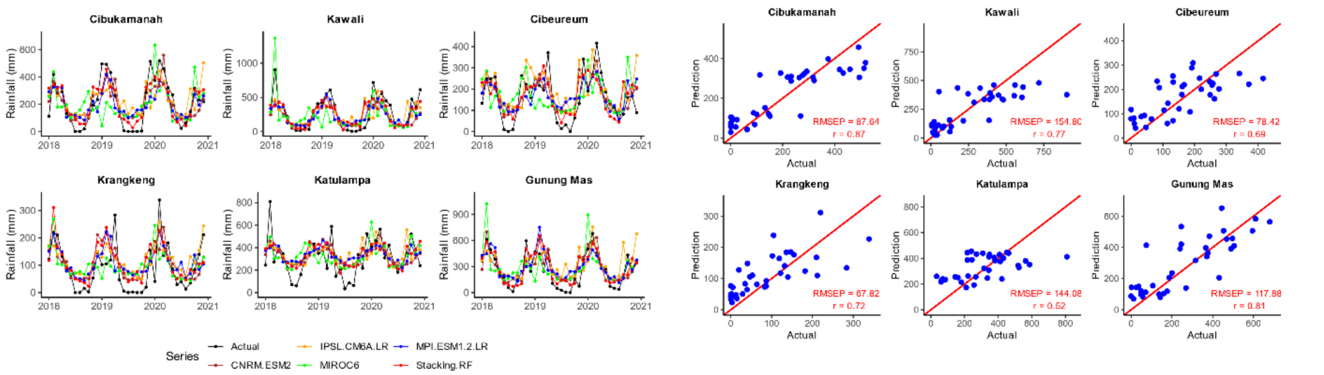


Fig. 4: Boxplot of (a) RMSEP; (b) correlation of single model and ensemble model (Δ: The differences between RMSEP or correlation ensemble model and single model)

model is used as an explanatory variable to predict rainfall.

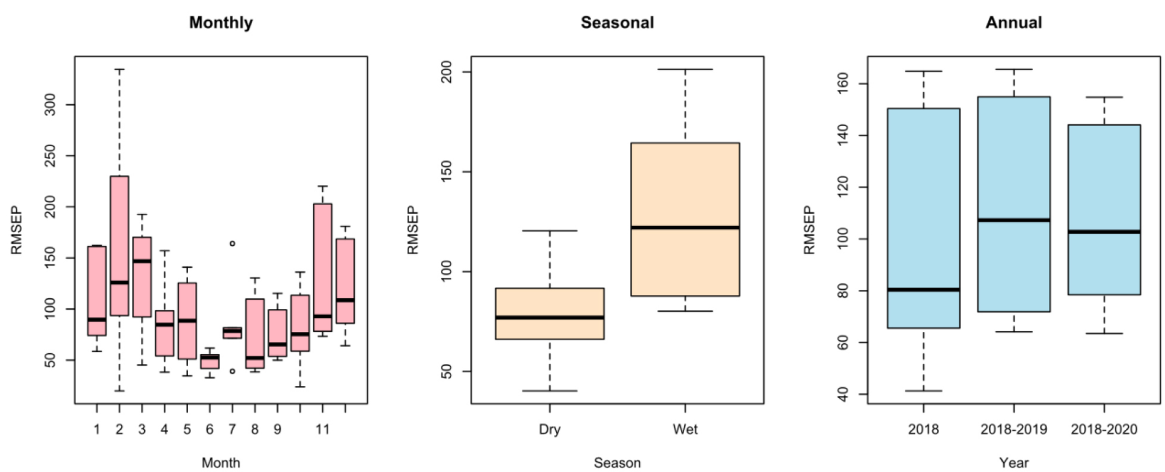
Plot predicted data versus actual data

Comparison plots of actual data and predictions from each DCPP model after bias correction and Stacking-RF before bias correction for the period January 2018 to December 2020 can be seen in Fig. 5a. All predicted values have a pattern that tends to be the same as actual data but less able to follow actual data with extreme values. Overall, the prediction value of the Stacking-RF model is better able to predict rainfall data with lower extreme values when compared to a single model. This can be seen from the red points, which are closer to the black points at some of the lower extreme values than the other coloured points, which represent a single model. The RMSEP and coefficient correlation of Stacking-RF without bias correction, which is the best model in this study, can be seen in Fig. 5b.



(a) (b)

Fig 5: Comparison plot of actual data and predictions from (a) each DCPP model after bias correction and Stacking-RF before bias correction; (b) Stacking-RF before bias correction



(a) (b) (c)

Fig 6: Boxplot of RMSEP Stacking-RF before bias correction (a) monthly; (b) seasonal; (c) annual

Best model performance

The RMSEP of the Stacking-RF model before bias correction, which is the best model, can be seen in Fig. 6. Model evaluation was carried out on monthly data from 2018 to 2020 to see the model's predictive capabilities on a monthly, seasonal, and annual basis. Indonesia, which is a tropical country, consists of two seasons, namely the dry season (May-October) and the wet season (November-April) (Mulsandi *et al*., 2024). The RMSEP of the model in Fig. 6a has a similar pattern to the rainfall boxplot in Fig. 2b. Months with high rainfall tend to have higher RMSEP, and vice versa. This finding is also supported by Fig. 6b, which shows that the dry season has a lower RMSEP than the wet season. On an annual basis, the model has a reasonably stable RMSEP for prediction periods of 1 to 3 years.

CONCLUSION

This research focuses on studying bias correction and ensemble methods in statistical downscaling to improve the ability to predict future rainfall. Based on the research that has been carried out, we obtained information that the bias correction can improve the performance of the Tweedie-LASSO single model but not the Stacking-RF ensemble. Ensemble models incorporating DCPP models from diverse sources unequivocally outperform single DCPP model predictions regarding rainfall accuracy. Leveraging an ensemble model reduces the bias inherent in individual models, rendering the bias correction step unnecessary. This finding contrasts sharply with the mandatory bias correction when relying solely on a single DCPP model for rainfall prediction. This combination of bias correction and ensemble models is highly recommended for conducting statistical downscaling (SD), as it significantly enhances the accuracy of rainfall predictions at stations and other areas.

ACKNOWLEDGEMENT

The author would like to thank the Agency for Meteorology Climatology and Geophysics (BMKG), the University of California and the Earth System Grid Federation (ESGF) website for providing the necessary data and the IPB University for supporting the research.

Conflict of Interests: The authors declare that there is no conflict of interest related to this article.

Funding: This research was funded by the Indonesia Endowment Fund for Education Agency (LPDP)

Data availability: DCPP and CHIRPS can be obtained from the link provided. Rainfall data is available upon request**.**

Authors contribution: **D. Dewanti**: Data collection, Analysis, Interpretation, Writing-original draft and Review; **A. Djuraidah**: Conceptualization, Methodology, Review, Supervision; **B. Sartono**: Conceptualization, Methodology, Supervision; **A. Sopaheluwakan**: Data collection, Conceptualization, Methodology, Supervision.

Disclaimer: The contents, opinions and views expressed in the research article published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

Publisher's Note: The periodical remains neutral with regard to jurisdictional claims in published maps and institutional affiliations

REFERENCES

BMKG. (2021). BMKG Deputy for Climatology. https://iklim.bmkg.go.id/publikasi-klimat/ftp/brosur/LEAFLETINGGRISB.pdf

BNPB. (2024). Indonesian Natural Disaster Data Geoportal. https://gis.bnpb.go.id/

Bonat, W. H., and Kokonendji, C. C. (2017). Flexible Tweedie regression models for continuous data. *J. Stat. Comput. Simul.*, 87(11): 2138–2152. https://doi.org/10.1080/00949655.2017.1318876

Breiman, L. (2001). Random forest. *Mach. Learn*., 45(1): 5–32. https://doi.org/10.1023/A:1010933404324

Dar, M. U., Aggarwal, R., and Kaur, S. (2018). Comparing bias correction methods in downscaling meteorological variables for climate change impact study in Ludhiana, Punjab. *J. Agrometeorol.*, 20(2): 126-130. https://doi.org/10.54386/jam.v20i2.523

Dunn, P. K. (2004). Occurrence and quantity of precipitation can be modelled simultaneously. *Int. J. Climatol.*, 24(10): 1231-1239. https://doi.org/10.1002/joc.1063

Dzupire, N. C., Ngare, P., and Odongo, L. (2018). A Poisson-Gamma model for zero inflated rainfall data. *J. Probab. Stat.*, 2018 (1012647): 1-12. https://doi.org/10.1155/2018/1012647

Fernandez-Delgado, M., Cernadas, E., Barro, S., and Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. *J. Mach. Learn. Res*., 15: 3133-3181.

Gu, J., Liu, S., Zhou, Z., Chalov, S. R., and Zhuang, Q. (2022). A stacking ensemble learning model for monthly rainfall prediction in the Taihu Basin, China. *Water*. 14(3): 492. https://doi.org/10.3390/w14030492

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Skaugen, T. E. (2012). Technical note: downscaling RCM precipitation to the station scale using quantile mapping – a comparison of methods. *Hydrol. Earth Syst. Sci. Discuss.*, 9: 6185–6201. https://doi.org/10.5194/hessd-9-6185-2012

Hayati, M., Wigena, A. H., Djuraidah, A., and Kurnia, A. (2021). A new approach to statistical downscaling using Tweedie compound Poisson Gamma respone and lasso regularization. *Commun. Math. Biol. Neurosci.*, 2021(60): 1-16. https://doi.org/10.28919/cmbn/5936

Lu, M., Hou, Q., Qin, S., Zhou, L., Hua, D., Wang, X., and Cheng, L. (2023). A stacking ensemble model of various machine learning models for daily runoff forecasting. *Water*. 15(7): 1265 https://doi.org/10.3390/w15071265

McCullagh, P., and Nelder, J. (1989). Generalized Linear Models (2nd ed.). London: Chapman and Hall.

Mulsandi, A., Koesmaryono, Y., Hidayat, R., Faqih, A., and Sopaheluwakan, A. (2024). Detecting Indonesian Monsoon Signals and Related Features Using Space–Time Singular Value Decomposition (SVD). *Atmosphere*. 15(187). https://doi.org/10.3390/atmos15020187

Nur, I. A., Hidayat, R., Latifah, A. L., and Misnawati. (2021). Effect of bias correction and ensemble method on rainfall data from four output regional climate model (RCM) CORDEX-SEA over Sumatera. *JPSL,* 11(1): 49-56. https://doi.org/10.29244/jpsl.11.1.49-56

Qian, W., Yang, Y., and Zou, H. (2016). Tweedie's compound Poisson model with grouped elastic net. *J. Comput. Graph. Stat.*, 25(2): 606-625. https://doi.org/10.1080/10618600.2015.1005213

Rakhmalia, R. I., Soleh, A. M., and Sartono, B. (2020). Pendugaan curah hujan dengan teknik statistical downscaling menggunakan clusterwise regression sebaran tweedie (rainfall estimation with statistical downscaling techniques using tweedie distribution clusterwise regression). *IJSA,* 4(3): 473-483. https://doi.org/10.29244/ijsa.v4i3.667

Sa'adi, Z., Shahid, S., Pour, S. H., Ahmed, K., Chung, E.-S., and Yaseen, Z. M. (2020). Multi-variable model output statistics downscaling for the projection of spatio-temporal changes in rainfall of Borneo Island. *J. Hydro. Environ. Res.*, 31(2020): 42-75. https://doi.org/10.1016/j.jher.2020.05.002

Swarinoto, Y. S., Koesmaryono, Y., Aldrian, E., and Wigena, A. H. (2012). Ensemble prediction system model for monthly rainfall total using weightes value (case of Indramayu district). *J. Meteor. Geofis.*, 13(3): 189-200.