



Comparison Performance of the Multi-Regional Climate Model (RCM) in Simulating Rainfall and Air Temperature in Batanghari Watershed

Unggul Handoko^{1,4,*}, Rizaldi Boer¹, Edwin Aldrian², Arnida Lailatul Latifah³,
Bambang Dwi Dasanto¹, Apip Apip⁴, Misnawati Misnawati⁵

¹Applied Climatology, Graduate School of Bogor Agricultural University, Bogor, Indonesia;

²Agency for The Assessment and Application of Technology, Serpong, Indonesia;

³ Research Center for Informatics LIPI, Cibinong, Indonesia;

⁴ Research Center for Limnology LIPI, Cibinong, Indonesia;

⁵Indonesian Agroclimate and Hydrology Research Institute, Bogor, Indonesia

*Corresponding author email: unggul@limnologi.lipi.go.id

Received : 18 November 2018

Accepted : 4 June 2019

Online : 4 June 2019

Abstract – The scientists believe that RCM output is directly used as input for climate change impact models, while it consists of systematic errors. Consequently, RCM still requires bias correction to be used as an input model. The purpose of this study was to analyze the RCM performance before and after bias correction, its best performance from several models, as well as to clarify the importance of bias correction before it is used to analyze climate change. As a result of this, the method used for bias correction was Distribution Mapping method (for rainfall) and Average Ratio-method (for air temperature). While the *Generalized Extrem Value* distribution (GEV) was used to analysis extreme rainfall. To determine the performance of the model before and after bias correction, statistical analysis was used namely R², NSE, and RMSE. Furthermore, ranking for every single model and Taylor Diagram was used to determine the best model. The results showed that the RCMs performance improved with bias correction. However, CSIRO-Mk3-6-0, CCSM4, GFDL-ESM2M, and MPI-ESM-MR models can be ignored as ensemble models, because they demonstrated poor performance in simulating rainfall. From this study, it was suggested that the best model in simulating daily and monthly rainfall was ACCESS1-0, while MIROC-ESM-CHEM (daily air temperature) and ACCESS1-0 (monthly air temperature) were best models used in simulating air temperature.

Keywords: RCM, bias correction, performance, rainfall, air temperature.

Introduction

The study of climate change is become as a popular topic and is beginning to be realized as part of human needs. This study is used to investigate the response of the climate system caused by several forces, simulate interactions of the atmosphere, oceans, and the earth surface. Furthermore, this model is used to make predictions over the next few decades (Flato *et al.*, 2013; IPCC, 2007). The climate model, widely used to build long-term data, was known as Global Climate Model (GCM). GCM was

generally suitable for reconstructing past climate conditions, as well as future climatic conditions on a broad scale. To increase its output on a more detailed scale, there was a need for a downscaling approach, by statistics or dynamically (Deidda *et al.*, 2013). The statistics and dynamic downscaling process were two approaches used to produce high-resolution regional climate models (RCM) based on large-scale information sourced from data reanalysis or GCM (Tang *et al.*, 2016). RCM tends to produce climate data with a spatial resolution between 25-50 km (Olsson *et al.*, 2014) or even 10-25 km (Murdock and Spittlehous, 2011). It is a reliable model for describing regional and even smaller scale climates. However, RCM still has systematic errors. Therefore it still requires bias correction to be used as input for impact models (Thiemeßl *et al.*, 2010; Ahmed *et al.*, 2012).

Scientists in need of RCM climate data with daily resolution tend to use products from *NASA Earth Exchange Global Daily Downscaled Projects* (NEX-GDDP) or *Coordinated Regional Climate Downscaling Experiments* (CORDEX). NEX-GDDP was a web provider of downscaling data from GCM originating from CMIP5. Its output has a spatial resolution of 25 km x 25 km and daily temporal resolution. The climate parameters produced were rainfall, maximum and minimum temperature with the baseline period (1950-2005), and projections (2006-2099). The characteristics of CORDEX were almost the same as NEX-GDDP, such both had a spatial and temporal resolution, as well as scenarios used for projection. However, the downscaling technique was used to distinguish between both products. The NEX-GDDP uses statistical downscaling techniques, while CORDEX uses dynamical downscaling techniques. Based on the spatial resolution, which was 25 km x 25 km, both were categorized in the Regional Climate Model.

Currently, many scientists have assumed that RCM output can be directly used as input for impact models, with the consideration that the spatial resolution produced was finer than the GCM, making it suitable for local scale studies. Also, there was a need for research on RCM performance before and after bias correction. Therefore, to determine the best performance of the model, it was necessary to evaluate their comparison and observational data. This was important because the best model of evaluation results tends to be used as input in the impact model (Meher *et al.*, 2017). As a result of this, the objectives of this study are to analyze the RCM performance before and after bias correction, and the best models in simulating rainfall and air temperature.

Materials and Methods

Time and site

The research was conducted from May to October 2018 in Batanghari watershed, which is located in Jambi Province and West Sumatra Province (100.5°E -104.5°E and 0°S – 3°S). The map of Batanghari watershed is seen in Figure 1.

Data collection

The data used in this study were rainfall data observations of 39 stations obtained from several sources (BMKG, BBWS Sumatera VI, and Ministry of Environment and Forestry), observed air temperature grid data downloaded via HTTP: //sacad .database.bmkg.go.id /grid /web, grid rainfall data and grid air temperature of the RCM model. In this study, 13 RCM models were used, namely ACCESS1-0, bcc-csm1-1, BNU-ESM, CanESM2, CCSM4, CESM1-BGC, CSIRO-Mk3-6-0, GFDL-ESM2M, inmcm4, MIROC-ESM-CHEM, MPI-ESM-MR, MRI-CGCM3, NorESM1-M. Data RCM can be downloaded from ftp://ftp.nccs.nasa.gov/BCSD/, except for CSIRO-MK3-6-0, GFDL-ESM2M, and MPI-ESM-MR obtained from BMKG. All rainfall and air temperature data, both observations and models have a grid spatial resolution of 0.25° and period from 1 January 1982 to 31 December 2005.

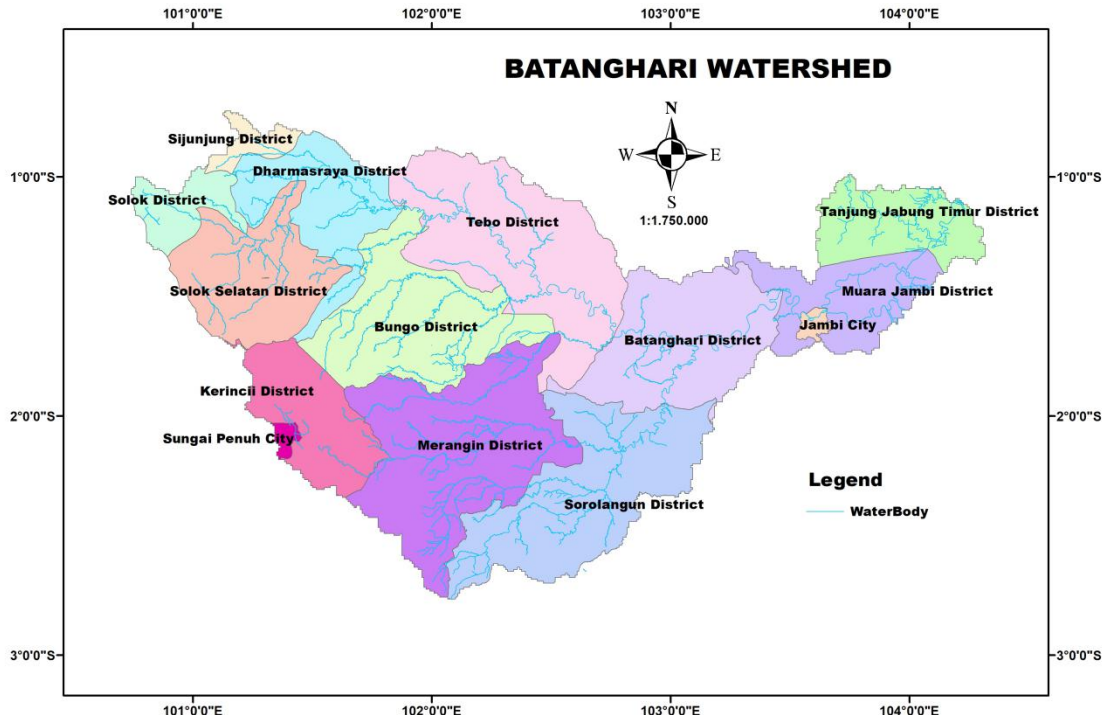


Figure 1. Location of the Batanghari watershed in Jambi and West Sumatra Provinces

Analysis of RCM performance before and after bias correction

Gridding observational data were adjusted to the grid from the output data of the RCM model. It was needed for easier to make it easier to make bias corrections of each observation grid (Dasanto *et al.*, 2014). The rainfall RCM data was corrected by distribution mapping method. Furthermore, for the purpose of extreme analysis, this method was more appropriate because it considers the distribution of rainfall, by assuming that its observation and RCM data were gamma distribution and probability density function as follows (Piani *et al.*, 2010):

$$pdf(x) = \frac{1}{b^a \Gamma(a)} x^{(a-1)} e^{\left(\frac{-x}{b}\right)}, x > 0$$

$$= 0, x \leq 0$$

Where, x = mean daily rainfall, letters a and b are the form and scale parameters respectively, $\Gamma(a)$ = Gamma function, solved by a factorial function ($\Gamma(a) = (a-1)!$)

Furthermore, the *cumulative distribution function* (CDF) was calculating by integrating the pdf equation above. After that, it makes the transfer function of the cumulative gamma distribution between rainfall observations and model. The transfer function used to correct the model was the polynomial regression equation (Misnawati *et al.*, 2018). Extreme rainfall was analyzed using the *Generalized Extreme Value distribution* (GEV). It was an effective method to describe very rare/extreme events (Supari, 2012). In this study, the intended extreme event was the maximum daily rainfall in one year. The GEV distribution equation is as follows.

$$f(x) = \begin{cases} \frac{1}{\alpha} \exp(-(1+kz)^{-1/k}) (1+kz)^{-1-1/k} & k \neq 0 \\ \frac{1}{\alpha} \exp(-z - \exp(-z)) & k = 0 \end{cases}$$

Where, $z = (x - \mu) / \sigma$ and k, σ, μ were of form, scale and location parameters, respectively. Scale parameters must be positive. The GEV distribution definition range depends on k :

$$1 + k \frac{(x - \mu)}{\sigma} > 0 \quad \text{for } k \neq 0$$

$$-\infty < x < +\infty \quad \text{for } k = 0$$

Various values for form parameters (k) correspond to the Gumbel, Frechet, and Weibull distributions as follows. $k > 0$ (Frechet), $k < 0$ (Weibull), $k = 0$ (Gumbel). The threshold for a particular return period utilized the following formula:

$$X_T = \mu - \frac{\alpha}{k} \left[1 - \left\{ -\log\left(1 - \left(\frac{1}{T}\right)\right) \right\}^{-k} \right]$$

T = return period

The RCM temperature data was corrected using Average Ratio-method (Lenderink *et al.*, 2007). The equation of the mean ratio method was:

$$T^* = T_{\text{mod}} \times \left[\frac{\mu_m T_{\text{obs}}}{\mu_m T_{\text{mod}}} \right]$$

Where, T^* = corrected temperature model, T_{mod} = temperature model, $\mu_m T_{\text{obs}}$ = monthly temperature mean observation, and $\mu_m T_{\text{mod}}$ = monthly temperature mean model.

After correcting the data model, the model accuracy test was conducted. Phogat *et al.*, 2016 stated that the assessment of model performance is essential because reliable statistical comparison of observed data with simulated behavior of a model reflects the performance and consistency. In this study, the accuracy of the model was tested with three statistical parameters namely mean square error, R squared, and Nash-Sutcliffe Efficiency.

$$\text{Mean square error} = \text{MSE} : \frac{1}{n} \sum_{i=1}^n [Z_{\text{sim}}(i) - Z_{\text{obs}}(i)]^2$$

$$R_{\text{squared}} = \left[\frac{\frac{1}{n} \sum_{i=1}^n Z_{\text{obs}}(i) \cdot Z_{\text{sim}}(i) - M_{\text{obs}} M_{\text{sim}}}{\left(\frac{1}{n} \sum_{i=1}^n Z_{\text{obs}}^2 - M_{\text{obs}}^2 \right) \left(\frac{1}{n} \sum_{i=1}^n Z_{\text{sim}}^2 - M_{\text{sim}}^2 \right)} \right]^2$$

Where, Z_{sim} = the value of the model and Z_{obs} = the value of observation.

The model showed to have good performance assuming the error value was small, and the correlation was high. Mukaka (2012), classifies the correlation (R) as in Table 1.

Table 1. Correlation value and its criteria to show the closeness of the relationship between the model and observation

Correlation	Criteria
0.9 to 1.0 (-0.9 to -1)	Very high correlation
0.7 to 0.9 (-0.7 to -0.9)	High correlation
0.5 to 0.7 (-0.5 to -0.7)	Moderate correlation
0.3 to 0.5 (-0.3 to -0.5)	Low correlation
0.0 to 0.3 (0 to -0.3)	Correlation can be ignored

Also, the Nash-Sutcliffe Efficiency (NSE) was also used to determine the closeness of the relationship between observation and model. A perfect model is similar to the observation indicated by the NSE value = 1. Its values range $\sim$ to 1, therefore, assuming a model has an NSE value close to 1, then the performance of the model was good. The equation for calculating its value was as follows (Moriassi *et al.*, 2007):

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^n (Y_i^{obs} - Y^{mean})^2} \right]$$

Where, Y_i^{obs} = the i_{th} observation for the constituent being evaluated, Y_i^{sim} = the i_{th} model for the constituent being evaluated i , Y^{mean} = the mean of observed data for the constituent being evaluated, n = the total number of observations. The NSE value criteria, as shown in Table 2 (Zhong and Dutta, 2015, Waseem *et al.*, 2017).

Table 2. NSE value and its criteria to indicate the acceptability of the model

NSE	Criteria
$\sim$ to 0	Unacceptable model
0 to 0.5	Unsatisfactory model
0.5 to 0.65	Satisfactory model
0.65 to 0.75	Good model
0.75 to 1	Very good model

Analysis of the best performance RCM models.

The best performance of the RCM model was analyzed by the "*ensemble performance ranking*" method. This method ranks the RMSE statistic parameters for the entire RCM model ensemble based on its climatology (Herger *et al.*, 2018). While the statistical parameters used for ranking in this study were MSE, R^2 , and NSE (calculations have carried out in step 1). Furthermore, the score for each single RCM model was obtained from the summation of ranks for three statistical parameters. The smallest score indicated the best performance.

In addition to using the "*ensemble performance ranking*" method, the performance of the best models was analyzed graphically using the Taylor Diagram, which indicates a pattern of closeness between models with observations based on correlation parameters (R), root means square error (RMSE) and amplitude standard deviation (Std) (Taylor, 2001).

Results

Comparison of RCM performance and rainfall

Comparison performance of each RCM model was carried out daily and monthly by making a long-term mean climatology from 1982 to 2005, with rainfall and temperature as the analyzed parameter. To write research results more systematically, the presentation of each climate parameter was made separately. Figure 2 shows the pattern of long-term daily and monthly mean rainfall between observations and 13 RCM models before and after bias correction. From Figure 2a, the model rainfall means value (solid line) indicates overestimate value compared to the observational value (blue dashed line). However, assuming it was observed, from the 13 models, a good number displayed similar patterns as the observation, the difference was that rainfall models are generally higher than observations. There are only three models whose patterns were not similar to observations, namely

CSIRO-Mk3-6-0, GFDL-ESM2M, and MPI-ESM-MR. Furthermore, after the bias correction was carried out, the pattern of daily rainfall mean was seen in Figure 2b, which was the same as an observation. However, there were still some models that show the value of outlier rainfall on certain days, such as CSIRO-Mk3-6-0, MPI-ESM-MR, and CCSM4.

Figure 2c shows the monthly rainfall mean pattern between observations and models before the bias correction was made. As with the daily rainfall mean, monthly rainfall means models also shows an overestimate value compared to observation. In this, there were three models which were far from observational patterns, namely CSIRO-Mk3-6-0, GFDL-ESM2M, and MPI-ESM-MR. However, the other ten models were not so far away from the observation pattern. Figure 2d shows the pattern of monthly rainfall mean between observations with the model after the bias correction was made. It is seen in the figure that the pattern of the model has followed the observation and the range of the mean value of rainfall models, which was also close to the mean value of rainfall observations. However, there were two models which in certain months failed to follow the observation pattern. The two models were CSIRO-Mk3-6-0 and MPI-ESM-MR.

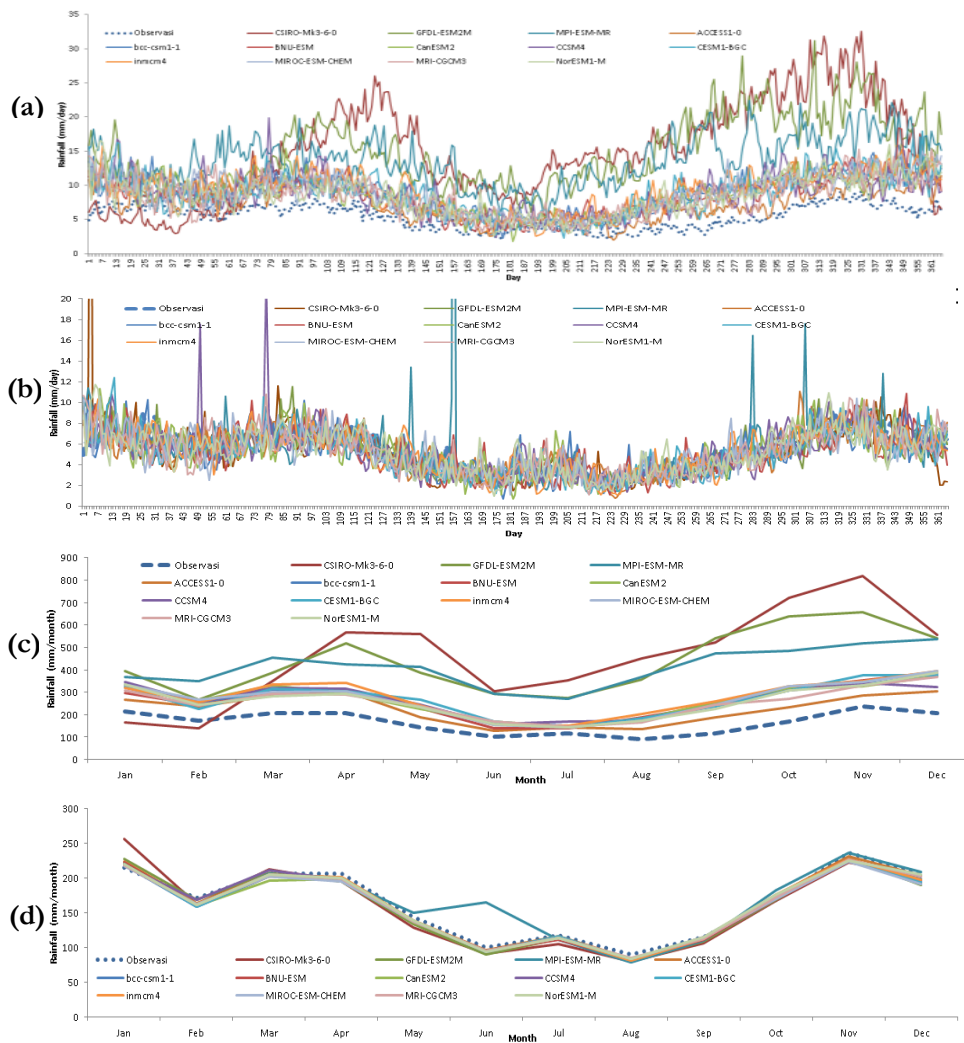


Figure 2. A pattern of observational rainfall and model; (a) uncorrected daily mean, (b) corrected daily mean, (c) uncorrected monthly mean, (d) corrected monthly mean.

Comparison of the performance of the RCM model before and after being corrected based on statistics are shown in Table 3. It was a summary of the statistical values with three parameters between uncorrected and corrected RCM. Based on Table 3, it was clear that the bias correction will improve the performance of the model in simulating the value of rainfall. Most of the R^2 value models before correction were already high, i.e., there are ten models with R^2 values above 0.8 (daily) and 0.9 (monthly). Values above 0.8 and 0.9 were included in the category of high and very high correlation, respectively. There were only three models with R^2 values below 0.8, namely CSIRO-Mk3-6-0, GFDL-ESM2M, and MPI-ESM-CHEM, with low and moderate correlation criteria. After correction, increasing of R^2 value for daily rainfall was not adequately significant, as it still ranges between 0.8. While for monthly rainfall, ten models increased to 1, with three models above 0.9 after the bias correction was conducted.

Table 3. Comparison of Statistical Analysis of the mean rainfall observations and models (Uncorrected and Corrected)

Model	Uncorrected			Corrected		
	R^2	NSE	RMSE	R^2	NSE	RMSE
<i>Mean daily rainfall</i>						
ACCESS1-0	0.89	-0.94	2.45	0.89	0.53	1.20
bcc-csm1-1	0.87	-3.72	3.81	0.87	0.43	1.33
BNU-ESM	0.84	-3.63	3.78	0.85	0.33	1.44
CanESM2	0.85	-3.85	3.87	0.85	0.33	1.44
CCSM4	0.84	-3.99	3.92	0.83	0.13	1.63
CESM1-BGC	0.82	-4.39	4.08	0.86	0.40	1.35
CSIRO-Mk3-6-0	0.46	-44.67	11.87	0.73	-1.50	2.78
GFDL-ESM2M	0.62	-35.10	10.55	0.88	0.47	1.27
inmcm4	0.83	-4.85	4.25	0.87	0.44	1.31
MIROC-ESM-CHEM	0.84	-4.10	3.97	0.84	0.29	1.48
MPI-ESM-MR	0.66	-23.71	8.73	0.59	-3.96	3.91
MRI-CGCM3	0.84	-2.99	3.51	0.88	0.45	1.30
NorESM1-M	0.84	-3.30	3.64	0.86	0.36	1.41
<i>Mean monthly rainfall</i>						
ACCESS	0.98	-0.77	63.87	1.00	0.99	5.05
bcc-csm1-1	0.94	-4.13	108.73	1.00	0.98	7.12
BNU-ESM	0.93	-3.86	105.82	1.00	0.97	7.86
CanESM2	0.94	-4.17	109.17	1.00	0.98	7.56
CCSM4	0.97	-3.76	104.70	1.00	0.98	6.14
CESM1-BGC	0.94	-4.34	110.92	1.00	0.97	7.75
CSIRO-Mk3-6-0	0.46	-52.54	351.29	0.98	0.90	15.41
GFDL-ESM2M	0.71	-37.07	296.21	0.99	0.96	9.35
inmcm4	0.93	-5.24	119.91	1.00	0.98	6.53
MIROC-ESM-CHEM	0.94	-4.40	111.59	1.00	0.97	7.86
MPI-ESM-MR	0.76	-27.53	256.44	0.96	0.82	20.29
MRI-CGCM3	0.95	-2.96	95.54	1.00	0.98	6.15
NorESM1-M	0.94	-3.51	101.92	1.00	0.98	6.33

Table 4. The return period of Observation and RCM Extreme Rainfall (mm/day) and NSE Value

Model	Return Period					NSE
	5	10	25	50	100	
Observation	45.64	54.53	71.55	90.31	116.47	-
ACCESS1-0	45.97	50.35	57.24	63.56	71.08	0.91
bcc-csm1-1	49.15	51.21	53.15	54.24	55.08	0.83
BNU-ESM	48.17	52.43	58.08	62.48	67.02	0.89
CanESM2	48.60	51.43	54.21	55.82	57.13	0.84
CCSM4	60.80	89.70	169.30	291.90	520.60	-5.74
CESM1-BGC	57.62	73.33	103.08	135.56	180.45	0.76
CSIRO-Mk3-6-0	126.20	211.80	447.40	811.30	1491.70	-80.00
GFDL-ESM2M	108.42	134.36	171.71	203.14	237.88	-0.50
inmcm4	47.66	53.64	61.91	68.61	75.79	0.93
MIROC-ESM-CHEM	49.51	51.34	52.95	53.79	54.41	0.83
MPI-ESM-MR	49508.80	51340.50	52954.50	53793.20	54410.10	-429671.16
MRI-CGCM3	57.43	67.70	83.90	98.78	116.52	0.98
NorESM1-M	47.51	49.97	52.43	53.87	55.05	0.83

The NSE value of all uncorrected RCM models was still negative, which means that it was poor at simulating the value of rainfall both daily and monthly. While a negative NSE value indicates that the model was unacceptable. However, after correction, a good number increased to positive with the highest NSE value at the ACCESS 1-0 model for both daily and monthly meant, which were 0.53 (for daily) and 0.99 (for monthly). The CSIRO-Mk3-6-0 and MPI-ESM-MR models NSE values remained negative when simulating daily rainfall. Table 3 also shows that the RMSE value of all models decreased after the bias correction was conducted, which indicates that the performance of the model after the correction has increased. A significant decrease in RMSE was seen in monthly rainfall simulations. Table 4 shows the performance of several models indicated by the value of NSE in simulating extreme rainfall based on GEV distribution with numerous return periods. Based on Table 4, the corrected model was able to simulate extreme rainfall properly. Nine out of from thirteen models analyzed showed NSE values above 0.75, which were included in the very good category in simulating extreme rainfall. Even the MRI-CGCM3 model has an NSE value close to 1, to be exact 0.98. However, the remaining four models (CCSM4, CSIRO-MK3-6-0, GFDL-ESM2M, and MPI-ESM-MR) have poor performance, as seen from NSE less than zero (negative value) on all four models.

Air temperature

Variability in air temperature was not adequate compared to rainfall. Therefore, the bias correction applied to the air temperature was simpler by using the average ratio method between observation and the model. Figure 3 compared the average pattern of air temperature between the models before and after the bias correction. Figure 3a shows that some models were overestimated and underestimated before the bias correction was conducted. Figure 3b displays the daily air temperature mean pattern after bias correction, with the observation and model pattern same, although daily variability still exists. However, figures 3c and 3d show a monthly mean pattern of air temperature before and after correction. With the bias correction, the pattern of monthly air temperature means was the same between observation and models. But there were models in certain months different from

the observation pattern, such as the bcc-csm1-1 model. A summary of statistical analysis for daily and monthly mean air temperatures, which indicates that it was very significant is presented in Table 5. The correlation value of all models was above 0.9, even for the monthly air temperature mean with a value of 1.

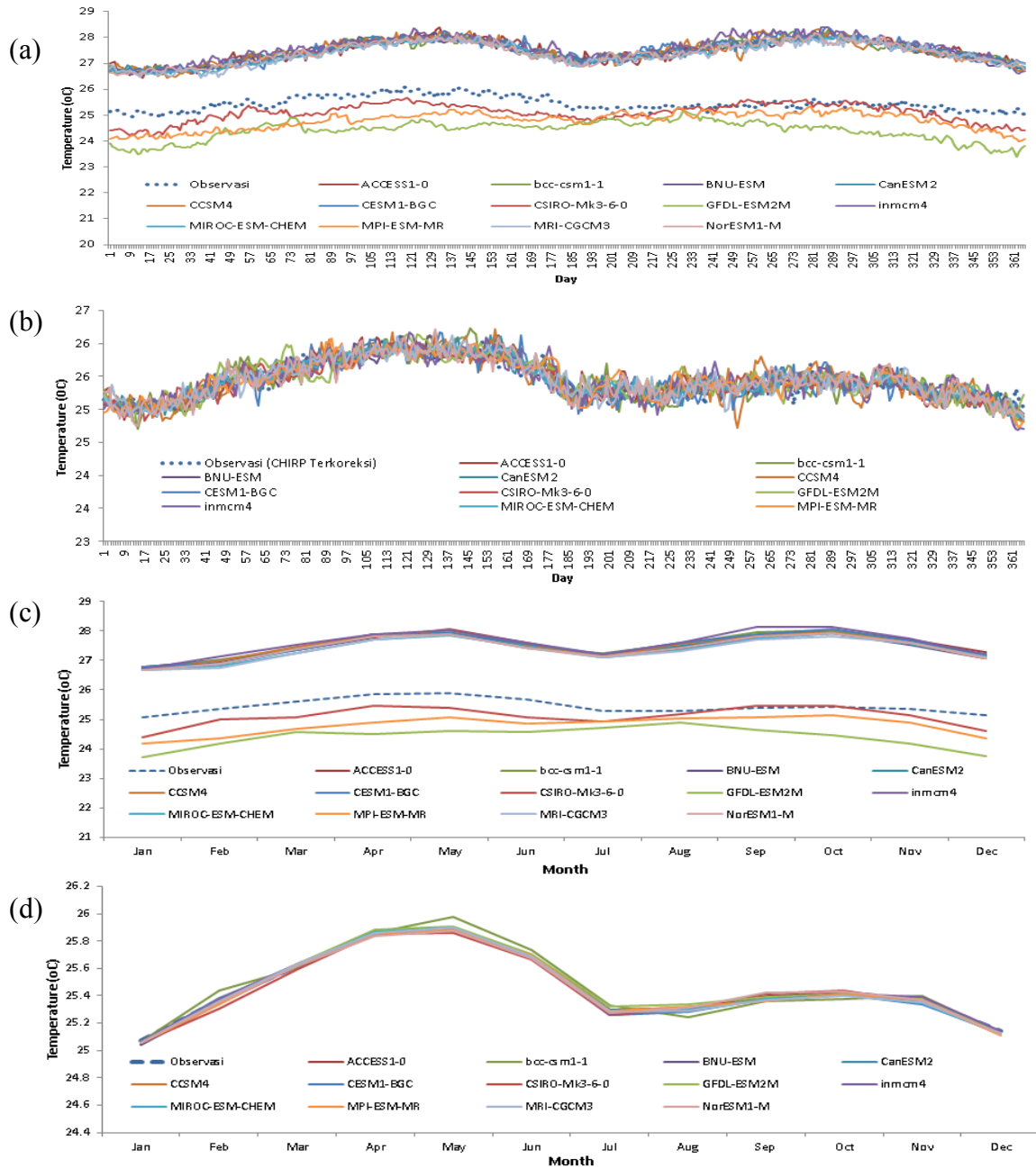


Figure 3. A pattern of observation air temperature and model; (a) uncorrected daily mean, (b) corrected daily mean, (c) uncorrected monthly mean, (d) corrected monthly mean.

Table 5. Comparison of Statistical Analysis of the mean air temperature
 Observations and models (Uncorrected and Corrected)

Model	Uncorrected			Corrected		
	R ²	NSE	RMSE	R ²	NSE	RMSE
<i>Mean daily temperature</i>						
ACCESS1-0	0.786	-55.535	2.072	0.928	0.693	0.153
bcc-csm1-1	0.752	-56.051	2.081	0.916	0.618	0.170
BNU-ESM	0.823	-53.095	2.027	0.939	0.746	0.139
CanESM2	0.803	-54.726	2.057	0.941	0.760	0.135
CCSM4	0.800	-54.679	2.056	0.912	0.606	0.173
CESM1-BGC	0.748	-56.069	2.082	0.931	0.712	0.148
CSIRO-Mk3-6-0	0.795	-1.630	0.447	0.934	0.744	0.139
GFDL-ESM2M	0.687	-15.044	1.104	0.918	0.660	0.161
inmcm4	0.761	-60.764	2.166	0.920	0.675	0.157
MIROC-ESM-CHEM	0.779	-49.725	1.962	0.952	0.807	0.121
MPI-ESM-MR	0.681	-6.064	0.732	0.927	0.704	0.150
MRI-CGCM3	0.780	-48.878	1.946	0.934	0.730	0.143
NorESM1-M	0.788	-51.752	2.001	0.925	0.689	0.154
<i>Mean monthly temperature</i>						
ACCESS	0.812	-66.783	2.064	1.000	0.998	0.011
bcc-csm1-1	0.780	-67.434	2.074	0.994	0.966	0.046
BNU-ESM	0.853	-63.890	2.020	1.000	0.997	0.013
CanESM2	0.833	-65.876	2.051	1.000	0.998	0.012
CCSM4	0.839	-65.593	2.046	0.999	0.997	0.014
CESM1-BGC	0.769	-67.431	2.074	0.999	0.995	0.017
CSIRO-Mk3-6-0	0.830	-1.869	0.425	0.998	0.989	0.026
GFDL-ESM2M	0.722	-18.019	1.094	0.999	0.993	0.021
inmcm4	0.791	-73.076	2.158	0.999	0.997	0.014
MIROC-ESM-CHEM	0.796	-59.913	1.957	1.000	0.997	0.013
MPI-ESM-MR	0.711	-7.224	0.719	0.999	0.995	0.019
MRI-CGCM3	0.806	-58.780	1.939	1.000	0.997	0.015
NorESM1-M	0.818	-62.248	1.994	0.999	0.995	0.018

Performance of the best RCM models and Rainfall

The best performance analysis of several RCM models based on ranking and graph analysis using Taylor Diagrams has been conducted. Research shows that each model has its performance. Table 6 was the ranking of every single model from three statistical parameters for daily, and monthly rainfall mean. Based on Table 6, it was known that the ACCESS1-0 model was the best model in simulating daily and monthly rainfall at the study site, while the worst model was CSIRO-Mk3-6-0 and MPI-ESM-MR. In addition to the three models, the ranking sequence varies in simulating daily and monthly rainfall. However, some models experienced a drastic decline in performance when simulating monthly rainfall, namely GFDL-ESM2M, which ranked 2 and 11. Furthermore, figures 4a and 4b clarifies the

best model based on its proximity. They were Taylor Diagrams used to indicate the closeness of the model with observations for daily and monthly rainfall. It was seen that the MPI-ESM-MR (K) and CSIRO-Mk3-6-0 (G) are the farthest models used to simulate daily and monthly rainfall. Besides these, the CCSM4 (E) was also far from observation for daily rainfall simulations, while the GFDL-ESM2M was far from observation for monthly rainfall simulations. Therefore, the four models were not good at simulating rainfall.

Table 6. Ranking and scores for every single model in simulating rainfall based on three statistical parameters

Ranking	Daily Rainfall		Monthly Rainfall	
	Model	Skor	Model	Skor
1	ACCESS 1-0	3	ACCESS 1-0	3
2	GFDL-ESM2M	6	CCSM4	9
3	MRI-CGCM3	9	MRI-CGCM3	9
4	inmcm4	13	bcc-csm1-1	14
5	bcc-csm1-1	14	inmcm4	14
6	CESM1-BGC	18	NorESM1-M	16
7	NorESM1-M	21	CanESM2	20
8	CanESM2	24	CESM1-BGC	26
9	BNU-ESM	27	BNU-ESM	27
10	MIROC-ESM-CHEM	30	MIROC-ESM-CHEM	27
11	CCSM4	33	GFDL-ESM2M	33
12	CSIRO-Mk3-6-0	36	CSIRO-Mk3-6-0	36
13	MPI-ESM-MR	39	MPI-ESM-MR	39

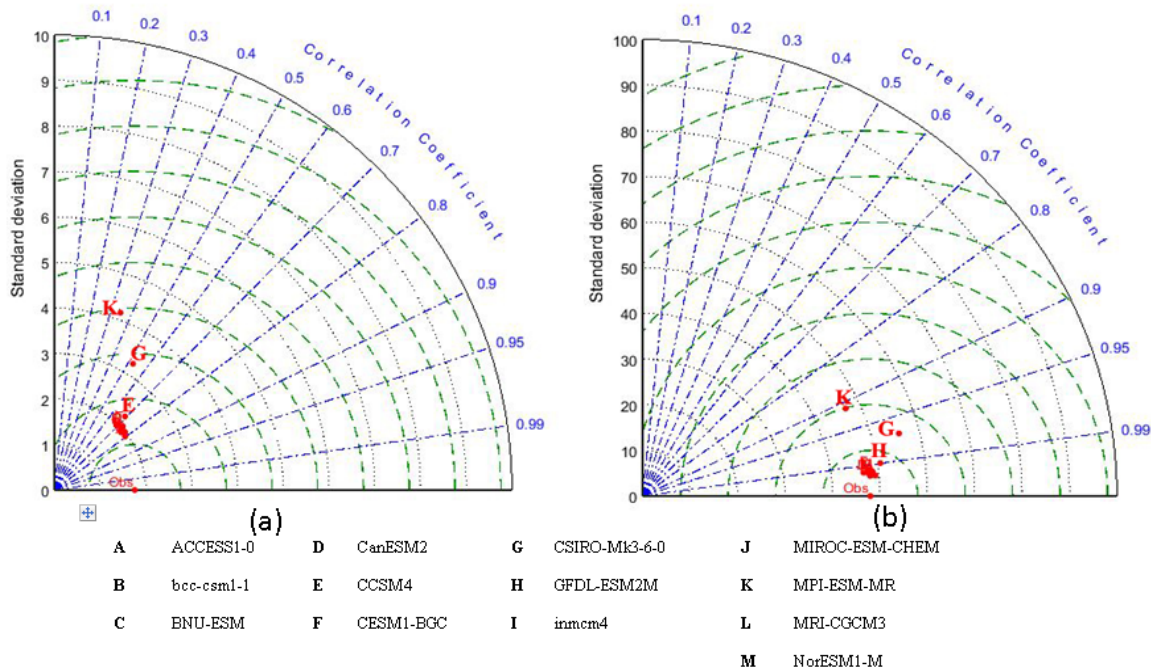


Figure 4. Taylor diagram; (a) daily rainfall, (b) monthly rainfall.

Air temperature

The best daily and monthly model for simulating air temperature from the thirteen shown in Table 7 is the MIROC-ESM-CHEM and ACCESS1-0. While the model that was not the best in simulating daily and monthly air temperatures were CCSM4 and bcc-csm1-1 respectively. To further clarify and determine the best and worst model in simulating daily and monthly air temperatures based on proximity to observation, Taylor Diagram in Figures 5 is very helpful visually. It was seen that the position of CCSM4 (E), bcc-csm4 (B) and CSIRO-Mk3-6-0 (G) models were the furthestmost from observation position (in simulating daily air temperatures). While the CSIRO-Mk3-6-0 (G) and bcc-csm4 (B) models were the furthestmost from observation position (in simulating monthly air temperatures). Therefore, the three models (CCSM4, bcc-csm4, CSIRO-Mk3-6-0) were not good at simulating air temperatures compared to others, as all of them were close to observation.

Table 7. Ranking and scores for every single model in simulating temperature based on three statistical parameters

Ranking	Daily Temperature		Monthly Temperature	
	Model	Skor	Model	Skor
1	MIROC-ESM-CHEM	3	ACCESS 1-0	6
2	CanESM2	6	CanESM2	7
3	BNU-ESM	9	MIROC-ESM-CHEM	8
4	CSIRO-Mk3-6-0	12	BNU-ESM	9
5	MRI-CGCM3	15	CCSM4	18
6	CESM1-BGC	18	inmcm4	18
7	MPI-ESM-MR	22	MRI-CGCM3	19
8	ACCESS 1-0	23	CESM1-BGC	23
9	NorESM1-M	27	NorESM1-M	29
10	inmcm4	30	MPI-ESM-MR	30
11	GFDL-ESM2M	33	GFDL-ESM2M	31
12	bcc-csm1-1	36	CSIRO-Mk3-6-0	36
13	CCSM4	39	bcc-csm1-1	39

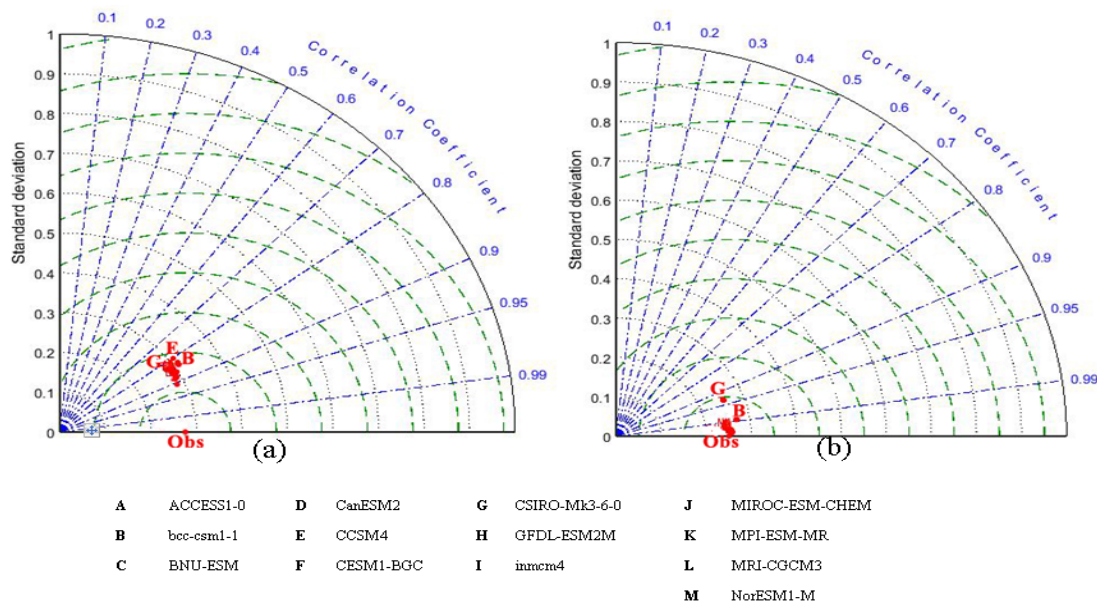


Figure 5. Taylor diagram of daily air temperature (a) and monthly air temperature (b)

Discussion

Comparing RCM performance before and after bias correction.

Every RCM have different characteristics because they are from different types of downscaling or GCM sources. Figure 2a shows the existence of 10 models whose patterns were similar to observation with only three models comprising of different patterns, namely CSIRO-Mk3-6-0, GFDL-ESM2M, and MPI-ESM-MR. These models came from CORDEX while the other was from NEX-GDDP. Therefore, different RCM data sources show different output of characters. CORDEX in producing RCM uses a dynamic downscaling technique while NEX-GDDP uses statistics downscaling. These results indicate that the downscaling statistics were better at simulating rainfall compared to dynamic downscaling, which is consistent with researches conducted by Attique (2018) and Roux *et al.*, (2018). Attique stated that statistics downscaling produces satisfactory simulation results in projecting monthly rainfall cycles while dynamic downscaling shows simulation of rainfall, which has a high bias for both historical and future periods. In line with Attique, Roux stated that statistics downscaling produced better simulations in many cases compared to dynamic downscaling. The three models above also show that the model was better at simulating during the dry season compared to the rainy season. During the rainy season, the graph was far above the observation, meaning that during this period, the resulting error was greater. This result was consistent with Aldrian *et al.* (2004) in simulating rainfall throughout Indonesia using the MPI model. He showed that the best simulation results occur in the dry season.

Based on Figure 2 and Table 3, the daily and monthly rainfall mean patterns between observations and models show that the bias correction increases the similarity pattern of rainfall models with observation. Furthermore, it was clear that the bias correction improves the performance of the model in simulating the value of rainfall. The performance improvement of the model was indicated by the increasing value of R^2 , NSE, and the smaller value of RMSE. When the values of R^2 and NSE are closer to 1, it indicates that the model was getting closer to observation. However, there were still some models that differ slightly in pattern with observations even though biased corrections have been made such as CSIRO-Mk3-6-0, MPI-ESM-MR, CCSM4 (for the daily rainfall mean pattern) and CSIRO-Mk3-6-0, MPI-ESM-MR (for monthly rainfall mean pattern). Beside pattern differences, the NSE values of CSIRO-Mk3-6-0 and MPI-ESM-MR for daily rainfall simulations still show negative values, even though the bias correction has been carried out. Since the beginning, the performance of both models before and after bias correction was different from observation. Therefore, the model with a different observation pattern will be difficult to improve even though bias correction has been conducted.

The bias correction in this study uses the distribution *mapping method* with advantages in considering the type of rainfall distribution to have a good performance in simulating extreme rainfall compared to other methods. Table 4 shows that by using the *mapping method*, the model was able to perform very well in describing extreme rainfall in the study area. The NSE values of nine models were above 0.75, which means that the model's ability was very good at simulating extreme rainfall as observed. Only four models of NSE were below 0, it means that the four models were unacceptable in simulating extreme rainfall from the beginning. Also, the results of this method equation are applied to projected rainfall models.

Figure 3 shows that patterns of air temperature similar to rainfall. Different RCM data sources show different output characters. The data source of the RCM model has an overestimate value from the NEX-GDDP while the underestimate comes from CORDEX. Table 5 shows a very high correlation between air temperature observation and the model, with NSE indicating a very good increase. The NSE values of all models before the correction are negative, both daily and monthly. A negative value means that the model cannot be accepted. However, after the bias correction was made, the values of all models increase to 0.606 - 0.807 (for daily) and 0.966 - 0.998 (for monthly). Furthermore, the error value after the bias correction also decreases significantly from 0.7 - 2.082 became 0.121- 0.173 (for daily) and from 0.425 - 2.158 became 0.01 - 0.046 (for monthly). Statistical values from bias correction of air temperature indicate that all models are very good in simulating air temperatures, both daily and monthly. This was different from the model when simulating rainfall was several models cannot be accepted as input for the impact model, while for the air temperature all are accepted. This was because the mechanism and representation of temperature in climate models were simpler than rainfall, which was more complex. It was very dependent on the specific location conditions and the rain formation

scheme. Besides, the variability of rainfall was very high, both in the space and time scale, therefore, its simulation was more complicated than the air temperature.

Performance of the best RCM models

In terms of climate change impact studies, using ensemble models, it was believed that an ensemble model provides a good simulation compared to just using a single model. The use of a single model causes overconfident, while the ensemble model improves simulation performance in simulating climate data. Therefore, the ensemble models are more powerful than single (Bhattacharjee and Zitchik, 2015; Pierce in Her *et al.*, 2016; Miao *et al.*, 2014, Duan and Phillips, 2010). However, assuming too many models were used it causes inefficient and inaccurate simulations of climate data. Therefore, the best performance analysis of the RCM model can be used as a reference in choosing some optimal ensemble models while maintaining important ensemble characteristics and minimizing bias to obtain an accurate mean ensemble (Herger *et al.*, 2018).

Based on Table 6 and Figure 4, the resulting pattern of Taylor Diagram was the same as that from the ranking and model performance in simulating the extreme rainfall discussed above, namely MPI-ESM-MR, CSIRO-Mk3-6-0, CCSM4, and GFDL-ESM2M models that were not good at simulating rainfall in the study area. Based on the description above, the best model in simulating rainfall both daily and monthly was the ACCESS1-0. It was based on ranking always occupies the top rank. These results indicate the similarity of research conducted (Perez *et al.*, 2014). Perez tested the performance of 42 CMIP5 models in terms of simulating historical climates, interannual variations, and consistency of projections for the future. Of the 42 models, ACCESS1-0 was the best model in simulating these three things. However, the ACCESS1-0 model in simulating extreme rainfall was not good compared to MRI-CGCM3. The MRI-CGCM3 model was the best model capable of simulating extreme rainfall and then followed by the ACCESS1-0 model.

When compared with the rainfall ranking, the best and the worst ranking for this air temperature varies depending on the day or month. This was possible because the variability in air temperature was very small compared to rainfall. With this small difference, it is sensitive in the ranking, although the lowest model in Figure 5 was far, however, all models were accepted as input because based on the three statistical parameters as stated, they were reliable and accurate in simulating air temperature.

Conclusion

Bias correction increases model performance as indicated by the increase in NSE and R^2 values as well as the smaller error value. However, there were still some models that performed poorly even though bias correction has been conducted such as CSIRO-Mk3-6-0 and MPI-ESM-MR. Furthermore, by using the ranking method and from the visual analysis by Taylor diagram, it was known that the best model in simulating daily and monthly rainfall was ACCESS1-0. Furthermore, for simulating air temperature, the best models were MIROC-ESM-CHEM (daily air temperature) and ACCESS1-0 (monthly air temperature).

Acknowledgment

This study was part of the Priority Research Program of the Indonesian Institute of Sciences (LIPI), Sub-Program IV, entitled "Evaluation and Projection of Climate Change Impact on Flood Risk with High Precision for the Formulation of Flood Disaster Mitigation Strategy" (PI: Dr. Apip, M.Eng), funded by LIPI through the fiscal year of 2015-2017. Also, the author also thanks wishes to express his profound gratitude to the High-Performance Computer (HPC) Laboratory, Research Center for Informatics LIPI, www.grid.lipi.go.id, for all the facilities provided.

References

- Ahmed, K.F., Wang, G., Silander, J, Wilson, A.M., Allen, J.M., Horton, R., Anyah, R. 2012. Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in The U.S. Northeast. *Global and Planetary Change*, 100: 320-322.
- Aldrian, E., Gates, L.D., Jacob, D., Podzun, R., Gunawan, D. 2004. Long-term simulation of Indonesian rainfall with the MPI regional model. *Climate Dynamics*, 22: 795-814.

- Attique R. 2018. Comparison between statistical and dynamical downscaling of rainfall under Representative Concentration Pathways scenarios over the Gwadar- Ormara basin, Pakistan. Thesis. Faculty of Geo-Information Science and Earth Observation of the University of Twente.
- Bhattacharjee, P.S., Zaitchik, B.F. 2015. Perspectives on CMIP5 model performance in the Nile River headwaters regions. *International Journal of Climatology*, 35: 4262–4275.
- Dasanto, B.D., Boer, R., Pramudya, B., Suharnoto, Y. 2014. Evaluasi curah hujan TRMM menggunakan pendekatan koreksi bias statistik. *Jurnal Tanah dan Iklim*, 38(1): 15-24.
- Deidda, R., Marrocu, M., Garoletti, G., Pusceddu, G., Langousis, A., Lucarini, V., Puliga, M., Speranza, A. 2013. Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. *Hydrology and Earth System Sciences*, 17: 5041–5059 DOI:10.5194/hess-17-5041-2013.
- Duan, Q.Y., Phillips, T.J. 2010. Bayesian Estimation of Local Signal and Noise in Multimodel Simulations of Climate Change. *Journal of Geophysical Research*, 15(D18): 1-15.
- Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and M. Rummukainen. 2013. Evaluation of climate models. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Her, Y., Yoo, S.H., Seong, C., Jeong, J., Cho, J., Hwang, S. 2016. Comparison of uncertainty in multiparameter and multimodel ensemble hydrologic analysis of climate change. *Hydrology and Earth System Sciences Discussion*, DOI:10.5194/hess-2016-160
- Herger, N., Abramowitz, G.A., Knutti, R., Angelil, O., Lehmann, K., Sanderson, B.M. 2018. Selecting a climate model subset to optimise key ensemble properties. *Earth System Dynamics*, 9: 135–151, DOI:10.5194/esd-9-135-2018
- IPCC. 2007. Climate models and their evaluation, In: *Climate Change 2007: The Physical Science Basis*. available at: <http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter8.pdf> (accessed 14 August 2018)
- Lenderink, G., Buishand, A., Deussen, W.V. 2007. Estimate of future discharges of the river Rhine using two scenarios methodologies: direct versus delta approach. *Hydrology and Earth System Sciences*, 11 (3):1145-1159.
- Meher, J.K., Das, L., Akhter, J., Benestad, R.E., Mezghani, A. 2017. Performance of CMIP3 and CMIP5 GCMs to Simulate Observed Rainfall Characteristics over Western Himalayan Region. *Journal of Climate*, 30: 7777–7799 DOI :10.1175/JCLI-D-16-0774.1.
- Miao, C., Duan, Q., Sun, Q., Huang, Y., Kong, D., Yang, T., Ye, A., Di, Z., Gong, W. 2014. Assessment of CMIP5 climate models and projected temperature changes over northern Eurasia. *Environmental Research Letters*, 9: 1-12.
- Misnawati, Boer, R., June, T., Faqih, A. 2018. Perbandingan metodologi koreksi bias data curah hujan CHIRPS. *Limnotek Perairan Darat Tropis di Indonesia*, 25(1): 18-29.
- Moriasi, D.N., Arnold, J.G., Liew, M.W.V., Bingner, E.L., Harmel, R.D., Veith, T.L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, 50(3): 885–900.
- Mukaka, M.M. 2012. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3): 69-71.
- Murdock, T.Q. and Spittlehouse, D.L. 2011. Selecting and using climate change scenarios for British Columbia. *Pacific Climate Impacts Consortium*, University of Victoria, Victoria, BC, 39 pp.
- Olsson, J., Berg, P., Kawamura, A. 2017. Impact of RCM spatial resolution on the reproduction of local, subdaily precipitation. *Journal of Hydrometeorology*, 16: 534-547 DOI: 10.1175/JHM-D-14-0007.1

- Perez, J., Menendez, M., Mendez, F.J., Losada, I.J. 2014. Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region. *Climate Dynamics*, 43: 2663-2680. DOI 10.1007/s00382-014-2078-8
- Phogat, V., Skewes, M.A., Cox, J.W., Simunek, J. 2016. Statistical assessment of a numerical model simulating agro hydrochemical processes in soil under drip fertigated mandarin tree. *Irrigation and Drainage Systems Engineering*, 5(1): 1-9. DOI:10.4172/2168-9768.1000155
- Piani, C., Haerter, J.O., Coppola, E. 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, 99(1-2): 187-192.
- Roux, R.L., Katurji, M., Reza, P.Z., Quenol, H., Sturman, A. 2018. Comparison of statistical and dynamical downscaling results from the WRF model. *Environmental Modelling and Software*, 100: 67-73.
- Supari. 2012. Spatiotemporal Characteristics of Extreme Rainfall Events Over Java Island, Indonesia (Case: East Java Province). Thesis. The Graduate School Gadjah Mada University Faculty Of Geo-Information And Earth Observation University Of Twente.
- Tang, J., Niu, X., Wang, S., Gao, H., Wang, X., Wu, J. 2016. Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. *Journal of Geophysical Research: Atmosphere*, 121: 2110–2129. DOI:10.1002/2015JD023977
- Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, 106: 7183–7192.
- Thiemeßl, M.J., Gobiet, A., Leuprecht, A. 2010. Empirical-Statistical Downscaling and Error Correction of Daily Precipitation from Regional Climate Models. *International Journal of Climatology*, DOI: 10.1002/joc.2168
- Waseem, M., Man, N., Usman, M. 2017. A Review of Criteria of Fit for Hydrological Models. *International Research Journal of Engineering and Technology (IRJET)*, 4(11): 1765-1772.
- Zhong, X., Dutta, U. 2015. Engaging nash-sutcliffe efficiency and model efficiency factor indicators in selecting and validating effective light rail system operation and maintenance cost models. *Journal of Traffic and Transportation Engineering*, 3: 255-265.