Bias correction and statistical downscaling of earth system models using quantile delta mapping (QDM) and bias correction constructed analogues with quantile mapping reordering (BCCAQ)

F Fauzi, H Kuswanto, and R M Atok

Department of Statistics, Faculty of Mathematics, Computing and Data Science, Institut Teknologi Sepuluh Nopember (ITS), Indonesia
E-mail: heri_k@statistika.its.ac.id

Abstract. Earth System Models (ESM) is a model that can simulate, predict climate change that occurred in the past, present, and create climate change scenarios in the future. ESM output has not been able to represent local scale climate. Statistical Downscaling (SD) is a static downscaling process in which data on a large-scale grids in a certain period and time period is used as a basis for determining data on small-scale grids. SD results still have a sizeable bias, so we need a method that works to reduce the bias. The bias correction method used in this research is Quantile Delta Mapping (QDM) and bias correction constructed analogues with quantile mapping reordering (BCCAQ). This study downscale rainfall and maximum temperature data generated from Beijing Normal University Earth System Model (BNU-ESM) and ERA-Interim as the proxy of the observation. The skill is verified by means of Taylor Diagram showing the correlation value, the Root Mean Square Error (RMSE), and standard deviation. Based Taylor Diagram the QDM has better performance compared to the BCCAQ method. The performance of downscaling and bias correction during September-October-November (SON) is the best compared to other seasons.

1. Introduction

Earth System Models (ESM) are models that can simulate climate and predict future climate change. The ESM generates output on global scale with low resolution does not represent local climate characteristic. Downscaling technique is a technique that is able to change the grid with large scale units into grids with smaller scale units [1]. There are two types of dynamic downscaling techniques and downscaling statistics. Statistical downscaling techniques provide an efficient and effective computational way to produce sensible hydroclimatoloty from large Global Climate Model (GCM) ensembles [2]. The simplest and easiest method of downscaling is the interpolation method [3]. In recent years, several methods have been developed that combine downscaling statistics with bias correction to better represent the local climate. [4]. The bias correction and statistical downscaling technique have been applied by many researchers [5,6,7,8].

The quantile mapping correction algorithm is used to correct the systematic distribution bias in the output of rainfall from the climate model. Although effective in reducing bias, damage to the projection trends of future models is found as a result of quantile mapping [9]. Therefore Cannon et al. [9] introduced the Quantile Delta Mapping (QDM) method which explicitly maintains relative
changes in rainfall. Bias Correction / Constructed Analogues with Quantile Mapping (BCCAQ) is a hybrid downscaling method that uses the output from CA and quantile mapping at high-scale resolution. Constructed Analogues (CA) and Climate Imprint (CI) algorithms plus QDM are run independently. The BCCAQ provides a better representation than analogue constructed bias correction (BCCA), double BCCA (DBCCA), bias correction ans spatial downscaling (BCSD), BCSD using minimum / maximum temperature (BCSDX), climate imprint method (CI), and corrected bias (BCCI) [10]. BCCAQ method is applied to downscale and correct the bias of the daily simulation scale of temperature and rainfall of 10 km to 800m, the result is that BCCAQ is able to reduce the observation error [11].

This study focuses on evaluating the skills of the bias correction method and statistical downscaling QDM and BCCAQ in reducing bias in the ESM output data. The skill evaluated based on correlation values, Root Mean Square Error (RMSE), and standard deviation. Correlation aims to determine the pattern of the relationship between the results of correction bias (QDM and BCCAQ) with observation. RMSE aims to determine the closeness or similarity between the results of the bias corrected (QDM and BCCAQ) with observational data. The ESM output used in this study is Solar Radiaton Management (SRM) scenario generated from Beijing Normal University Earth System Models (BNU-ESM) with a grid resolution of 2.8˚×2.8˚. The BNU-ESM models can simulate Southeast Asia weather well. The observational data in this study are Era-Interim reanalysis dataset from the European Center for Medium-Range Weather Forecast (ECMWF) with a resolution of 0.25˚×0.25˚. The variables used in this study are rainfall and maximum temperature.

2. Literature Review

Bias Correction / Analog Construction with Quantile Mapping Reordering (BCCAQ) is a downscaling method that combines several methods including Climate Analogues (CA), Climate Imprint (CI), Quantile Delta Mapping, and Quantile Mapping (QM). The process of BCCAQ method are to combine the daily QDM output of each high-scale grid and rearrange it according to the daily CA rank[10].

2.1. Climate imprint (CI) for rainfall and temperature

Multiplication of interpolated rainfall with the average monthly rainfall getting daily rainfall,

\[ P_{\text{daily}} = P_{\text{interpolated}} \times P_{\text{monthly}} \]  \hspace{1cm} (1)

where \( P_{\text{interpolated}} \) is the bilinear interpolation output from the daily rainfall ratio of GCM (\( P_{\text{ratio}} \)). \( P_{\text{ratio}} \) is daily rainfall divided by monthly rainfall. The daily temperature is obtained from the average monthly temperature minus the interpolation temperature,

\[ T_{\text{daily}} = T_{\text{monthly}} - T_{\text{interpolated difference}} \]  \hspace{1cm} (2)

where \( T_{\text{interpolated difference}} \) is the bilinear interpolation output from the difference in ESM output temperature(\( T_{\text{difference}} \)). \( T_{\text{difference}} \) is obtained from the difference between the monthly temperature and the daily temperature.

2.2. Constructed analogues (CA)

The analogue method has been used in the predictions of seasons and seasons since 1969 by Lorenz but produces unfavorable results. This is because analogues are chosen from a database of rough atmospheric patterns [12]. While the use of analogue methods for the downscaling process itself began [13]. One of the goals of the downscaling method is to obtain statistical data that is suitable for the atmosphere's scale [14]. The equation of CA [15]:

\[ \hat{P}_{\text{downscaled}} = P_{\text{analogues}} \left[ \left( Z_{\text{analogues}} Z_{\text{analogues}}^\top \right)^{-1} Z_{\text{analogues}} \right] Z_{\text{obs}} \]  \hspace{1cm} (3)
where $\hat{P}_{\text{downscaled}}$ is constructed high-resolution analogue, $P_{\text{analogues}}$ is a series of historical high-resolution analogues, $Z_{\text{analogues}}$ is a matrix of column vectors consisting of subsets of low resolution. Vector dimension of $P_{\text{analogues}}$ and $\hat{P}_{\text{downscaled}}$ are $p_{\text{VIC}} \times 1$, where $p_{\text{VIC}}$ the number of grids in a high-resolution. The dimension of the $Z_{\text{analogues}}$ are $p \times n$. Dimension of matrix $Z'_{\text{analogues}} Z_{\text{analogues}}$ is $n \times n$.

2.3. Quantile mapping and quantile delta mapping

Equating the cumulative distribution (CDF) function between ESM outcomes ($F_{m,h}$) and observation ($F_{o,h}$) is the basic idea of quantile mapping (QM), where $x_{o,h}$ is a representation historical observation data, and $x_{m,h}$ is historical model data. Following is the correction bias formula for the model projection data,

$$\hat{x}_{m,p}(t) = F_{o,h}^{-1}\left[F_{m,h}\left[x_{m,p}(t)\right]\right]$$

(4)

where $\hat{x}_{m,p}$ is the correction bias for the ESM output projection at time $t$. If the projection value is outside the historical range, then the Wood et al. [4] parametric distribution approach or the Boé et al. [16] constant correction approach, then the equation becomes,

$$\hat{x}_{m,p}(t) = F_{o,h}^{-1}\left[F_{m,h}\left[\frac{x_{m,h}x_{m,p}(t)}{\bar{x}_{m,p}(t)}\right]\right]$$

(5)

where $\bar{x}_{m,h}$ is the mean ESM output of the historical period, while $\bar{x}_{m,p}(t)$ represents the average output of ESM in the projected period at time $t$. The basic idea of Quantile Delta Mapping (QDM) is to maintain relative changes in climate simulation models.

$$\tau_{m,p}(t) = F_{m,h}^{-1}\left[x_{m,p}(t)\right], \tau_{m,p}(t) \in [0,1]$$

(6)

The probability that no continuation is related to the value at time $t$ is represented by $\tau_{m,p}$. The relative change in ESM output at time $t$ is then given by,

$$\Delta_m(t) = \frac{F_{m,h}^{-1}\left[\tau_{m,p}(t)\right]}{F_{m,h}^{-1}\left[\tau_{m,p}(t)\right]} = \frac{x_{m,p}(t)}{\bar{x}_{m,p}(t)}$$

(7)

The inverse of CDF estimated from the value observed $x_{o,h}$ over the historical period is a quantile correction at time $t$,

$$\hat{x}_{o,m,h,p}(t) = F_{o,h}^{-1}\left[\tau_{m,p}(t)\right]$$

(8)

The multiplicative process is applied to relative changes $\Delta_m(t)$ to correct for bias in future projections,

$$\hat{x}_{m,p}(t) = \hat{x}_{o,m,h,p}(t)\Delta_m(t)$$

(9)

equation 7 and equation 8 are substituted for equation 9,

$$\hat{x}_{m,p}(t) = F_{o,h}^{-1}\left[F_{m,h}\left[x_{m,p}(t)\right]\right]\left[\frac{x_{m,p}(t)}{F_{m,h}^{-1}\left[F_{m,h}\left[x_{m,p}(t)\right]\right]}\right]$$

(10)

ADDITIVES are applied to equations 7 and 9 for temperature variables.

2.4. Evaluate skills

To evaluate skills resulting from bias correction and statistical downscaling using Taylor diagrams that show correlation coefficients, root mean square error (RMSE) and standard deviations (SD).
\[
R_{xx} = \frac{\frac{1}{N} \sum_{i=1}^{N} (\hat{X}_i - \bar{X})(X_i - \bar{X})}{SD_{\hat{X}} SD_X}
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \frac{\hat{X}_i - \bar{X}}{N} \right)^2}
\]

\[
SD_{\hat{X}} = \sqrt{\frac{\sum_{i=1}^{N} (\hat{X}_i - \bar{X})^2}{N}}
\]

\[
SD_X = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N}}
\]

Where, \( X_i \) = the ERA-Interim of \( i \), \( \hat{X}_i \) = corrected value of \( i \) (results of bias correction).

### 3. Methodology

This research uses daily rainfall data and maximum daily temperature from two sources. The first source is the historical scenario of the climate change experiment ensemble Coupled Model Intercomparison Project Phase 5 (CMIP5). Historical scenario has several models, in this study the model chosen is the Beijing Normal University Earth System Model (BNU-ESM) with a grid resolution of 2.8° × 2.8°. The BNU-ESM model began in 1950–2005. The second source is the ERA-Interim reanalysis from the European Mid-Term Weather Forecast Center (ECMWF) with a grid resolution of 0.25° × 0.25°. Era-Interim reanalysis began in 1979–2005. This research will be analyzed separately between daily rainfall and maximum daily temperature. The resolution of the BNU-ESM model will be downscaled to the ERA-Interim resolution using the CI method to obtain local-scale climate. QDM and BCCAQ methods are used to reduce bias. The bias correction method skills in this study were assessed using Taylor diagrams. The steps of analysis are described as follows.

1. Downsampling the historical scenario data using the CI method.
2. Correcting the bias of CI downscale results on ERA-Interim reanalysis data using the QDM method.
3. Applying bias correction with the QDM method.
4. Downsampling the historical scenario using the CA method.
5. Correcting the bias of CA downscale results in the ERA-Interim reanalysis data using the QM method.
6. QDM bias correction results (step 2) at each high-scale grid point are rearranged in a particular month according to the QM bias correction results rank (step 5).
7. Applying bias correction with the BCCAQ method.
8. Dividing the results of the BCCAQ, QDM, and ERA-Interim dataset into December-January-February (DJF), March-April-May (MAM), June-July-August (JJA), and September-October-November (SON).
9. Evaluating skill methods per season (DJF, MAM, JJA, and SON) and all period using correlation coefficient and RMSE in Taylor diagram.

### 4. Results and Discussion

This section will explain the results of the QDM and BCCAQ methods, and evaluate the skills of the QDM and BCCAQ methods using the Taylor diagram.
4.1. Rainfall and maximum temperatures patterns

Time series plot is used to determine whether there are seasonal patterns or recurring events in each year, month, or day of data. Figure 1 is a time series plot for monthly rainfall and the maximum monthly temperature model of the BNU-ESM and ERA-Interim from 1995-2005.

![Time series plots](image)

**Figure 1.** Time series plots (a) rainfall (mm) (b) maximum temperature (°C).

Based on Figure 1, the highest rainfall from the BNU-ESM and ERA-Interim models occurs in December to February for each year due to the rainy season period. The decrease in rainfall occurs after February and rises again in December. The highest maximum temperatures of the BNU-ESM and ERA-Interim models occur in April to August for each year. The rainfall and maximum temperature in BNU-ESM and ERA-Interim is similar from year to year in most months with different intensities. The difference in rainfall between the BNU-ESM and the ERA-Interim is very clear, the BNU-ESM rainfall is higher than the ERA-Interim, as well as at the maximum temperature. This difference clearly shows a very large bias between BNU-ESM and ERA-Interim. The bias in rainfall tends to be greater than the bias at maximum temperature. To reduce bias, we applied statistical downscaling and correction bias methods (QDM and BCCAQ).

4.2. The result of bias corrections

The bias correction results are shown in Figure 2 of the rainfall map and Figure 3 of the maximum temperature map.

![Rainfall and temperature maps](image)

**Figure 2.** (a) BNU-ESM raw data (b) ERA-Interim raw data (c) result of BCCAQ method (d) result of QDM method.
Based on Figure 2, there is a similarity between rainfall results from downscaling and bias correction with ERA-Interim. The highest rainfall was happened on the Papua Island, parts of Kalimantan Island, and parts of Sulawesi Island. Based on Figure 3, the average highest maximum temperatures happened in southern Papua Island, Bali Island, East Nusa Tenggara Island, West Nusa Tenggara Island, and Ambon Island. The highest average maximum temperature is between 27-28°C. The lowest maximum temperature is in Sumatra Island, part of Kalimantan Island, part of Sulawesi Island, Java Island, and central Papua Island.

4.3. Evaluation skill
In order to evaluate the skill, RMSE and correlation are calculated. Figure 4 is map of correlation between result of downscaling and bias correction methods (QDM and BCCAQ) for ERA-Interim.

Based on Figure 4, the highest correlation for rainfall and maximum temperature is in southern part of Indonesia. Correlation maps cannot be used to determine the best method because there are no
significant differences results and hence, therefore a Taylor diagram is needed. The closeness of the ESM output to observations is illustrated by Taylor's diagram which is verified in terms of correlation, root mean square error (RMSE), and standard deviation [17].

Figure 5. Taylor Diagram for rainfall (a) All Period (b) season.

Figure 6. Taylor Diagram for maximum temperature (a) All Period (b) season.

Based on Figures 5 and 6, the dashed yellow line shows the value of RMSE, the gray line shows the correlation value, and the dashed black line indicates the standard deviation value. The best method is the method nearby to the point marked "observed" on the x-axis. Based on Figure 5 and 6, the best methods for bias correction is QDM in DJF, MAM, JJA, SON, and all periods. The optimum result is obtained during SON, with the correlation is about 0.39, RMSE is about 1.2, and standard deviation is about 1.3, as well as for maximum temperature, with correlation coefficient is about 0.40, RMSE is about 0.46, and standard deviation is about 0.50. If apart from the season, the maximum temperature has better than rainfall, with correlation coefficient 0.5, standard deviation is about 0.5 and RMSE is about 1.4.

5. Conclusion
The bias correction method and downscaling statistics can reduce the bias between BNU-ESM and ERA-Interim. There are similarities in the pattern between the results of BCCAQ and QDM with ERA-Interim. The southern part of Indonesia is a region that has a high correlation. Taylor Diagram results show QDM has better performance than BCCAQ to correct bias. SON is a season that has better performance than other seasons. Although QDM skills are not bad, other methods might be
applied to improve downscaling and bias correction skills. Besides trying other methods, a suggestion for future research is to be able to use ERA-5 reanalysis data to improve downscaling and bias correction skills, because ERA-5 is a reconciliation of ERA-Interim reanalysis.

Acknowledgment
The authors gratefully acknowledge the financial support from the ministry of Research and Technology, and Indonesian ministry Education and Culture. Partial support from DECIMALS project is also acknowledged.

References
[1] Wigena A H 2006 Pemodelan Statistical Downscaling Dengan Regresi Projection Pursuit Untuk Peramalan Curah Hujan Bulanan (Kasus Curah Hujan Bulanan Di Indramayu) (Bogor: Institut Pertanian Bogor).
[2] Salath E P, Mote W and Wiley M W 2007 Review of Scenario Selection and Downscaling Methods for the Assessment of Climate Change Impacts on Hydrology in the United States Pacific Northwest International Journal of Climatology 27 pp 1611–1621.
[3] Meentemeyer R K and Hunter R D 2005 Climatologically Aided Mapping of Daily Precipitation and Temperature Journal of Applied Meteorology 44 pp 1501–1510.
[4] WoodAW, Leung LR, SridharVand LettenmaierDP2004Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs Climate Change 62 pp 189–216.
[5] Ahmed K Z, Wang G, Sildar S, Wilson A M, Allen J M, Horton R and Anyah R 2013 Statistical Downscaling and Bias Correction of Climate Model Outputs for Climate Change Impact Assessment in the U . S . Northeast Global and Planetary Change 100 pp 320–332.
[6] Harsa H 2017 Pendugaan Parameter Distribusi Gamma Pada Quantile Mapping Menggunakan Self Organizing Map untuk Koreksi Bias Data Curah Hujan (Bogor: Istitut Pertanian Bogor).
[7] Maraun D 2013 Bias Correction, Quantile Mapping, and Downscaling : Revisiting the Inflation Issue Journal of Climate 26 pp 2137–2143.
[8] Piani C, Haerter J O and Coppola E 2010 Statistical Bias Correction for Daily Precipitation in Regional Climate Models over Europe Theoretical and Applied Climatology 99 pp187–192.
[9] Cannon A J, Sobie S R and Murdock T Q 2015 Bias Correction of GCM Precipitation by Quantile Mapping : How Well Do Methods Preserve Changes in Quantiles and Extremes Journal of Climate 28 pp 6938–6959.
[10] Werner A T and Cannon A J 2016 Hydrologic Extremes – an Intercomparison of Multiple Gridded Statistical Downscaling Methods Hydrology and Earth System Sciences 20 pp 1483-1508.
[11] Sobie S R and Murdock T 2017 Hight Resolution Statistical Downscaling Southwestern British Columbia Journal of Applied Meteorology and Climatology 56 pp1625-1641.
[12] van den Dool H M 2003 Performance and Analysis of the Constructed Analogue Method Applied to U . S . Soil Moisture over 1981 – 2001 Journal of Geophysical 108 pp 1–16.
[13] Zorita E 1995 Stochastic Characterization of Regional Circulation Patterns for Climate Model Diagnosis and Estimation of Local Precipitation Journal of Climate 8 pp 1023–1042.
[14] Zorita E and Hans Von Stroch 1999 The Analog Method as a Simple Statistical Downscaling Technique : Comparison with More Complicated Methods Journal of Climate 12 pp 2474–2489.
[15] Hidalgo H G, Dettinger M D and Cayan D R 2008 Downscaling with Constructed Analogues : Daily Precipitation and Temperature Fields over the United States California Energy Commission, PIER Energy-Related Enviromental Research CEC-500-2007-123.
[16] Boe J, Terray L, Habets F and Martin E 2007 Statistical and Dynamical Downscaling of the Seine Basin Climate for Hydro-Meteorological Studies *International Journal of Climatology* **27** pp 1643–1655.

[17] Taylor K E 2001 Summarizing multiple aspects of model performance in a single diagram *Journal of Geophysical* **106** pp 7183-7192.