Assessing the implementation of bias correction in the climate prediction

Nurul Nadrah Aqilah Tukimat\textsuperscript{1,2}

1Faculty of Civil Engineering and Earth Resources, Universiti Malaysia Pahang, 26300 Gambang, Pahang
2Centre for Earth Resources Research Management (CERRM), Universiti Malaysia Pahang, 26300 Gambang, Pahang

E-mail: nadrah@ump.edu.my

Abstract. An issue of the climate changes nowadays becomes trigger and irregular. The increment of the greenhouse gases (GHGs) emission into the atmospheric system day by day gives huge impact to the fluctuated weather and global warming. It becomes significant to analyse the changes of climate parameters in the long term. However, the accuracy in the climate simulation is always be questioned to control the reliability of the projection results. Thus, the Linear Scaling (LS) as a bias correction method (BC) had been applied to treat the gaps between observed and simulated results. About two rainfall stations were selected in Pahang state there are Station Lubuk Paku and Station Temerloh. Statistical Downscaling Model (SDSM) used to perform the relationship between local weather and atmospheric parameters in projecting the long term rainfall trend. The result revealed the LS was successfully to reduce the error up to 3% and produced better climate simulated results.

1. Introduction
Nowadays, the global warming and climate changes issues become most significant threat to the whole countries. The rises in the global temperature creates other disasters such as drought, flood, thunderstorm, and flash flood in everywhere and anytime. These changes also to have continuous impacts on different aspects of society, including health, agriculture, water resources, and energy demand [1]. Thus, the forecasting of the future changes of the climatic events become significant to provide beneficial data input especially when involve with the planning and managing of water resources.

Statistical Downscaling Model (SDSM) has been introduced by Wilby et al. [2] to downscale the spatial and temporal resolution which provided by General Circulation Model (GCMs) into finer resolution scale suit to the local climatic information. The main purpose is to look forward the changes pattern of climate trend affected by the greenhouse gases (GHGs) emissions and the implication to the hydrological system. The SDSM generally is considered by many researchers because the model applied simpler mathematical equation that is easier to understand the relationship pattern between predictor and predictand. Besides, the model does not require high computational demand to view the simulation result because the output is presented in the finer resolutions. Thus, it could reduce the total cost of the project but at the same time has higher ability to produce better climate simulation. This model has been applied widely in Malaysia as a statistical climate model in various sectors and purposes [3–5].
However, the accuracy of the climate simulation in the SDSM depends on the predictors selections. The predictors should have better association with the local surface climate to generate good agreement during calibration and validation process. Some criteria that are essential for the predictors to be reliably simulated by GCM are readily available to achieve from GCM output, and strongly correlated with the surface variable of interest. Disagreement between model and observations can be related to an inadequate selection of the value of some parameters that are not precisely known. Besides, inaccuracy in the estimation and forecasting will create errors in the sampling measurement and processing and produce unreliable simulation results.

Previous study applied various methods in effort to achieve the closest calibration among predictor and predictand such as Multiple Correlation Matrix (M-CM), Classical Least Square (CLS), Multiple Linear Regression (MLR), Nonlinear Programming (NLP), singular value decomposition analysis, and Canonical Correlation Analysis (CCA) [6,7]. Another way to overcome the simulation gaps is with the bias correction (BC) treatment. Mahmood et al. [8] applied Linear Scaling (LS) to forecast future temperature and detected the hotter climate in the future relative to the present climate. Chen et al. [9] revealed the Quantile Mapping (QM) and Local Intensity Scaling (LOCI) produced better performances particularly in the downscaling of precipitation extremes than the traditional multiple linear regression moethods. Bennett et al. [10] proved the quantile mapping BC able to improve the spatial correlation between modelled and observed of seasonal and annual rainfall. Meanwhile Lafon et al. [11] compared the performances of 4 different BC methods there are linear, nonlinear, gamma, and empirical distribution-based quantile mapping to reduce the RCM-simulated precipitation gaps for seven catchments spread across Great Britain. The results showed that the mean and standard deviation of daily precipitation can be corrected robustly while the correction of skewness and kurtosis of daily precipitation are much more sensitive to the choice of a bias correction method and the selection of a particular calibration period. Thus, LS method was applied in this study to treat the gaps in the simulation and future projection results. It is very significant to determine the importance of the bias treatment in the climate simulation.

2. Methodology of the Study
The study have been analysed as stated in the Figure 1 below. LS method was selected to treat the gaps between observed and simulated results produced by SDSM modelling. The selection of this method because of the simplicity, accuracy, parameter considerations and previous literatures in maintaining the reliability of the results after treatment. The equation of the LS and LOCI methods as shown below:

\[
\text{LS method} = P_{raw} \frac{\mu(P_{obs,m})}{\mu(P_{raw,m})}
\]

where \(P_{raw}\) and \(P_{obs}\) are the raw and observe rainfall on the mth month and \(\mu\) is the average rainfall for i month.
The SDSM model was chosen to simulate the future rainfall at the grid box of 28X x 34Y. This model is widely used in the context of hydrological issue due to climate scenarios because it provides station scale climate information from grid resolution GCM-scale as input based on multiple regression techniques. These regressions emerged from the empirical statistical relationship between atmospheric circulation pattern (predictors) and local scale parameters (predictand). Besides, the simplicity and computationally undemanding of SDSM tools made it acceptable and recognizable in terms of climate model projection. Daily rainfall data and daily atmospheric predictors from NCEP reanalysis data calibrates and validates in the period year 1982 - 2011 to develop the downscaling relationship. Then, the GCMs data are utilized to generate future trend for interval year periods of ∆2050. The A2 scenario chosen for this study provides an upper bound on future emissions, and it is selected from an impacts-and-adaptation point of view; if it is adaptable to large climate change, it will have no problem with smaller climate change and lower end scenario, although low emissions scenario gives less information from this point of view. As stated by Tukimat and Alias [12], the A2 scenario has more potential to develop bigger impacts of global warming compared to other emission groups.

2.1. Study Area
The study have been tested at Lubuk Paku rainfall station (3527092) and Temerloh rainfall station (3424081). These stations have been located in the Pahang River Basin which are influencing the water level in the Pahang River. The rainfall distributions at this region is influenced by the wind direction and 2 monsoons pattern known as Northeast (Oct – Mar) and Southwest (Jun – Aug) monsoons. The annual rainfall is 2400mm/year with average temperature of 28°C. The highest and lowest rainfall intensity is focused on Dec and Jan, respectively. The average wind speed is 5mph with the relative humidity of 85%.

3. Results and Discussions
3.1. Calibration and Validation Processes
About 30-years length of historical data have been considered at Lubuk Paku station (year 1982 to 2011) and Temerloh station (year 1975 to 2004) to forecast the future rainfall changes in year 2040 to 2069. The SDSM model was applied to downscale the GCMs resolution suit to the local climate condition. Based on the screening process, each locations requires different atmospheric predictors to provide better predictand-predictor’s relationships. The selection of the predictors were based on the correlation value
between predictand-predictor relationships for every month. Table 1 shows the predictor selection for each stations.

Table 1. Predictors selection for each rainfall stations.

| Predictor                          | Lubuk Paku | Temerloh |
|------------------------------------|------------|----------|
| Surface airflow strength (p_f)     | /          | /        |
| 850hPa airflow strength (p8_f)     | /          | /        |
| Specific Humidity (shum)           | /          | /        |
| Zonal velocity (p_u)               | /          | /        |
| Meridional velocity (p_v)          | /          | /        |
| 850hpa Relative humidity (r850)    | /          | /        |

Figure 2 and Figure 3 express the performances of simulated values in the calibration and validation processes at Lubuk Paku and Temerloh stations. Both stations were produced good simulated results in the calibrations but not in the validation. Calibration and validation have difference processes where the calibration is the stage of the equation development from 15years historical record using required variables meanwhile the validation is the stage to validate the reliability of that equation using another 15 years length of historical records. Referring to the Lubuk Paku station, the error was focused on Nov and Dec with 5%. Even very good calibration but the error become bigger in the validated results. Meanwhile in Temerloh station, both processes produced significant errors where BC needs to improve the quality of the simulated result.

Table 2 show the percentage error of every month for both stations. The table shows the SDSM was successfully to produce good simulated results with less than 15% error as proved that the selected predictors had good association with the local stations. At Lubuk Paku station, the bigger gaps were identified start from July to Oct during validation process. Meanwhile for Temerloh station was started from Nov to Feb. The average errors were 7.6% for Lubuk Paku station and 9.9% for Temerloh station. Even the bias was not in critical however the treatment was still significant to ensure the reliability of the projected results.

Figure 2. Calibrated (1982-1996) and validated (1997-2011) results without BC treatment at Lubuk Paku station.
3.2. BC Treatment Implementation and Climate Projection Results

In this study, LS method was implemented to treat the gaps between daily observed and simulated values. Graph in the Figure 4 shows the performances of the corrected simulation values after considered the BC at both stations. In general, the LS method was successfully to treat the simulated results more closely to the observed for every month. For Lubuk Paku station, the error was successfully to reduce to 4.4%, about 3.2% was treated by BC. Meanwhile, the gaps of the simulated value at Temerloh station was also successfully to reduce to 5.3%, about 4.6% difference from untreated results. It shows that the BC treatment is significant to treat the gaps in the simulation process.

Figure 5 shows the projected climate results after considered the BC treatment at both stations. In the projection results, the correction factors values were considered to consistent for every station. Based on the results, the rainfall intensity at both locations were expected to change in the c2050s. The rainfall patterns are also expected differ from the historical trend. At Lubuk Paku, the rainfall intensity was expected to reduce in average 12%. The heavy rainfall was estimated to occur started from Nov and then declined in the rest months. Differ in Temerloh station whereby the future rainfall trend was estimated to be fluctuated through a year. The highest rainfall trend was expected occur in June during Southwest monsoon and then pick up again in Nov and Jan during Northeast monsoon. The annual rainfall in the Δ2050s is expected to rise up to 40% becomes 2212mm/year compared to the historical.
4. Conclusion
An approaching of the statistical climate modeling nowadays becomes significant with the technologies growth. The simulation results produced from this technology can be as an importance data input especially for the long term planning and management in the water resources and land development. Thus, the model accuracy have been controlled to enhance the reliability of the production results. In this case study, LS method was applied and successfully treated the large biases in the simulation. Based on the result, the error was successfully to reduce to 4.4% and 5.3% in Lubuk Paku and Temerloh stations, respectively. The consistent LS values were used to treat the projected climate result in year
Δ2050 at both locations. It proved that the BC treatment is significant to treat the gaps in the simulation process.

Acknowledgement
This research was supported by Drainage and Irrigation Department (DID) and Ministry of Higher Education (MHE) for grant vot RDU150113.

References
[1] Raneesh K Y and Thampi S G 2013 Bias correction for RCM predictions of precipitation and temperature in the Chaliyar River Basin J. Climatol. Weather Forecast. 1–7
[2] Wilby R L, Hay L E and Leavesley G H 1999 A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado J. Hydrol. 225 67–91
[3] Tukimat N N A, Harun S and Shahid S 2017 Modeling Irrigation Water Demand in a Tropical Paddy Cultivated Area in the Context of Climate Change J. Water Resour. Plan. Manag. 143 5017003
[4] Tukimat N N A and Harun S 2015 Climate change impact on rainfall and temperature in Muda irrigation area using multicorrelation matrix and downscaling method J. Water Clim. Chang. 6 647–60
[5] Hamidon N, Harun S, Malek M A, Ismail T and Alias N 2015 Prediction of paddy irrigation requirements by using statistical downscaling and cropwat models: A case study from the Kerian Irrigation Scheme in Malaysia J. Teknol. 76 281–8
[6] Tukimat N N A and Harun S 2013 Multi-correlation matrix (M-CM) for the screening complexity in the statistical downscaling model (SDSM) Int. J. Eng. Sci. Innov. Technol 2 331–42
[7] Busuioc A, Giorgi F, Bi X and Ionita M 2006 Comparison of regional climate model and statistical downscaling simulations of different winter precipitation change scenarios over Romania Theor. Appl. Climatol. 86 101–23
[8] Mahmood R, Jia S and Babel M S 2016 Potential impacts of climate change on water resources in the kunhar river basin, Pakistan Water 8 23
[9] Chen J, Brisette F P, Chaumont D and Braun M 2013 Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America Water Resour. Res. 49 4187–205
[10] Bennett J C, Grose M R, Post D A, Ling F L N, Corney S P and Bindoff N L 2011 Performance of quantile-quantile bias-correction for use in hydroclimatological projections MODSIM 2011-19th International Congress on Modelling and Simulation-Sustaining Our Future: Understanding and Living with Uncertainty pp 2668–75
[11] Lafon T, Dadson S, Buys G and Prudhomme C 2013 Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods Int. J. Climatol. 33 1367–81
[12] Tukimat N N A and Alias N A 2016 Assessment the potential of SRES scenario for Kuala Sala, Malaysia J. Mech. Civ. Eng 13 6–12