Malaysian Journal of Civil Engineering 30(1):13-22 (2018)

DOWNSCALING OF DAILY AVERAGE RAINFALL OF KOTA BHARU KELANTAN, MALAYSIA

Muhammad Noor* & Tarmizi Ismail

Department of Hydraulic and Hydrology, Faculty of Civil Engineering, Universiti Teknologi Malaysia

*Corresponding Author: *mnkakar@gmail.com,*

Abstract: Downscaling Global Circulation Model (GCM) output is important in order to understand the present climate as well as future climate changes at local scale. In this study, Random Forest (RF) was used to downscale the mean daily rainfall at Kota Bahru meteorological station located in Kelantan Malaysia. The RF model was used to downscale daily rainfall from GCM of Coupled Model Intercomparison Project Phase 5 (CMIP5), BCC-CSM1.1. The potential predictors were selected using stepwise regression at grid points located around the study area. Quantile mapping was used to remove the bias in the prediction. The results showed that the RF model was able to establish a good relation between observed and downscaled rainfall. The Quantile mapping was found to perform well to correct errors in prediction. The statistical measures of performance of downscaling and bias correction approaches show that they are able to replicate daily observed rainfall with Nash-Schutclif efficiency greater than 0.75 for all the months. It can be concluded that RF and Quantile mapping are reliable and effective methods for downscaling rainfall.

Keywords: *Statistical downscaling, rainfall, random forest, quantile mapping*

Article history:
Received 28 March 2017
Received in revised form 20 February 2018
Accepted 10 March 2018
Published online 30 April 2018

1.0 Introduction

Climate change due to global warming will modify the climate in terms of both the mean and variability of rainfall. Small changes in rainfall variability and the mean can produce relatively large changes in the probability of extreme rainfall events. As the primary impacts of climate change on human life, environments, economies and societies result from extreme events, changes in extreme weather events due to global warming have become a major concern in recent years. Therefore, reliable projection of future changes in extreme indices at the local scale is a major challenge in any climate region. It is anticipated that variability in inter-annual and inter-seasonal rainfall due to climate change will cause more hydrologic extremes on the east coast of Peninsular Malaysia and make people’s livelihoods and local infrastructure more vulnerable.

Therefore, downscaling extreme rainfall indices is very important for the creation of rational countermeasures in the context of climate change (Pour et al., 2014; Ahmed et al., 2015).

The Global Circulation Models (GCMs) are currently considered as the most important and appropriate tool for the modelling and assessing the impact of climate change (Goyal et al., 2012). The raw outputs of climate change simulation from GCMs, however, fail to provide reliable information on spatial scales (Maraun et al., 2012). Statistical downscaling is a technique widely used to overcome the spatial resolution problem of GCMs. Statistical downscaling methods are widely used for their simplicity, easiness, flexibility, quickness and providing local scale information. Several regression-based statistical methods are developed and applied such as principal component analysis, artificial neural networks, multiple linear regression, Support Vector Machine (SVM), Random Forest (RF) and canonical correlation analysis (Mahmood and Babel, 2013). GCM simulations are not directly used because of biased representation. This biasness is due to the assumptions made during the development phase of circulation models due to lack of incomplete knowledge about atmospheric phenomena. Therefore, GCM models may not simulate climate variables accurately, and there is always a difference between observed and simulated climate variables. Thus, it is necessary to use bias correction methods to remove biasness from GCM output. In this study, Quantile mapping was used for the bias correction. The model was calibrated and validated using Room Mean Square Error (RMSE) and Nash-Schutclif efficiency (NSE) statistics.

The main objective of this study was to evaluate the performance of RF algorithm for downscaling of precipitation. This study will be a baseline for modelling and assessing the impacts of climate change in the future for Kota Bahru region in Kelantan Malaysia.

2.0 Study Area and Data

45 years (1971-2015) hourly observed rainfall data was collected from Department of Irrigation and Drainage (DID), Malaysia for Kota Bahru meteorological station located in Kelantan. The 100 predictors data of Coupled Model Inter comparison Project Phase 5 (CMIP5), for GCM, BCC-CSM1.1 was was downloaded from IPCC portal. The historical daily rainfall data for 35 years from 1971 to 2005 and projected data of 95 years from 2006 to 2099 for 100 predictors was extracted from the downloaded data of the GCM BCC-CSM1.1. The study area is indicated in the map shown in Figure 1.

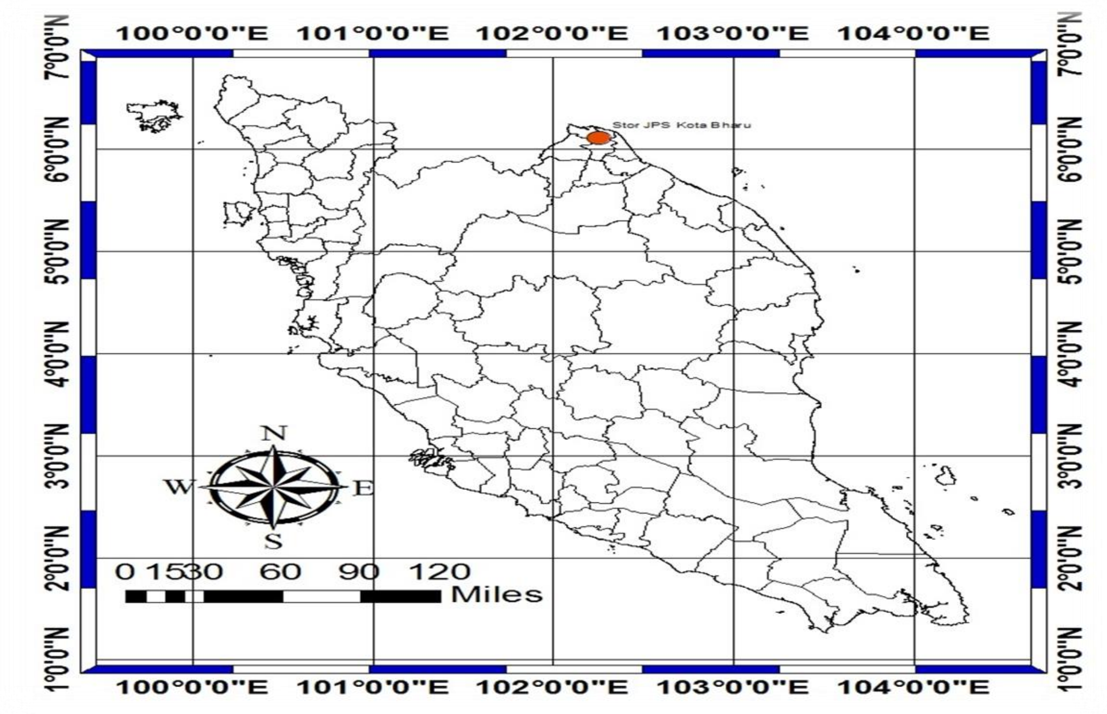

Figure 1: Location of Study Area Shown by Red Circle (Station ID 6019004)

3.0 Methodology

Historical hourly precipitation data of 45 years for peninsular Malaysia was set as observed data. Historical and projected data of Coupled Model Inter comparison Project Phase 5 (CMIP5), for Global Climate Model (GCM) BCC-CSM1.1 was downloaded from IPCC portal. Using step wise regression the daily data for 100 predictors was then extracted from the downloaded data of GCM. After calibration of data the Random Forest algorithm was used to downscale the model data. QM method was used for bias correction. The bias corrected data was then compared with the observed data. Figure 2 shows the flow chart for the study

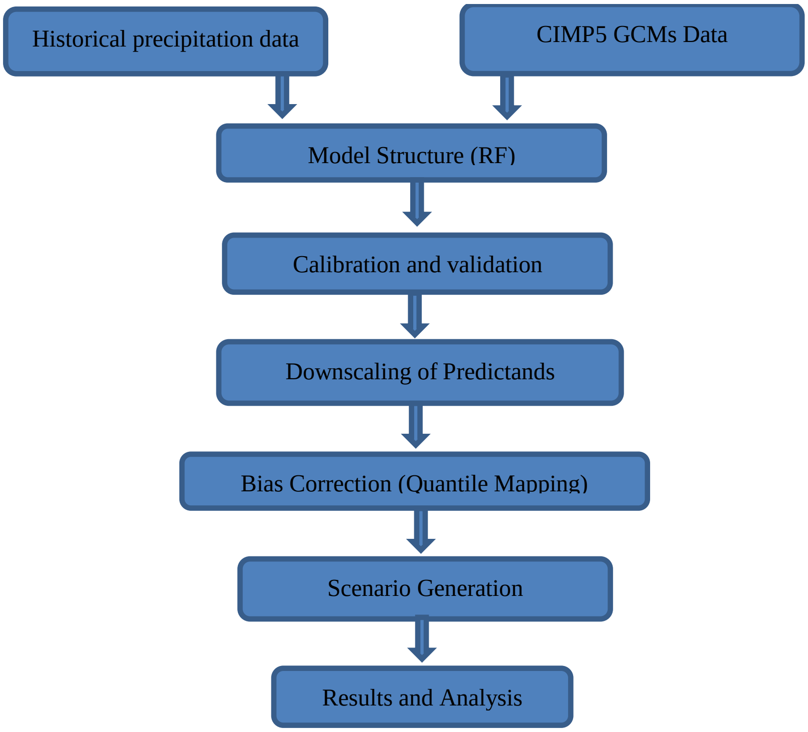

Figure 2: Methodology Flow Chart

3.1 Random Forest

Random Forest is a method used for constructing randomness in decision forests with classification or regression trees. The error is estimated after each iteration and the random forest having tree that gives least error is considered as the best model for prediction (Liaw and Wiener, 2002). The “Random Forest” package written in R program was employed in this study for developing RF downscaling model.

3.2 Bias Correction

QM is a bias correction method that relies on cumulative distribution function. The Principal of QM is to adjust the distribution of the model output with observed data set. QM was formulated by (Gudmundsson et al., 2012) as:

$$P_o = h\,(p_m) \qquad (1)$$

Where, P_o is observed precipitation, h is transfer function and P_m is downscaled precipitation.

The “qmap” package written in R program was used in this study for quantile mapping.

3.3 Calibration and Validation of Model

The model was validated and calibrated using the statistical Indices. Two different indices were used for this purpose. These are explained briefly as below.

3.3.1 Statistical Indices

The performance of precipitation data and downscaling models were assessed using Root Mean Square Error (RMSE) and Nash-Sutcliffe efficiency (NSE) statistics. The equations used for calculating these parameters are given below:

3.3.1.1 Room Mean Square Error

RMSE summaries the mean difference in the units of observed and estimate values and therefore can provide overall performance (Willmott, 1982). A model is considered more accurate if RMSE is closer to zero (Johnston, 2004; Chen and Liu, 2012).

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}\left(x_{sim,i} - x_{obs,i}\right)^2\right]^{\frac{1}{2}} \tag{2}$$

Where, $x_{sim,i}$ and $x_{obs,i}$ are the modelled and observed data.

3.4 Nash-Schutclif efficiency

The NSE proposed by (Nash and Sutcliffe, 1970) can have values between -∞ and 1.0 being the optimal values. Values between 0.0 and 1.0 are generally viewed as acceptable level of performance, whereas values < 0.0 indicates that the mean observed value is a better predictor than the simulated value, which indicates unacceptable performance of model (Moriasi et al., 2007)

$$NSE = 1 - \frac{\sum_{i=1}^{N}\left(xs_{sim,i} - x_{obs,i}\right)^2}{\sum_{i=1}^{N}\left(x_{obs,i} - \overline{x_{obs}}\right)^2} \qquad (3)$$

Where, $x_{sim,i}$ and x_{obs},i are the modelled and observed data, while $\overline{x_{obs}}$ and $\overline{x_{sim}}$ are mean observed and mean modeled data and N is the number of observations.

4.0 Results and Discussions

4.1 Downscaling

The precipitation data from Global Climate model was downscaled using RF method. The QM method was used for the bias correction of the downscaled data. The comparison of the daily observed data and monthly observed data with the bias corrected daily downscaled data and bias corrected monthly downscaled data respectively, showed a very good relationship. Figure 2(a) shows the comparison of the daily observed data and bias corrected daily downscaled daily data, while figure 2(b) shows the comparison of the monthly observed data and bias-corrected monthly downscaled data. Comparisons of the results showed that this procedure is effective for downscaling precipitation. The overall, downscaling result was found satisfactory in term of average daily and monthly precipitation.

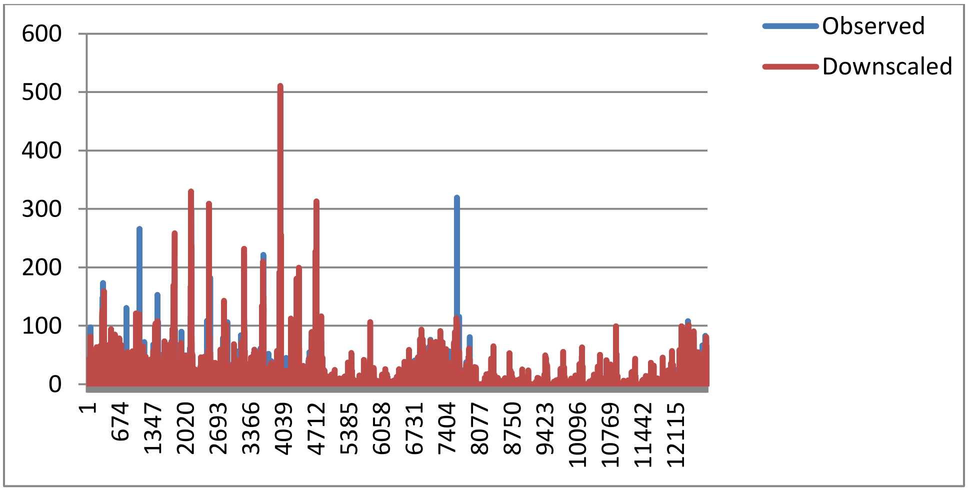

Figure 3 (a): Observed and Downscaled Daily Rainfall

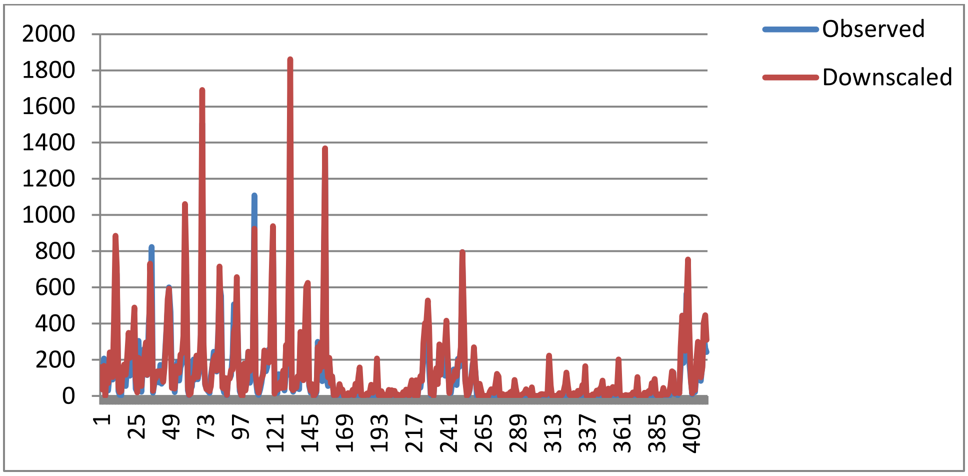

Figure 3 (b): Observed and Downscaled Monthly Rainfall

4.2 Performance of Downscaling model

The performance of the downscaling model was evaluated using two statistics indices. A model is considered more accurate if RMSE is closer to zero. NSE values between 0.0 and 1.0 are generally viewed as acceptable level of performance. Both the statistics indices performed very well. The results for both the indices are tabulated in Table 1.

Table 1: Performance of Downscaling model

Method	*V1*
RMSE	76.26
NSE	0.88

4.3 Projected precipitation

The yearly projected future rainfall shows a good relationship between four Representative Concentration Pathways (RCPs) 2.6, 4.5, 6.0, and 8.5. Figure 3 shows the projected yearly precipitation relationship amongst these RCPs.

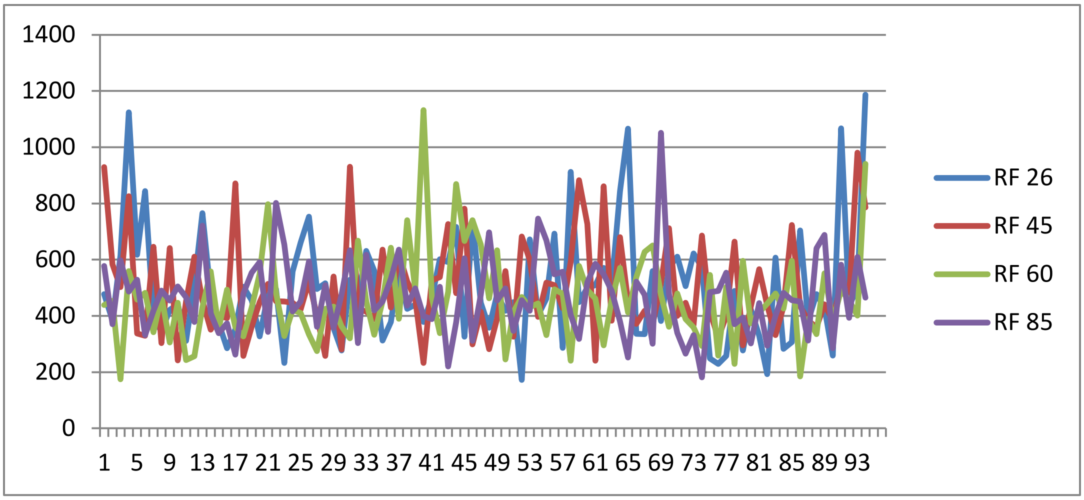

Figure 4: Yearly Projected Rainfall for Different Scenarios

Table 2 shows the average annual projected rainfall for four RCPs for three different periods of 2010-2039, 2040-20969 and 2040-2099. From the table it can be seen that for RCP 26 the average annual future projected rainfall is expected to decrease during period of 2040-2069 and increases for period of 2070-2099. For RCP 45 and RCP 60 there is an increase and for RCP 85 there is decrease in annual average projected rainfall for both the time periods 2040-2069 and 2070-2099. It was observed that due impacts of changing climate under various RCPs the intensity of average annual rainfall will be more in future. This increase in intensity due to expected extreme events can increase the susceptibility of the current urban storm water management infrastructure. Therefore, necessary measures should be taken so as to minimize the human and economic losses.

Table 2: RCP Scenarios

Scenario	*2010-2039*	*2040-2069*	*2070-2099*
RCP 26	589.22352	577.6887315	1126.616368
RCP 45	362.795541	658.0872502	618.9663744
RCP 60	395.545767	511.8591405	676.363412
RCP 85	474.0374551	415.8296497	358.3860252

5.0 Conclusion

The present study was set out to evaluate the performance of Random Forest algorithm in terms of ability to downscale and precipitation. The RF model was calibrated and validated using daily precipitation recorded at meteorological of Kota Bahru station located in Kelantan Malaysia. The result shows perfect match between downscaled and observed precipitation. Errors in downscaled daily and monthly average values were very less for all the months and days of the year. This indicates that Random Forest algorithm can be used to downscale and project precipitation. It can be concluded that RF method is reliable and effective method for downscaling rainfall.

6.0 Acknowledgment

The authors would like to acknowledge Universiti Teknologi Malaysia (UTM) for financial support through GUP vote no. 10H36

References

Ahmed, K., Shahid, S. Haroon, S.B. and Wang, X.-J. (2015) *Multilayer perceptron neural network for downscaling rainfall in arid region: A case study of Baluchistan, Pakistan.* Journal of Earth System Science 124 (6), 1325-1341

Chen, F.-W. & Liu, C.-W. (2012) *Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan.* Paddy and Water Environment 10(3):209-222.

Goyal, M., Burn, D. and Ojha, C.S.P. (2012). *Evaluation of machine learning tools as a statistical downscaling tool: temperatures projections for multi-stations for Thames River Basin, Canada*. Theoretical and Applied Climatology 108, 519-534.

Gudmundsson, L., Bremnes, J., Haugen, J. & Engen-Skaugen, T. (2012). *Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations–a comparison of methods*. Hydrology and Earth System Sciences 16(9):3383-3390.

Johnston, K. (2004) ArcGIS 9: using ArcGIS geostatistical analyst. Esri Press.

Liaw, A. and Wiener, M. (2002). *Classification and regression by Random Forest*. R news 2(3):18-22.

Mahmood, R. and Babel, M. (2013). *Evaluation of SDSM developed by annual and monthly sub-models for downscaling temperature and precipitation in the Jhelum basin, Pakistan and India*. Theoretical and Applied Climatology 113(1-2), 27-44.

Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M. and Brienen, S. (2010). *Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user*. Reviews of Geophysics 48(3), RG3003.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. (2007*) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations.* Transactions of the ASABE 50(3):885-900.

Nash, J. E. & Sutcliffe, J. V. (1970*) River flow forecasting through conceptual models part I—A discussion of principles*. Journal of Hydrology 10(3):282-290.

Pour S.H., Harun, S.B. and Shahid, S. (2014) *Genetic Programming for the Downscalling of Extreme Rainfall Events on the East Coast of Peninsular Malaysia*. Atmosphere, DOI: 10.3390/atmos5040914.

Willmott, C. J. (1982). *Some comments on the evaluation of model performance*. Bulletin of the American Meteorological Society 63(11):1309-1313.