Statistical downscaling of rainfall under transitional climate in Limbang River Basin by using SDSM

T Tahir¹, A M Hashim*¹, K W Yusof¹

¹Department of Civil and Environmental Engineering, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia

*E-mail: mustafa_hashim@utp.edu.my

Abstract. Climate change is a global phenomenon that has affected hundreds of people around the globe. In transitional climatic patterns, it is essential to compute the severity of rainfall in the regions prone to hydro-meteorological disasters. Therefore, the main aim of this study is to assess the severity of rainfall under three Representative Concentration Pathways (RCPs) from Global Climate Model data of CanESM2 in Limbang River basin. Furthermore, the objective is to check the capability of Statistical Downscaling Model (SDSM) in the tropical region. The historical data of nine weather stations were used for the period of 30 years (1976 - 2005) and Global Climate Model data of CanESM2 under RCPs of RCP2.6, RCP4.5 and RCP8.5 for the period of 2071-2100. The model was calibrated for the period of 1976-1995 and validated for the period of 1996-2005. After successful calibration and validation of SDSM, the future rainfall was simulated separately for all the three scenarios of RCPs. The obtained results have shown the values of R² and RMSE for the model calibration and validation ranged between 0.58 – 0.86 and between 1.49 and 4.7, respectively for all stations. The obtained future rainfall data from 2071 – 2100 was then compared with the base period rainfall from 1976 - 2005. It was shown that under RCP2.6 scenario there will be an increase of 8.13%, while 14.7% rise in the RCP4.5 scenario during the period of 2071-2100. An abrupt increase of about 40.6% was observed under the robust scenario of RCP8.5. Therefore, it is concluded that future pattern of rainfall in Limbang River basin under all the scenarios is constantly increasing due to the climate change.

1. Introduction

Warming of the climate system is undeniable, as it is now obvious from observations which indicate an increase in global average air and ocean temperatures, pervasive melting of snow and ice, and rising global mean sea level. According to the recently published report of Intergovernmental Panel on Climate Change (IPCC) 2013, the rise in temperature from the year 1990 to 2100 is approximately 1.4°C to 5.8°C [1]. A stove atmosphere can grasp extra moisture, and universally water vapor rises to 7% for each degree centigrade of warming. How this will interpret into fluctuations in worldwide precipitation is unclear, but the overall capacity of precipitation is expected to upsurge by 1-2% by each rise in the degree of warming. There are plenty of indication that reflects, the areas that are even now rainy are possible to become rainier. But, facts on how copious rainier and what influences there will be on a local scale are more challenging to determine [2].

The current Intergovernmental Panel on Climate Change (IPCC) Valuation Report used a numeral global climate model (GCM) to evaluate upcoming climate with different emission scenarios and concluded that it is very likely that trends in extreme precipitation will continue to increase [3]. Although GCM models are very supportive in the study and forecasts about future variations in climate, the model results are built on a larger grid scale (250 to 600 km) [4]. So, due to its coarse resolution, the
results are not fine enough to be used directly to examine the variation of hydrological impacts of local or regional scale.

The downscaling process bridges the gap between the local scale climatic variables and global climatic parameters as the resolution of GCM is very coarse in nature. In last 20 years, several downscaling techniques have been developed, which were able to reduce the mismatch between spatial and temporal local and coarse scale. Up till now, downscaling has been divided into two main categories; statistical downscaling and Dynamical downscaling. In dynamically downscaling, a higher-resolution Regional Climate Models (RCMs), having approximately resolution 5-50 Km is tied with GCMs [5]. The foremost constraint of RCMs is that it is computationally challenging as GCMs. The scenarios developed by RCMs are correspondingly complex to the choice of boundary circumstances (like soil moisture, orography, land use and so on) used to start tests. Therefore, all these factors are responsible for limiting the study by using RCMs. The analysis in this paper focuses on statistical downscaling approach because it is cheap, computationally unchallenging and willingly convenient. In addition, it is also easy to get station scale climatic data from GCM scale outputs. Statistically, downscaling techniques are much easier than dynamically, to downscale outputs from GCMs. It gives the local-scale information, and it is helpful in climate change studies [6]. On the other hand, the constraint of statistically downscaling is that it demands the past weather-related data for a longer period (i.e. up to 30 years) to obtain a suitable relation with large-scales variable.

There are several statistical models developed up till now viz Artificial Neural Networks (ANN), Support Vector Machine (SVM) and so on. statistical downscaling model (SDSM) is one of them. SDSM characterizes the combination of various regression base and stochastic based weather originator approaches [7]. SDSM is being vastly used in the studies to downscale different climatic parameters like precipitation and temperature to predict future variation in hydrological conditions [8-9]. Most of scientist and researchers used a different kind of statistical models and come to the point that SDSM is one of reasonable and reliable approach among all others. For instance, Canadian global climate model predictor was used to compare three different kinds of models and recapitulate that SDSM is a more reliable as well as robust method as compared to LARS-WG (Long Ashton Research Station Weather Generator) and mathematical approach ANN (Artificial Neural Network) [10]. In a recent study, a statistical downscaling modeling adopted to investigate the pertinence and modification in patterns of climate from 2011 to 2040 of different climatic parameters including precipitation and temperature in the Brunei Muara, Brunei Darussalam [11]. In this study, SDSM were selected by considering two distinct climatic scenario A2 and B2 of HadCM3 as well as CanESM2 scenarios of RCP2.6, RCP4.5, and RCP8.5. The results of SDSM indicated that in 2017-2046 and 2047-2076 the intensity of the extreme event will increase significantly.

Flooding is presumed as the most significant natural hazards in Malaysia. The country is privileged that it is not affected by typhoons and earthquake [12]. The increasing annual rainfall trends in Malaysia trigger more floods particularly in the states of Sabah and Sarawak. Sarawak, however, experience high rainfall (3,000-4,000 mm) compared to the Peninsula Malaysia [13]. The average annual rainfall in Limbang is approximately 4600 mm [14]. Different settlements are under the threat of frequent flooding due to intense and continuous rainfall. Limbang is one of the district in Sarawak region that has been affected by recent intense flooding events. The total area of the Limbang is approximately 3976 km². Due to being squeezed in between Brunei Darussalam at its north and coastal areas, it is pertinent to work on early flood warning. Therefore, this study is based on the two main objectives; (i) to assess the severity of rainfall by downscaling using SDSM in Limbang River basin and to check the model capabilities in the tropical region, and (ii) to forecast the climatic variability at local scale under numerous emission scenario viz, RCP2.5, RCP4.5, and RCP8.5.
1.1. Study Area

Limbang is situated in the northeast of Sarawak with a longitude of 4°45’0” N and the latitude is 115°0’0” E. It has an area of 4200 square (km²). Sarawak stretches 800 km along the Northwest coast of Borneo, covering an area of 124,450 square kilometers. Limbang has an equatorial climate with four seasons; two monsoons and two inter-monsoons. The temperature is relatively uniform throughout the year - within the range of 23 °C early in the morning to 32 °C during the day. The northeast monsoon brings heavy rainfall during the months November to February. While the southwest monsoon from June to October is usually receives less rainfall. The average rainfall per year is between 3,300 mm and 4,600 mm. Humidity is consistently high in the lowlands ranging from 80% to 90% [15]. The historical data of precipitation on daily basis was gathered from nine climate stations for the period of 30 years (1976 to 2005). The description of rainfall stations has been described in table 1 and figure 1 illustrates the geographical location.

Table 1. Location of rainfall stations in Limbang River Basin

| Station Name       | Station ID  | Latitude     | Longitude    |
|--------------------|-------------|--------------|--------------|
| Lubai Tengah       | 4649025     | 004°36’35’’   | 114°54’55’’  |
| Lubok Lalang       | 4450001     | 004°24’40’’   | 115°01’20’’  |
| Medamit Nanga      | 4449012     | 004°28’55’’   | 114°54’45’’  |
| Merbau             | 4351003     | 004°21’00’’   | 115°03’30’’  |
| Pandaruan          | 4650023     | 004°41’15’’   | 115°01’10’’  |
| Rutoh              | 4053001     | 004°05’00’’   | 115°30’40’’  |
| Setuan             | 4250001     | 004°14’50’’   | 115°04’30’’  |
| Ukong              | 4548004     | 004°33’00’’   | 114°51’15’’  |
| Ulu Medamit        | 4351002     | 004°20’30’’   | 115°11’45’’  |

Figure 1. Geographical locations of rainfall stations in Limbang catchment
Several climatic models have been developed by the Canadian Centre for Climate Modelling and Analysis (CCCMA). These models are widely used to investigate the change in climatic parameters and its variability. These models can predict quantitatively, long-term variation in future climate under different emission scenarios like GHG, aerosols and so on. For this study, CanESM2 climatic variables have been used and it is acquired from the website of Canadian climate data and scenarios (www.cccds-dscc.ec.gc.ca). CanESM2 have historical data of 26 climatic parameters for the period of 1976-2005. Moreover, it provides the future climatic data under three different scenarios of Representative Concentrations Pathway (RCP) from 2006-2100. These RCPs are RCP2.6, RCP4.5 and RCP8.5. All these three scenarios are defined based on a future projection of greenhouse gas emissions. The RCP2.6 based on the very low emission of GHG concentration. The RCP4.5 considered as a steady scenario, in which entire radiative forces keep in a stable state before 2100 by implementing numerous technologies and strategies for declining the GHG emission. The RCP8.5 scenario characterized by constantly rising in the emission of GHG with the passage of time. The data is accessible at the grid resolution of 2.8125° latitude × 2.8125° longitude.

2. Methodology

SDSM is a tool used to downscale the rainfall, from GCMs to the regional level. It comprises of stochastic approach as well as multiple linear regression (MLR). The foremost step in SDSM is to develop a quantitative connection between predicted (observed) and predictor. SDSM is further divided into three sub-models i.e. annual, seasonal and monthly sub-model. All these models drive a regression equation. But the difference is that annually sub model derive an only single equation for a whole year, whereas seasonal, derives the equation separately for each season. On the other hand, monthly sub model derives the equation for each single month separately. Furthermore, the sub-models may be conditional and/or unconditional based on the type of parameter to be downscaled. For instant, the conditional model is ideal for downscaling of rainfall and unconditionally is appropriate for downscaling of temperature.

2.1. Screening of Predictors

Screening of predictor is a most pertinent step in SDSM [16]. In this process, the most relevant atmospheric parameters are chosen with the help of MLR model, based on P-value, histograms, scatter plots, correlation matrix, and partial correlation. In the present study, a correlation matrix was preferred between predictands and CanESM2 predictors. The parameter of a predictor which is highly correlated, have best scatter plot and the minimum P – value was carefully chosen for rainfall. The selected parameters for all the stations are precipitation (prcp), surface specific humidity (Shum) and wind velocity at 500 hPa. The predictor prcp is the dominant predictor in all the station so it may be said that prcp is the super predictor for this area.

2.2. Calibration and validation of SDSM

The data used for calibration of rainfall in this study was for 20 years (1976-1995) and the period from 1996 to 2005 is being used to validate the model. The model developed for daily downscaling for rainfall by using SDSM. Monthly, a regression equation was applied in all eight stations and the conditional submodel is used. In SDSM there are two methods to optimize the model and that is Ordinary Least Square (OLS) and Dual Simplex (DS). OLS is being used because it is faster than DS [17]. The root mean square error (RMSE) and coefficients of correlation (R²) was used to check the performance of historical and simulated data of model during calibration as well as validation period. The model which was calibrated for the period from 1976-1995 is used as the base and simulate the daily rainfall for the period of 1996-2005 with the help of NCEP and CanESM2 predictors. In order to get good results, 100 ensembles were used to derive the average which was then used for the validation from 1996-2005 with observed data. During calibration, the model was found to overestimate the rainfall at all stations and the details are shown in figure 2. Also, this overestimation is resolved with the help of biased correction. This method was widely used in many studies to de-biased the overestimated rainfall [18]. The description of R² and RMSE of rainfall data listed in table 2. These variations were due to the heterogeneous characteristics of rainfall [19].
Table 2. Results of $R^2$ and RMSE for model validation

| Station name | $R^2$ | RMSE  |
|--------------|-------|-------|
| Lubai        | 0.731 | 1.499 |
| Lubok        | 0.773 | 1.475 |
| Medamit      | 0.670 | 1.626 |
| Merabu       | 0.862 | 2.683 |
| Panduran     | 0.586 | 4.450 |
| Rutho        | 0.812 | 1.754 |
| Setuan       | 0.691 | 2.460 |
| Ukong        | 0.742 | 4.734 |
| Ulu          | 0.843 | 2.587 |

3. Results and discussions

Our results were based on three different scenarios of CanESM2, i.e. RCP2.6, RCP4.5 and RCP8.5 which modeled in the SDSM to find out the future rainfall for the time span of 2071-2100 under various carbon emissions. The period from 1976-2005 is selected as a base period because a number of researchers have been taken this period and believed that this is sufficient to assess the transition in climate [20]. So, the prediction of future rainfall is depending on the comparison of these two-time spans i.e 1976-2005 and 2071-2100. The results show that the overall amount of rainfall will be increase in this region of Limbang for the period of 2071 – 2100 as compared to the base period due to the transition in climate. While there is a variation in seasonal rainfall, under CanEsm2 scenarios of RCP2.6, RCP4.5, and RCP8.5. It has been predicted that the variation in rainfall is slightly different for two scenarios i.e. RCP2.6 and RCP4.5 as compared to the worst scenario of RCP8.5 where an abrupt rise in rainfall is expected. The approximate predicted change up to the 21st century will be 8.13%, 14.7% and 40.6% under three scenarios of RCP2.6, RCP4.5, and RCP8.5 respectively in Limbang region. Sarawak can be divided into four seasons according to the occurrence of rainfall. The season which receives more rainfall among all others is from December – February (DJF) because of the Northwest monsoon. On the other hand, the season which receives the lowest rainfall among others is from June – August (JJA). Rest of
the seasons March-May (MAM) and September – November (SON) are said to be in between the monsoon and they receive less rainfall as compared to DJF. The future rainfall for the season DJF is expected to increase under all RCPs, scenario except Lubai, Medamit, Panduran, Rutho, Setuan and Ukong. The RCP 4.5 scenario is responsible to decline the future rainfall for the stations Lubai, Panduran and Rutho. While in Medamit, Panduran, and Rutho RCP 2.6 scenario is responsible to reduce the rainfall. The scenario RCP 8.5 reduce the rainfall in Lubai, Rutho and Ukong during the DJF season. Contrary to this during the MAM season the rainfall is expected to increase under every scenario of GHG emission except Panduran station. In this station, the rainfall is decreased under the RCP 2.6 and RCP 4.5 scenario. While in JJA season there will be a large flocculation in rainfall pattern under all the RCPs scenarios especially in Lubok and Merabu where all the scenarios are responsible to lessen the future rainfall. Moreover, Panduran and Ukong projected to receive less rainfall under the RCP 2.6 scenario, and the Panduran also receive less rainfall under the RCP 4.5 scenario. In comparison to other seasons, the SON season subjected to increase the rainfall in future under all the scenarios except Lubok and Panduran. The Lubok station projected to experience a decrease in a future rainfall under a transitional change in all the RCPs scenarios. While in Panduran the RCP 2.6 and the RCP 4.5 scenarios are responsible to decrease the future rainfall. The detailed information of predicted rainfall change in all the stations is illustrated in figure 3.

Figure 3. Difference in observed and RCP2.6, RCP4.5, and RCP8.5 rainfall for the period of 2071-2100. The x-axis represent seasons and Y-axis reflects the rainfall (mm)

4. Conclusion
The most extensively used downscaling tool SDSM was applied in this study to project the future transition in rainfall in the Limbang River catchment under CanESM2 scenarios of RCP2.6, RCP4.5, and RCP8.5. The downscaling of these parameters is more than crucial to investigate the impact of climate change on the catchment of the basin. In this study, the historical data of nine stations was used from the period of 1976-2005 and then a quantitative approach is applied in the SDSM to select the predictors to measure the future change. Moreover, calibration and validation are done and from the findings, it is believed that all the RCPs are used to simulate the variation of climate in future. The result of SDSM projected that rainfall will be increased under all the RCPs scenario. But the seasonal rainfall varies from one station to another, especially during the JJA season where there will be less rain in most
specific RCP2.6 and RCP8.5 scenario. Considering all these scenarios, under Coupled Model Intercomparison Project Phase 5 (CMIP5) model of CanESM2 scenario the rainfall is expected to increase from the period of 2071 – 2100 due to the change in climate by comparison to the base period of 1976 – 2005. The finding of this study could help the policymakers and the responsible authorities in the better planning of water management and drainage system under the transition climate of future.

Acknowledgments
Authors are thankful to Canadian Centre for providing the GCM data and DID Sarawak for facilitating us with rainfall data. The authors are also would like to appreciate Universiti Teknologi PETRONAS, Malaysia for providing the financial assistance in carrying out this research.

References
[1] Kaspar F, Meinshausen M and Schulz M 2013 IPPC Introduction Climate Change 2013 Phys Sci Basis Contrib Work Gr 1 to Fifth Assess Rep Intergov Panel Clim Chang pp 1-90
[2] Myhre G, Shindell D, Bréon F-M, Collins W, Fuglestvedt J and Huang J Anthropogenic and Natural Radiative Forcing. Clim Chang 2013 Phys Sci Basis Contrib Work Gr 1 to Fifth Assess Rep Intergov Panel Clim Chang pp 659–740
[3] Zhou T, Gong D, Li J and Li B 2009 Detecting and understanding the multi-decadal variability of the East Asian Summer Monsoon - Recent progress and state of affairs Meteorol Zeitschrift 18 455
[4] Gebremeskel S, Liu YB, de Smedt F, Hoffmann L and Pfister L 2005 Analysing the effect of climate changes on streamflow using statistically downscaled GCM scenarios Int J River Basin Manag. 2 271
[5] Chu JT, Xia J, Xu CY, Singh VP 2010 Statistical downscaling of daily mean temperature, pan evaporation and precipitation for climate change scenarios in Haihe River, China Theor Appl Climatol. 99 149
[6] Tavakol-Davani H, Nasseri M and Zahraie B 2013 Improved statistical downscaling of daily precipitation using SDSM platform and data-mining methods Int J Climatol. 33 2561
[7] Wilby RL, Whitehead PG, Wade AJ, Butterfield D, Davis RJ, Watts G. 2006 Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK. J Hydrol. 330 204
[8] Wilby RL, Dawson CW. 2013 The statistical downscaling model: Insights from one decade of application. Int J Climatol. 33 1707
[9] Mahmood R and Babel MS 2013 Evaluation of SDSM developed by annual and monthly sub-models for downscaling temperature and precipitation in the Jhelum basin, Pakistan and India. Theor Appl Climatol. 113 27
[10] Khan MS, Coulibaly P and Dibike Y 2006 Uncertainty analysis of statistical downscaling methods using Canadian Global Climate Model predictors Hydrological Process 20 3085
[11] Hasan DSNA PA, Ratnayake U, Shams S, Nayan ZBH and Rahman EKA 2017 Prediction of climate change in Brunei Darussalam using statistical downscaling model. Theor Appl Climatol. 2
[12] Hiew KL. Flood Mitigation and Flood Risk Management in Malaysia 1996. International Workshop of Floodplain in Risk Management p 205–16
[13] Zubaidi J 2010 Climate Change:Potential Impacts On Water Resources And Adaptation Strategies In Malaysia. Clim Chang Impacts To Water Environ Adapt Options
[14] Hussain M, Yusof KW, Mustafa MR and Afshar NR 2015 Application of statistical downscaling model (SDSM) for long term prediction of rainfall in Sarawak, Malaysia. WIT Trans Ecol Environ. 196 269
[15] The Official Portal of the Sarawak Government. [cited 2017 Mar 9].
[16] Toggweiler J, Key R. 2001 Ocean circulation: Thermohaline circulation. Encycl Atmos Sci. 4 1549
[17] Wilby RL and Dawson CW 2007 SDSM 4.2-A decision support tool for the assessment of regional climate change impacts, Version 4.2 User Manual. (Lancaster Univ
Lancaster/Environment Agency Engl Wales 1)

[18] Goyal MK, Burn DH and Ojha CSP 2012 Evaluation of machine learning tools as a statistical downscaling tool: Temperatures projections for multi-stations for Thames River Basin, Canada. *Theor. Appl. Climatol.* 108 519

[19] Wilby RL, Dawson CW and Barrow EM 2002 SDSM — a Decision Support Tool for the Assessment of Regional Climate Change Impacts *Environ Model Softw.* 17 145

[20] Hussain M and Yusof KW 2017 Evaluation of CMIP5 models for projection of future precipitation change in Bornean tropical rainforests. *Theor. Appl. Climatol.* 1