Analysis of Potential Extreme Drought using Integrated Statistical Model

N N A Tukimat¹, A S Othman¹ and S N Rahmat²

¹Faculty of Civil Engineering Technology, Universiti Malaysia Pahang. 26300 Gambang, Pahang, Malaysia
²Faculty of Civil Engineering and Environment, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Johor, Malaysia

Abstract. Frequent extreme drought especially in urban area is majorly connected with the changes of the global climate and drastic releases of greenhouse emissions in the earth system. It becomes significant in identifying how frequent the potential drought event in the long term and how big its impact to the existence water sources. Due to this concern, the integrated statistical model has been used to estimate the potential extreme drought in Pahang state, Malaysia. The Representative Concentration Pathways in three radiation levels known as RCP2.6, RCP4.5 and RCP8.5 provided by IPCC Fifth Assessment Report (AR5) were implemented to produce the plausible future weather scenarios in the different radiation levels. The results revealed the climate changes could alter the seasonal trend and intensity with small rises in average 7%/year (rainfall) and 0.2 °C/decade (temperature). Although the rainfall was expecting to increase however almost 42% of Pahang state is expected to receive lower rainfall intensity than the historical annual rainfall. Estimated the drought events potentially to occur in 20 % from upcoming 80 years with every station has high probability to drought at least twice times. For the RCPs performances, the RCP4.5 potentially to produce more frequent drought compared to other RCPs.

1. Introduction
Drought classifies as one of a disaster because it also gives a huge impact to the earth life as big as flood. Malaysian Meteorological Department (MMD) have been reported plenty of calamities events since year 1900s. At least 12 times of extreme drought have been recorded since year 1951 especially during South-West monsoon (SW). Even the dry season not something new for Malaysian people, however the occurrence of this disaster become more frequent compared to 100 years ago. Meanwhile, the strongest drought were occurred in year 1982 – 1983 and 1997 – 1998 which caused the economic losses, destroyed of paddy field and fisheries, restriction of recreation activities, air pollutant, water rationing, and tourism activities. In the year of 2015, the weather pattern in Malaysia during northeast monsoon (NE) was identically different. Most of the east coast states received low rainfall intensity and achieving below normal level. Meanwhile the temperature rises time by time up to 35°C as an early sign of the El-Nino formation in Malaysia. The high temperature encourages the evapotranspiration (ET) activities and affecting to the water utilization efficiency [1]. It is also affecting the electricity demand which one of the sources of electricity production were generated by the hydropower plant [2]. According to Energy Commission Report [3] the demand of electricity is estimated to increase 3.1% in every year. Thus, the drought event not even effects to the electricity supply but also gives an impact on the consumer. In fact, the drought in Malaysia is not categorised as
a major issue because rarely happened in the critical condition (such as El-Nino) and also not involve with high cost after impact likes flood disaster. However, the dry season nowadays is inconsistently happened and in longer period. Reported by National Oceanic Atmospheric Administration [4], the anomaly global surface temperature was fluctuated increases year by year since 1980s, after 200 years of Industrial Revolution which related to the beginning of the climate changes impact. The climate changes produce many abrupt changes on the water resources affected by the monsoon impact [5]. One of the biggest contributions of greenhouse gases comes from the human activities which contribute to the changes in the climate’s variability, pattern and extreme event formation [6]. Hence, the integrated statistical model which combining of Statistical Downscaling Model (SDSM) and Standard Precipitation Index (SPI) has been applied to investigate the potential extreme drought events in the long term. The SDSM is one of the climate models that been used to predict the changes of climate variability in responding to the expected dispersion of greenhouse gases and radiation into the atmospheric system. This model has potential to generate the reliable results while having limited sources [7-10]. It uses the Model Output Statistic known as MOS and perfect prog approach to predict the numerical weather in the short range [11]. The multiple regression equation between large-scale (predictor) and local climate (predictand) has been built to produce the time series of synthetic daily weather. There were numerous studies had been conducted to test and prove the reliability of the models especially in Malaysia [12-15]. Meanwhile the drought indices were determined using SPI measurement. It uses a statistical relationship in monitoring the deficit of rainfall in a different time scale. The rainfall distribution must be converted into normal distribution as a control for each location and period time scale. The rainfall deficit was based on the SPI value. The positive value shows the higher intensity than median (wet) while negative value shows the lower intensity than the median (dry). According to [16] and [17], an extreme drought was more presentable in spatial standardization and SPI method which showed better drought index than other methods. Therefore the integrated of SDSM-SPI had been applied in estimating the potential trend of drought severity in the long term.

2. Study Area
The focused area of this study was at Pahang State, Malaysia (Figure 1) because the weather trend in Pahang was influenced by the wind direction and the monsoons. Pahang is the largest state in Peninsular Malaysia. It was located at Eastern of Malaysia and nearest to the equator. Meanwhile Temerloh (one of the districts in Pahang state) has been recognised as a centre of Peninsular Malaysia. Tropical monsoon at Pahang state brings with the series of uniform temperature between 21 °C to 32 °C throughout the year. The dry season occurred during months of January to April meanwhile months of May to December would be the wettest. The area experienced wet and dry seasons throughout the year due to its geographical position which located in equator lines. The climatic cycle in the region are influenced by four seasonal changes, known as North-East (NE) monsoon, SW monsoon, and 2 inter-transition monsoons which influence the monthly rainfall intensity. The average annual rainfall at Pahang state is 2,540 mm with humidity of 84%.

Figure 1. Rainfall and temperature stations in Pahang state
3. Methodology

The framework of this study consists of three significant steps as shown in Figure 2. There were validated the predictand-predictor equations, generated the long-term changes of local climates using GCMs variables and analysed the probability of extreme drought events in the long term using SDSM-SPI analysis. The SDSM utilized a linear regression method in built the predictor-predictand relationship. It consists of two steps: 1) develop the predictor-predictand equation for validation process and 2) generating the daily weather with considered the level of GHGs and radiation forcing. The large-scale predictors were provided by two climate groups at the grid box of 28X x 33Y there were NCEP reanalysis (for validation process) and GCMs (for the long-term generation).

Figure 2. Framework of the study

The GCMs depends on three RCPs which provided by GCM-CanESM2 (AR5) to produce potential future time series of weather scenarios. The RCPs was based on the level of radiation forcing start from low (RCP2.6), intermediate (RCP4.5) and high (RCP8.5) emissions. The predictor-predictand equations were developed using multi-linear regression approach for generation the long term climates at the region. The rainfall \((y)\) on day \(t\) can be determined by listed below:

\[
y_t = F^{-1}[\Phi(Z_t)] \\
Z_t = \beta_0 + \sum \beta_j u_{\hat{t}} + \beta_{t-1} + \epsilon
\]

Where \(F\) is the empirical function of \(y_t\), \(\Phi\) is the normal cumulative distribution function, \(Z_t\) is the z-score on day \(t\), \(\beta\) is the regression parameter, \(u_{\hat{t}}\) is normalized predictor and \(\epsilon\) is the variable parameter. For the rainfall analysis, the equation was transformed to the fourth root to take account for the skewed nature of the rainfall distribution. Even the statistical downscaling has several limitations [18], however the SDSM model does not require high computational demand to view the simulation results but has ability to produce high quality of projection results. These advantages, as a whole, had made SDSM a reliable tool for climate downscaling and was selected as a downscaling tool to generate the future climate trend at the study site.

3.1. Drought Assessment using SDSM-SPI Analysis

The drought assessment was analysed based on the daily climate results produced by the SDSM. The classification of the drought index as shown in Table 1. The drought severity depends on the duration
of events occurred and its impact size to the society. The drought severity is measured with the summation of all SPI index during drought period as shown in Equation 1 below:

\[ S = \sum_{i=1}^{d} SPI_i \]  

(3)

Where; \( S \) = Drought severity, \( d \) = Drought duration, \( i \) = Starts with the first month of a drought and continues until the end of the drought duration. The severe magnitude of the longer drought period suggests \( i=1 \). Next, the Gumbel Extreme Value I was used to determine the probability distribution of extreme rainfall dataset by:

\[ K_T = -\frac{\sqrt{6}}{\mu} \left[ 0.5772 + \ln \left[ \ln \left( \frac{T}{\tau - 1} \right) \right] \right] \]  

(4)

where \( K_T \) = Frequency factor, and \( T \) = Return period

The K value is then used in the following equation to produce the magnitude because severity in unit mm the value need to multiply with time as shown in Equation 3.

\[ X_T = \mu + K\sigma \]  

(5)

where; \( \mu \) = Mean, \( \sigma \) = Standard deviation, and \( K \) = Frequency factor

### Table 1. SPI classification

| Classification | Drought Category |
|----------------|------------------|
| > 0            | Normal           |
| 0 to -1        | Moderate         |
| -1 to -1.5     | Severe           |
| > -1.5         | Extreme          |

### 3.2. Reliability of the Results using Statistical Analyses

There were 3 significant statistical analyses; mean absolute error (MAE), correlation coefficient (R), and Nash Sutchliffe Efficiency (NSE) have been used to evaluate the accuracy and reliability of the predictor-predictand equations as shown in Table 2 where \( X_{obs} \) refers to the ith month observed data, \( X_{est} \) is the ith month estimated data and \( n \) is the number of data. The function of MAE is to measure the accuracy of continuous variables through the average of errors between the two sets of data representing the whole disparity of two data sets. Meanwhile R and NSE were to evaluate the strength and efficiency of the predictand-predictor relationship between simulated and observed data.

### Table 2. List of statistical analyses

| Name       | Equation |
|------------|----------|
| %MAE       | \[ \frac{1}{n} \left[ \frac{\sum (X_{est} - X_{obs})}{X_{obs}} \right] \times 100 \]  
| R          | \[ \frac{n\sum X_{est} X_{obs} - (\sum X_{est})(\sum X_{obs})}{\sqrt{[n\sum X_{est}^2 - (\sum X_{est})^2][n\sum X_{obs}^2 - (\sum X_{obs})^2]}} \]  
| NSE        | \[ \frac{\sum (X_{est} - X_{obs})^2}{\sum (X_{obs} - X_{obs})^2} \]  

### 4. Results and Discussions

#### 4.1. Climates Simulation and Projection

The performances of the downscaled simulated of monthly rainfall and temperature by SDSM as shown in Figure 3. The simulated rainfall results were based on the relationship between combination of these five selected predictors; p_u (zonal velocity), p_v (meridional velocity), p5_u (zonal velocity
at 500hPa), r850 (specific humidity), and shum (specific humidity) with the local weather series. Each predictor was selected based on the highest monthly correlation value between predictand-predictor associations. Based on the correlation performance, the $p_u$ became as a domain predictor variable for all local stations. Even $r850$ and $shum$ were not produced higher correlation (in average 0.3) however the combination of these five selected predictors was successfully able to model relationship with the local stations. This finding confirms the works of [19] stated the characteristics of specific humidity, airflow indices, and zonal velocity that are commonly used as the precipitation predictors in the climate studies. Meanwhile [20] used wind speed, global radiation, and humidity as important parameters in measuring the weather trend changes. The difference in average between the simulated results with the historical data over region were in range 0.10 % to 23.7 % with high $R$ and NSE values more than 0.83 as shown in Figure 4. The simulated result during calibration process have very good agreement with the historical data but poor performance in the validation process at certain stations especially at Kechau and Jandaibak stations. Meanwhile the selected predictors for the temperature were different from the rainfall variables there were $p_z$ (vorticity), $p500$ (geopotential height at 500hPa), $r500$ (relative humidity at 500hPa), $shum_850$ (specific humidity at 850hPa) and temp (mean temperature). The temperature was classified as unconditional process which not so difficult to analyse compared to the rainfall. For the temperature, predictor of $p500$ became as a main predictor variable that gives huge impact to the formation of local temperature with the correlation value of 0.8. Figure 5 and 6 show the calibrated and validated performances for these two temperature stations. The biases between simulated and historical data were less than 1.9% with $R$ and NSE values closed to 1.0.

![Figure 3](image-url)
Figure 4. Calibrated and validated performances of 12 rainfall stations using statistical analyses

Figure 5. Comparison between simulated result with historical for calibrated (1984 to 1998) and validated (1999 to 2013) process in maximum, mean and minimum temperature (unit in Celsius)

Figure 6. Calibrated and validated performances of temperature using statistical analyses
4.2. Long Term Climate Generation

Figure 7 shows the average annual spatial distribution throughout Pahang state by three different RCPs. In general, the pattern of projected rainfalls by these RCPs were consistent throughout Pahang state with very small variance. The rainfall was expected to increase to 5.5 %, 6.1 %, and 6.7 % annually by RCP2.6, RCP4.5, and RCP8.5, respectively from year 2020 to 2099 (80 years). The wettest areas were concentrated at eastern Pahang with annual rainfall of ±2700 mm/year. It is consistent with the location of these regions which located nearest to the South China Sea and monsoon pathways. The temperature can be influenced by the land and sea breezes on the general wind flow pattern potentially to bring heavier rainfall quantity. Besides, the land topographic encourages the formation of orographic rain at these regions. All the RCPs projected agreed that the least areas to receive rainfall amount were focused on the middle part of Pahang state especially at Bentong, Bera and Pekan with minimum rainfall of 1665 mm/year, 30% lesser than the average historical annual rainfall. Although the annual rainfall intensity was predicted to rise annually however almost 42 % of Pahang state was expected to receive rainfall intensity below the historical annual rainfall which forming in two radiuses; 1) at Northern Pahang including Lipis, southern part of Cameron Highland, and northern of Raub 2) at middle of Pahang state including Bera, Maran, and small western part in Pekan. The rainfall intensity at these regions was less than 2000 mm/year only. Meanwhile the temperature was al expected to rise to 0.2 °C per decade and it is consistent to the Malaysia Meteorological Department [12] stated the projected temperature in Peninsular Malaysia is in range of 1.1 °C to 3.0 °C at the end of century with Dec-Jan-Feb (DJF) as the warmest season.

![Figure 7](image)

Figure 7. Prediction of average annual rainfall distribution for upcoming 90 years by RCP2.6, RCP4.5 and RCP8.5 (unit in mm)

4.3. Probability of the Long Term Drought Event

The potential occurrence of the drought in the future year at each region was estimated using integrated SDSM-SPI model as described by Figure 8. The SPI was analysed based on the cumulative precipitation of 12 months and were classified based on Table 1. The potential drought event was evaluated when the SPI reading lower than -1. According to the results, all RCPs claim that more than 80 % of upcoming 80 years are considered as a normal. Meanwhile another 20 % of the predicted years were drought. Although the percentage occurrence was not too extreme however most of the stations were highly potential to have drought in the future year at least twice times. As shown in Figure 8, each RCPs produced diverse percentage of probability drought events at different region. Each region has inconsistent probability to have the drought chances. Bhg Selatan station at Cameron Highland as the lowest percentage of drought chances compared to another regions. Cameron Highland known as a cool area because located at mountainous with the normal annual temperature is below than 25 °C. For the meantime Kuala Bera station was predicted to receive drought event more frequent with high probability of 2.2 %, 5.1 %, 9.3 % in extreme, severe, and moderate, respectively. About >3 % of extreme drought event was potentially to occur at several pointed stations such as
Palas, Kuantan, Kg Temai Hilir, Kg Salong, and Kg Jawi. Considering the SPI analysis, the RCP4.5 was predicted to produce the highest probability of drought event compared to others RCPs. For example, in Janda Baik station, extreme drought events were expected more frequent (35 occurrences in lasting more than 12 months) by RCP4.5 however 12 and 27 occurrences of drought events were predicted by RCP2.6 and RCP8.5, respectively. The critical years that have to be caution were predicted on year 2026, 2047, 2057, 2073, 2076, 2080 and 2089 which the SPI value drop to -2. Similar case with Kg Manchis station where the extreme drought predicted by RCP4.5 was slightly higher (47 occurrences) than estimated by the RCP2.6 (28 occurrences) and RCP8.5 (11 occurrences). It proven a fact that the higher radiation does not mean to contribute greater changes to the climates.
Figure 8. Potential annual dry length based on the SDSM-SPI analysis for year 2019 to 2099

5. Conclusion
The identification of the potential drought events in the long-term becomes significant especially for the water resources management and planning with monitor how frequently the events at each different region. The application of the integrated SDSM-SPI models was carried out the probability of extreme events with concerned the climate changes impact. The reliability of the climate projected results using SDSM model was controlled by less MAE with high R² and NSE values during calibration and validation processes. The temperature was well analyzed with less 1.9 % of MAE, R and NSE was closed to 1.0 due to unconditional process. Unlike to the rainfall analysis which categorized under conditional process. The calibrated results were successfully to produce good agreement but poor in
the validation performance. However, the statistical analyses still produced lower percentage of MAE (<23.7 %) with >0.83 of $R^2$ and NSE values. These were as a proved the selected predictors for each station has good correlations with the local climates. In the long-term climate analysis, there were 3 RCPs used which represented level of radiation in the atmosphere known as RCP2.6, RCP4.5 and RCP8.5. For the rainfall projection, all of them agreed the local rainfall was predicted to rise less than 7 % at the end of the century. The rainfall was expected to concentrate at eastern Pahang with 2700 mm/year. The hottest area was focused on the middle part of Pahang state with 1665 mm/year, 30 % dropped from the historical record. Due to non-uniform rainfall pattern, almost 42 % of Pahang state was expected to receive lower rainfall intensity. Estimated the drought events potentially to occur in 20 % from upcoming 80 years with every station has high probability to drought at least twice times. For the meantime, the RCP4.5 produced the more frequent drought events compared to other RCPs. The predicting of the long-term drought at the region can be as significant data input in optimizing the water management and minimizing the risk and cost in the future year.

6. References
[1] Z. Zou, Y. Yang, G. Y. Qiu, Quantifying the Evapotranspiration Rate and Its Cooling Effects of Urban Hedges Based on Three Temperature Model and Infrared Remote Sensing, Remote Sensing, Vol 11 (202) doi: 10.3390/rs11020202, 2019
[2] TNB. (2014). Tenaga Nasional Berhad Annual Report 2014. Retrieved from Jalan Pantai Baru, Kuala Lumpur:
[3] EC (2014). Energy Commission: Peninsular Malaysia Electricity Supply Industry Outlook 2014. Suruhanjaya Tenaga (Energy Commission). ISSN: 2289-7666
[4] NOAA, Climate at a Glance: Global Time Series, National Centres for Environmental information, http://www.ncdc.noaa.gov/cag/ 2017
[5] A. Amini, S. Zareie, P. Taheri, , K. Wan Yusof, and M.R. Mustafa, Drought Analysis and Water Resources Management Inspection in Euphrates-Tigris Basin in Book of River Basin Management, Daniel Bucur, IntechOpen, DOI: 10.5772/63148, 2016
[6] NOAA, New report finds human-caused climate change increased the severity of many extreme events in 2014. National Centres for Environmental information, http://www.noaanews.noaa.gov/ 2015
[7] N. N. A. Tukimat, S. Harun, and S. Shahid, Modeling Irrigation Water Demand in a Tropical Paddy Cultivated Area in the Context of Climate Change, Journal of Water resources Planning and Management, Vol. 143 (7) DOI: 10.1061/(ASCE)WR.1943-5452.0000753, 2017
[8] Z. Hassan, S. Shamsudin, S. Harun, M. A. Malek, N. Hamidon, Suitability of ANN applied as a Hydrological Model Coupled with Statistical Dowcsaling Model: A case study in the Northern Area of Peninsular Malaysia, Environmental Earth Science Vol 74(1) pp: 463-477, 2015
[9] S. Shahid, S. H. Pour, and S. Harun, Genetic programming for the downscaling of extreme rainfall events on the East Coast of Peninsular Malaysia. Atmosphere, Vol 5(4), pp: 914-936, 2014
[10] M. Sharma, P. Coulibaly, Y. Dibike, Assessing the need for Downscaling RCM Data for Hydrologic Impact Study. Journal of Hydrologic Eng. Vol 16(534) pp.1943-5584, 2010
[11] R. L. Wilby, and C. W. Dawson, SDSM 4.2-A decision support tool for the assessment of regional climate change impacts. User Manual. London, UK, 2007
[12] M, Noor, T. Ismail, S. Shahid, K. Ahmed, E-S. Chung, and N. Nawaz, Selection of CMIP5 multi-model ensemble for the projection of spatial and temporal variability of rainfall in Peninsular Malaysia, Theor. Appl. Climatol., Vol. 137 (331) pp. 1-14, 2019
[13] A. H. A. Tarmizi, S. N. Rahmat, A. T. Abd Karim, N. N. A. Tukimat, Climate change and its impact on rainfall, International Journal of Integrated Engineering, Vol 11 (1) pp. 170-177, 2019
[14] [14] T. Tahir, A. M. Hashim, and K. W. Yusof, Statistical downscaling of rainfall under transitional climate in Limbang River Basin by using SDSM, IOP Conference Series Earth Environment Science., Vol. 140 (1), 2018
[15] N. N. A. Tukimat, A. S. Othman, S. Wan Ahmad, K. Muthusamy, Evaluation of climate
variability performances using statistical climate models, *Sains Malaysiana*, Vol 47 (1) (2018), pp. 77-84, 2018

[16] B. Lloyd- Hughes, and M. A. Saunders, A drought climatology for Europe. *International Journal of Climatology*, Vol. 22(13), 1571-1592, 2002

[17] M. A. A. Zarch, H. Malekinezhad, M. B. Mobin, M. T. Dastorani, and M. R. Kousari, Drought monitoring by reconnaissance drought index (RDI) in Iran. *Water resources management*, Vol 25(13), pp: 3485, 2011.

[18] N. Wangsoh, W. Wathayu, and D. Sukawat, A Hybrid Climate Model for Rainfall Forecasting based on Combination of Self-Organizing Mao and Analog Method, *Sains Malaysiana* Vol 46(12) pp: 2541-2547, 2017

[19] P. Willem, and M. Vrac, M., Statistical Precipitation Downscaling for Small-Scale Hydrological Impact Investigation of Climate Change. *Journal of Hydrology*. Vol 402 pp 193-205, 2011

[20] F. M. Hamzah, M. S. Farahani, J. Othman, Identifying the Monotonic Trend in Climate Change Parameter in Kluang and Senai, Johor, Malaysia. *Sains Malaysiana*, Vol 46(10) pp: 1735-1741, 2017

**Acknowledgments**

This research is supported by Universiti Malaysia Pahang (grant vot RDU1803156), Ministry of Higher Education (MHE) (FRGS/KPT/2019 grant vot RDU1901141), Malaysian Meteorological Department (MMD), and Drainage and Irrigation Department (DID).