MODELLING MONTHLY RAINFALL IN PENINSULAR MALAYSIA USING TWEEDIE DISTRIBUTION

MUHAMAD HANIF BIN AZMI

A dissertation submitted in partial fulfilment of
the requirements for the award of the degree of
Master of Science

Faculty of Science
Universiti Teknologi Malaysia

JANUARY 2019

DEDICATION

Specially dedicated to my beloved parents,
Azmi bin Alias and Normala binti Shuib,
my family and friends for their support and encouragements.
Thank you to Dr. Shariffah Suhaila who has guide and inspire me
through my Master dissertation project.
I am really appreciating it

ACKNOWLEDGEMENT

First and foremost, I am very grateful to Allah the Almighty for giving me the strength to complete this dissertation. I take this opportunity to express my sincere gratitude and appreciation to my supervisor, Dr. Shariffah Suhaila who had gave me a lot of guidance and encouragement. I am really thankful and appreciate for her valuable guidance and encouragement extended to me.

A special thanks to my family. I feel a deep sense of gratitude for my parents and family members who form part of my vision and taught me good things that really matter in life. Their prayers and moral support has always been my strength. Their patience and sacrifice will remain my inspiration throughout my life.

I also would like to extend my sincere appreciation to my friends who kindly giving their suggestions, comments and support. Last but not least, I would like to thank to all of those who supported me in any respect during the completion of the thesis.

ABSTRACT

This study focused on modelling rainfall process by using Generalised Linear Model (GLM) when the response variable follows a distribution that comes from Tweedie family of distributions. The Tweedie family belong to the class of exponential dispersion models where the variance is proportional to some power of the mean. A special case of power-variance family of distributions, called the Tweedie distribution is used to handle the continuous real data with a discrete mass at zero. These distributions has been previously used to enable a single model for rainfall to be produced. In this context, the present study applied the Tweedie family of distributions to fit the monthly rainfall data from 10 selected rain gauge stations in Peninsular Malaysia which covers the period from January, 1980 to December, 2015. The aim of this study is to determine whether the distributions within the Tweedie family fit well the monthly Malaysian rainfall series. First, the possibility that different distributions are needed for each station was explored by estimating the index parameter, p. To do so, the profile likelihood plot is used to estimate the p index for which the log-likelihood is maximized and hence the appropriate distribution within the Tweedie family was identified. Within the Tweedie family, the Tweedie distribution was found appropriate to model both rainfall occurrence and rainfall amount simultaneously. Then, Tweedie GLM with sine and cosine functions of different harmonics are used to account for cyclical rainfall pattern. The results indicated that the stations in the eastern region are best described with one harmonics while two harmonics are required for the stations in the west and northwest region. The plots of the predicted mean monthly rainfall from the simulated data shows a resemblance to that of the observed data. Finally, reparameterization of Tweedie parameters was adopted and found to predict the amount of rain per rainfall event and probability of zero rainfall well for all months. The model provide a good description and useful information for describing the cyclical nature of rainfall pattern in the studied stations.

ABSTRAK

Kajian ini memberi tumpuan kepada pemodelan proses hujan dengan menggunakan Model Linear Umum (GLM) apabila pembolehubah tindak balas mengikuti taburan yang berasal dari taburan keluarga Tweedie. Keluarga Tweedie tergolong dalam kelas model taburan eksponen di mana varians berkadaran dengan beberapa kuasa min. Satu kes khas taburan keluarga kuasa-varians, yang dikenali sebagai taburan Tweedie digunakan untuk mengendalikan data sebenar yang berterusan dengan jisim diskret pada sifar. Taburan ini telah digunakan sebelumnya untuk membolehkan satu model hujan dihasilkan. Dalam konteks ini, kajian ini menggunakan taburan keluarga Tweedie untuk menyesuaikan data hujan bulanan dari 10 stesen pengesan hujan terpilih di Semenanjung Malaysia yang merangkumi tempoh dari Januari, 1980 hingga Disember, 2015. Tujuan kajian ini adalah untuk menentukan sama ada taburan dalam keluarga Tweedie sesuai dengan data bulanan hujan Malaysia. Pertama, kemungkinan bahawa taburan yang berbeza diperlukan untuk setiap stesen dianggarkan mengguna parameter indeks, p. Untuk berbuat demikian, plot kemungkinan profil digunakan untuk menganggarkan indeks p yang mana kemungkinan log dimaksimumkan dan dengan itu taburan yang sesuai dalam keluarga Tweedie telah dikenalpasti. Di dalam keluarga Tweedie, taburan Tweedie didapati sesuai untuk memaparkan kedua-dua kejadian hujan dan jumlah hujan secara bersamaan. Kemudian, Tweedie GLM dengan fungsi sinus dan kosinus yang berbeza harmonik digunakan untuk mengambil kira corak kitaran hujan. Hasilnya menunjukkan stesen-stesen di wilayah timur paling baik digambarkan dengan satu harmonik manakala dua harmonik diperlukan untuk stesen di wilayah barat dan barat laut. Plot purata hujan yang diramalkan dari data simulasi menunjukkan persamaan dengan data yang diperhatikan. Akhir sekali, pengukuran semula parameter Tweedie digunakan dan didapati meramalkan jumlah hujan setiap kejadian hujan dan kebarangkalian hujan sifar dengan baik bagi setiap bulan. Model ini memberikan gambaran yang baik dan maklumat yang berguna untuk menggambarkan corak kitaran hujan di stesen yang dikaji.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

EDM	-	Exponential Dispersion Model
GLM	-	Generalized Linear Models
MCME	-	Markov Chain Mixed Exponential
ARIMA	-	Autoregressive Integrated Moving Average
GAM	-	Generalized Additive Model
SOI	-	Southern Oscillation Index
CV	-	Correlation of Variation
MLE	-	Maximum Likelihood Estimate
CI	-	Confidence Interval

LIST OF SYMBOLS

p	-	Index parameter
μ	-	Mean
$V(\mu)$	-	Variance function
σ	-	Standard deviation
$\emptyset$	-	Dispersion parameter
X	-	Vector of predictor
β	-	Vector of parameter coefficient
η	-	Linear component of GLM
θ	-	Canonical parameter for EDM
ℓ	-	Log likelihood function
j	-	Harmonic number for the series
m	-	Month of the year
$D(y;\mu)$	-	Deviance of the response variable
χ^2	-	Chi-square
λ	-	Mean number of rainfall events
γ	-	Scale of the gamma distribution
α	-	Shape of the gamma distribution
$\alpha\gamma$	-	Mean amount of rain per rainfall event

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Malaysia experiences humid tropical climate and high intensity rainfall where abundant amount of rain is received throughout the year. The distinction between climates over Peninsular Malaysia is largely determined by uneven rainfall distribution. According to Suhaila and Jemain (2007), the annual rainfall amount in Malaysia is recorded approximately between 2000 mm to 4000 mm with the number of rainy day are observed between 150 to 200 days. The uneven rainfall distribution in Malaysia is greatly influenced by two monsoon seasons, Southwest monsoon which begins in the mid of May and ends in August and Northeast monsoon between November till February. The Southwest monsoon bring heavy rainfall to the west coastal regions of Malaysia. Meanwhile, the Northeast monsoon, causes the heaviest rainfall in the east coastal regions. It is important to note that the rainy seasons governed by monsoon winds typically bring heavy rainfall that may trigger severe flooding that could impact the changes in climate. In brief, it is anticipated that the variability in rainfall distribution due to climate change may make the Peninsular region more vulnerable to extreme events. Therefore, it is necessary to study rainfall process as to adapt with changing pattern of climate in the region.

The ability of rainfall modelling has promising applications for forecast analyses and simulation purposes in many areas of planning including hydrological systems, agriculture, weather derivatives and water resource systems. Also, modelling rainfall is beneficial in understanding the variability and uncertainty in rainfall processes. Indeed, understanding rainfall behavioral is vital to preserve the environment from natural disasters such as floods and droughts, and for better

management of water resources. Although this issue is crucial, but yet the studies for modelling rainfall data in Malaysia are still exhaustive. Hence, this study aims at enhancing the understanding in rainfall process by modelling rainfall data.

Rainfall data at different timescales have a mixture of discrete (exact zeros when no rain falls) and continuous (amount of rainfall in rainy days) components. In practice, two separate models are usually used to capture both rainfall components. The first model is best described as rainfall occurrence that simulates a sequences of wet and dry days, while the second model is better known as rainfall amount that simulates the rainfall intensities on wet days. In literatures, numerous studies have been conducted to model rainfall occurrence (Deni *et al*., 2009; Sinha *et al*., 2011; Basak, 2014) and there have also been various studies on modelling rainfall amount (Hussain et al., 2010; Suhaimi and Suhaila, 2016). Although many studies in the past has proved to be adequate to model rainfall occurrence and rainfall amount independently, however, these study provides only limited information and hence is not efficient enough to fully describe rainfall processes.

Modelling rainfall processes separately poses a lot of challenges. Direct modelling of rainfall processes is often problematic because rainfall data is continuously distributed and included observations that are identically zero. This makes it difficult to transform the data to normality by power transforms or to model it directly using continuous distribution. Consequently, it makes many of the available statistical methods for continuous data inappropriate for modelling rainfall since they do not have a point mass at zero. In addition, the use of two separate models will cause a large number of parameters to be estimated in the models and hence becoming more complex. The present study therefore seek to resolve this problems by modelling both rainfall components together.

A study by Dunn (2004) suggested that Tweedie family distributions is useful in modelling rainfall occurrence and rainfall amount simultaneously. Within the Tweedie family, the Poisson-gamma distribution (also called Tweedie distribution) was found sufficient to represent both rainfall components in a single complete rainfall model (Dunn, 2004; Hasan and Dunn, 2011). For instance, a Poisson-gamma distribution fits well to the monthly rainfall data for most Australian stations (Hasan

and Dunn, 2010). The study found that the simulated data have very similar properties to the observed data sets based on studying the probability of no rainfall and mean rainfall amounts. The distribution has been proven to be adequate for model both rainfall components simultaneously. Accordingly, the appropriateness of Tweedie distribution to adequately model Malaysian rainfall series is need to be studied in detail.

Tweedie distribution belong to the exponential dispersion model (EDM) family of distributions (Jørgensen, 1997) which form the basis of generalized linear model (GLM) (McCullagh and Nelder, 1989). Dunn (2004) points out several important properties of Tweedie distribution that makes them particularly exceptional for use in rainfall modelling. This includes their ability to model both occurrence and amount of rainfall concurrently. One of the advantages of using Tweedie distribution is that the motivation behind the setup of the model is simple and logical which considers total rainfall as a sum of rainfall amounts on a smaller timescale. In addition, they are sufficiently great to model skewed continuous variables with exact zeros and can be used to fit into the GLM framework. This is advantageous because fitting models and diagnostic testing are readily available for GLM and thus for Tweedie distribution models as well. Besides that, Tweedie distribution provide a mechanism in which finer scale structure can be understood through coarser scale data.

The reason for concern in Tweedie distribution is its applications associated with generalized linear models. GLM provides a flexible and rigorous framework to handle response with high levels of variability such as in rainfall data (Chandler and Wheater, 2002). In the initial formulation of GLM by Nelder and Wedderburn (1972) allowed to model those response variables that have distribution other than normal distribution. This allow for any other distributions from exponential family such as Poisson, exponential, gamma and Tweedie distribution to be used to fit a GLM. It is then possible to demonstrate the advantages of using Tweedie distribution which particularly useful in modelling rainfall. Indeed, application of Tweedie distribution associated with GLM have been found practical in modelling rainfall series (Hasan and Dunn, 2010; Hasan and Croke, 2013). However, fitting such Tweedie GLM to monthly rainfall totals for Malaysian stations are still lack in literatures. Thus,

considering this distribution, Tweedie GLM are proposed to model monthly Malaysians rainfall series.

1.2 Problem Statement

The critical issues of rainfall variability and uncertainty over the region of Peninsular Malaysia is considered crucial due to frequent incidents of huge floods and drought being recorded in recent year. Recognizing the importance of these issues has lead the study to model the rainfall distributions. Rainfall distributions are typically modelled into two separate components, either modelling rainfall occurrence or rainfall amount. However, there is a general consensus among researchers that modelling rainfall data separately may not reveal the true features of rainfall process or may serve only certain information of the rainfall features. Therefore, there is a critical need to have such a competent model that could simultaneously model both rainfall occurrence and rainfall amount at once.

To develop a practical statistical model for handling a mixture of both discrete and continuous components is the main challenge faced by researchers. In rainfall modelling, the difficulty involved when trying to accommodate the zero rainfall and non-zero rainfall amount together. Thus, some modifications or new approach is needed. One approach is to use Tweedie distribution that allows different component of rainfall to be modelled in a single complete stochastics model. In the case of monthly Malaysian rainfall series, it is seems necessary to employ Tweedie distribution since the series have exact zeros combined with continuous rainfall totals. Therefore, Tweedie generalized linear models are proposed in this study to handle such rainfall series.

1.3 Objective of the Study

The main objectives of this study are:

1. To examine monthly Malaysian rainfall series and determine the appropriate distribution within the Tweedie family distributions.
2. To fit rainfall occurrence and rainfall amount using Tweedie generalized linear model.
3. To reparameterize Tweedie distribution parameters to the Poisson and gamma parameters.

1.4 Scope of the Study

This study specifically examines rainfall data by modelling rainfall processes using GLM, when the response variable is assumed to have a distribution that comes from the Tweedie family of distributions. Monthly rainfall data available for the period from 1980 to 2015 are considered for all ten selected rain gauge stations across Peninsular Malaysia. Preliminary analyses are performed on these data sets to determine the preferred distributions within the Tweedie family to model the rainfall data. Then, generalized linear models are used to fit the model. Monthly rainfall shows an annual cycle, and thus to account for the cyclical patterns, the explanatory variables in the form of sine and cosine terms were included into the model framework. The parameter required to create the model were estimated using maximum likelihood estimation.

1.5 Significance of the Study

The significant contribution of this study is to provide rainfall model for describing the rainfall distribution in Peninsular Malaysia. The model that is suggested has a great application in rainfall analysis. The results from this study can be used to capture many desirable features of rainfall such as the number of rainfall events, the amount of rainfall per events and the probability of recording no rain. These rainfall

features provide valuable information in modelling purposes. For example, rainfall modelling assists in climate related and decision making capabilities especially in improving the accessibility of weather information.

More interestingly, rainfall modelling give large contribution in agricultural planning. In agriculture, better understanding of various rainfall characteristics is essential for optimizing farm production, precision farming and also timing for boost crop. It is not surprising that prolonged drought will hinder the productivity of the sector or may stop the planting activity while severe floods will damage the crops. The study of rainfall modelling is viewed to be able to overcome these disasters through proper planning in agriculture sector and hence aids to increase the crop productivity.

Besides that, rainfall models can be a great tool in making scientific decisions on long term planning and management in various sector where rain is considered as a major driving factor such as in water cycle management and hydrology planning. For example, simulated rainfall data can be used as an input into the management of water cycle to manage watersheds, to simulate the infiltration process of groundwater and to simulate stream flow.

1.6 Outline of the Study

This dissertation is organized in five chapters and structured as follows: Chapter 1 presents a brief introduction on the studied area including the purposes, motivation and significant of this study. Chapter 2 provide overview of past studies on modelling rainfall occurrence and rainfall amount. Following on, Chapter 3 introduced generalized linear model and discusses the framework of these increasingly popular models. Of particular importance is the discussion on the Tweedie family of distributions and their application on modelling rainfall. This forms a fundamental parts of this dissertation. In Chapter 4, the preliminary analyses are performed on the studied data sets to test their suitability to fit the model. This chapter ends with the model interpretation. Finally, Chapter 5 gives summarizes of this study and proposed some improvements that could be made. Recommendations for further research on this area are also suggested.

REFERENCES

Abubakar, U.Y., Lawal, A. and Muhammed A. (2013). The Use of Markov Model in Continuous Time for the Prediction of Rainfall for Crop Production. International Organization for Scientific Research. 7(1), 38-45.

Abdulkadir T.S., Salami A.W., Aremu A.S., Ayanshola A.M. and Oyejobi D.O. (2017). Assessment of neural networks performance in modeling rainfall amounts. *Journal of Research in Forestry, Wildlife & Environment,* 9(1), 1-8.

Aksoy, H. (2000). Use of gamma distribution in hydrological analysis. *Turkish Journal of Engineering and Environmental Sciences*, 24(6), 419-428.

Basak, P. (2014). On the Markov chain models for monsoonal rainfall occurrence in different zones of West Bengal. *Indian Journal of Radio & Space Physics, 43*, 349-354.

Barkotulla, M. A. B. (2010). Stochastic generation of the occurrence and amount of daily rainfall. *Pakistan Journal of Statistics and Operation Research*, 6(1), 61-74.

Brath, A., Montanari, A. and Toth, E. (2002). Neural networks and non-parametric methods for improving real-time flood forecasting through conceptual hydrological models. *Hydrology and Earth System Sciences Discussions, 6*(4), 627-639.

Chandler, R. E. and Wheater, H. S. (2002). Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. *Water Resources Research.* 38(10), 1192–1202.

Coe, R. and Stern, R. D. (1982). The use of rainfall models in agricultural planning. *Journal of Applied Meteorology.* 21, 1024-1031.

Deni, S. M., Jemain, A. A. and Ibrahim K. (2009). Fitting optimum order Markov chain models for daily rainfall occurrences in peninsular Malaysia. *Theoretical and Applied Climatology.* 97, 109–121.

Dobson, A. J. (2002). *An Introduction to Generalized Linear Models.* (3rd edition). Chapman and Hall, London.

Dunn, P. K. (2004). Occurrence and quantity of precipitation can be modelled simultaneously. *International Journal of Climatology.* 24, 1231–1239.

Dunn, P. K. (2017). Evaluation of Tweedie Exponential Family Models. Package 'tweedie'.

Dunn, P. K. and Smyth, G. K. (2005). Series evaluation of Tweedie exponential dispersion model densities. *Statistics and Computing.* 15, 267–280.

Dunn, P. K., and Smyth, G. K. (2008). Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. *Statistics and Computing*, 18(1), 73-86.

Fadhilah, Y., Zalina, M. D., Nguyen, V. T. V. and Zulkifli, Y. (2007b). Stochastic Modeling of Hourly Rainfall Series using Markov Chain Mixed Exponential (MCME) Model. Prosiding Simposium Kebangsaan Sains Matematik ke-15, Concorde Hotel, Shah Alam, Selangor. June 5 -7 2007, ms. 503-510.

Fowler, H. J., Kilsby, C. G., O'Connell, P. E. and Burton A. (2000). A stochastic rainfall model for the assessment of regional water resource systems under changed climatic conditions. *Hydrology and Earth System Sciences.* 4, 261-280.

Gabriel, K. R. and Neumann, J. (1962). A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of Royal Meteorology Society*, 88, 90-95.

Hasan, M. M., and Croke, B. (2013). Filling gaps in daily data: a statistical approach. *20th International Congress on Modelling and Simulation.*

Hasan, Md. M. and Dunn, P. K. (2010). A simple Poisson-Gamma model for modeling rainfall occurrence and amount simultaneously. *Agricultural and Forest Meteorology* 150, 1319-1330.

Hasan, Md. M. and Dunn, P. K. (2011). Two Tweedie distributions that are near-optimal for modelling monthly rainfall in Australia. *International Journal of Climatology*, 31(9), 1389-1397.

Hasan, Md. M. and Dunn, P. K. (2012). Understanding the effect of climatology on monthly rainfall amounts in Australia using Tweedie GLMs. *International Journal of Climatology*, 32, 1006-1017.

Hasan, Md. M. and Dunn, P. K. (2015). Seasonal rainfall totals of Australian stations can be modelled with distributions from the Tweedie family. *International Journal of Climatology*, 35(10), 3093-3101.

Husak, G. J., Michaelsen, J. and Funk, C. (2007). Use of the Gamma distribution to represent monthly rainfall in Africa for drought monitoring applications.

International Journal of Climatology: A Journal of the royal Meteorological Society, 27(7), 935-944.

Hussain, A. (2008). Stochastic Modelling of the Daily Rainfall Process: An Application of a MCME Model in Different Climatic Conditions. McGill University, Quebec, Canada. Master thesis.

Hussain, Z., Mahmood, Z. and Hayat, Y. (2010). Modeling the daily rainfall amounts of north-west Pakistan for agricultural planning. *Journal of Agriculture.* 27(2), 313-321.

Hyndmann, R. J. and Grunwald, G. K. (2000). Applications: Generalized additive of mixed distribution Markov models with application to Melbourne's rainfall. *Australian & New Zealand Journal of statistics*, 42(2), 154-158.

Jørgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of Royal Statistical. Society Series B Statistical Methodology*. 49, 127–162.

Jørgensen, B. (1997). *The Theory of Dispersion Models.* Chapman and Hall, London.

Juneng, L., Tangang, F. T. and Reason, C. J. C. (2007). Numerical case study of an extreme rainfall event over the east coast of Peninsular Malaysia. *Meteorology and Atmospheric Physics*, 98(1), 81-98.

Lall, S. (2013) A study of Tweedie family of distributions for rainfall modelling. Indian Agricultural Statistics Research Institute, New Delhi, India. Master thesis.

Lennox, S. M., Dunn, P. K., Power, B. D., and DeVoil, P. (2004). A statistical distribution for modelling rainfall with promising application in crop science. *In Proceedings of the 4th International Crop Science Congress.* Regional Institute Ltd.

Mahsin, M., Akhter, Y. and Begum, M. (2012). Modelling rainfall in Dhaka Division of Bangadesh using time series analysis, *Journal of mathematical modelling and application*, 1 (5), 67-73.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Model.* (2nd edition). Chapman and Hall, London.

Mielke, P. W. (1973). Another family of distributions for describing and analyzing precipitation data. *Journal of Applied Meteorology*, 12, 275-280.

Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A*, 135, 370-384.

Robertson, A. W., Kirshner, S. and Smtyh, P. (2003). Hidden Markov Models for Modelling Daily Rainfall Occurrence over Brazil. University of California. Technical report.

Roldan, J. and Woolhiser, D. A. (1982). Stochastic daily precipitation models. A comparison of occurrence process. *Water Resources Research*, 18(5), 1451-1459.

Sharda, V. N. and Das, P. K. (2005). Modelling weekly rainfall data for crop planning in a sub-humid climate of India. Agriculture Water Management, 76, 120-138.

Sinha, N. C., Islam, M. A. and Ahamed, K. S. (2011). Logistic Regression Models for Higher Order Transition Probabilities of Markov Chain for Analyzing the Occurrences of Daily Rainfall Data. *Journal of Modern Applied Statistical Methods*, 10(1), 337-348.

Stern, R. D. and Coe, R. (1984). A model fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society*, 147, 1–34.

Suhaila, J., and Jemain, A. A. (2007). Fitting Daily Rainfall Amount in Peninsular Malaysia Using Several Types of Exponential Distributions. *Journal of Applied Sciences Research.* 3(10), 1027-1036.

Suhaila, J., and Jemain, A. A. (2009). A comparison of the rainfall patterns between stations on the East and the West coasts of Peninsular Malaysia using the smoothing model of rainfall amounts. *Meteorological Applications*, *16*(3), 391-401.

Suhaila, J., Ching-Yee, K., Fadhilah, Y. and Hui-Mean, F. (2011). Introducing the mixed distribution in fitting rainfall data. *Open Journal of Modern Hydrology*, 1(02), 11.

Suhaimi, N. H. and Suhaila, S. (2016). Generalized Additive Model approach in modeling rainfall data in Johor area. *American Institute of Physics Conference Proceedings.*

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the gaussnewton method. *Biometrika*, 61(3), 439-447.

Wilks, D. S. (1990). Maximum likelihood estimation for the gamma distribution using data containing zeros. Journal of Climate, 3, 1495-1501.

Wilks, D. S. (1999). Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteotrology*, 93, 153-169.

Wong, C. L., Venneker, R., Uhlenbrook, S., Jamil, A. B. M. and Zhou, Y (2009). Variability of rainfall in Peninsular Malaysia. *Hydrology and Earth System Sciences Discussions*, 6(4), 5471-5503.

Wright, E. M. (1933). On the coefficients of power series having exponential singularities. *Journal of the London Mathematical Society*, *1*(1), 71-79.

Yunus, R. M., Hasan, M. M. and Zubairi, Y. Z. (2017a). Fitting monthly Peninsular Malaysian rainfall using Tweedie distribution. In *Journal of Physics: Conference Series*, 890(1).

Yunus, R. M., Hasan, M. M., Razak, N. A., Zubairi, Y. Z. and Dunn, P. K. (2017b). Modelling daily rainfall with climatological predictors: Poisson-gamma generalized linear modelling approach. *International Journal of Climatology*, 37(3), 1391-1399.

Yusof, F., Yung, L. M. and Yusop, Z. (2015). Tropical daily rainfall amount modelling using Markov Chain-Mixed Exponential (MCME). *Jurnal Teknologi*, *74*.

Yusuf, A. U., Adamu, L. and Abdullahi, M. (2014). Markov Chain Model and Its Application to Annual Rainfall Distribution for Crop Production. *American Journal of Theoretical and Applied Statistics*, 3(2): 39-43.