MODELLING THE DEPENDENCE STRUCTURE BETWEEN TWO RAINFALL STATIONS IN JOHOR

KONG CHING YEE

UNIVERSITI TEKNOLOGI MALAYSIA

MODELLING THE DEPENDENCE STRUCTURE BETWEEN TWO RAINFALL STATIONS IN JOHOR

KONG CHING YEE

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Mathematics)

Faculty of Science
Universiti Teknologi Malaysia

AUGUST 2019

ACKNOWLEDGEMENT

First of all, I want to express my deepest gratitude to my main supervisor, Dr. Shariffah Suhaila bt. Syed Jamaludin and co-supervisor, Professor Dr. Fadhilah bt. Yusof for their comments, guidance and invaluable knowledge in preparing and completing this thesis.

I appreciate the cooperation given by the Malaysian Meteorological Department (MMD) and Malaysian Irrigation and Drainage Department (DID) for providing the rainfall data. I also feel grateful to the Ministry of Education who sponsored me through MyPhD Scholarship and Universiti Teknologi Malaysia (UTM).

I would also like to express my appreciation to my beloved family especially my parents for their unlimited love, care and support throughout my study in UTM. Besides, I also feel grateful to my friends who have provided me suggestions to improve this thesis.

Finally, I would like to thank all the people who gave their cooperation, directly or indirectly, and to those who were involved in the commencement and improvement of this thesis.

ABSTRACT

Rainfall data consist of zero and real rain values. In many rainfall models, the zero values are not seriously considered in the analysis, which may lead to the loss of some important information. Therefore, to preserve sufficient information, the effect of zero values needs to be examined before its omission from the analysis. In this study, a bivariate mixed model that consists of continuous and discrete distribution is employed to disclose the portion of zero values in the analysis where the possibility of no rain phenomenon characteristics in the data can be included. The rainfall data used are taken from two rainfall stations that are classified into three cases: data with only positive values (non-zero values) recorded at both stations, data with positive values recorded in either one of the stations and all data values including zeroes recorded at both rainfall stations. The interstation correlation coefficients of the bivariate mixed distribution under these three cases are then used to detect the importance of the zero values. Results show that the case, data with only positive values recorded at both stations is the best. In addition, the rainfall characteristics of two stations that are nearby and located in the same river basin can be different due to their different spatial conditions. However, one of the limitations of bivariate distribution is that all its univariate marginal distributions are assumed to be the same type of distribution, yet there are neighbouring stations that have different types of distributions. Hence, the Copula model is then proposed to describe the dependency between two stations without considering the effect of the marginal distributions. Based on the rainfall data that contain only non-zero values for both stations, Galambos distribution is found as the best Copula model in describing the dependencies between the two stations in Johor area. Lastly, the dependencies parameters of bivariate mixed distribution and Copula distribution are proposed as spatial weighting methods in estimating the rainfall values at unsampled location.

ABSTRAK

Data hujan mengandungi nilai sifar dan nilai hujan. Terdapat banyak kajian pemodelan hujan mengecualikan nilai sifar dari analisis dan menyebabkan kehilangan maklumat penting. Oleh itu, untuk menjamin maklumat yang cukup diperolehi, kepentingan nilai sifar perlu diselidiki sebelum dikecualikan dari analisis. Dalam kajian ini, model taburan campuran dwi pembolehubah yang terdiri daripada taburan selanjar dan diskret digunakan untuk merangkumi nilai sifar dalam analisis supaya ciri fenomena tidak ada hujan dapat dimasukkan. Data hujan yang digunakan telah diambil daripada dua stesen hujan yang diklasifikasikan kepada tiga kes: data dengan nilai positif (nilai bukan sifar) sahaja yang direkodkan di kedua-dua stesen hujan, data dengan nilai positif yang direkodkan di salah satu stesen, dan semua nilai data termasuk sifar yang direkodkan di kedua-dua stesen hujan. Pekali korelasi antara stesen dari model taburan campuran dwi-pembolehubah diperoleh di bawah tiga kes ini digunakan untuk memeriksa kepentingan nilai sifar. Keputusan menunjukkan bahawa kes data dengan nilai positif sahaja yang direkodkan di kedua-dua stesen hujan adalah terbaik. Selain itu, ciri-ciri hujan di dua stesen yang berhampiran dan berlokasi di lembangan sungai yang sama boleh berbeza kerana keadaan reruang yang berbeza. Tetapi salah satu kelemahan model taburan dwi-pembolehubah adalah taburan marginal univariatnya mesti daripada taburan yang sama, namun terdapat stesen hujan berhampiran mempunyai taburan yang berbeza. Oleh itu, Copula model telah dicadangkan untuk menghuraikan perhubungan taburan kebarangkalian antara dua stesen hujan tanpa mempertimbangkan kesan taburan marginal. Berdasarkan data hujan yang hanya mengandungi nilai positif untuk kedua-dua stesen hujan, taburan Galambos telah didapati sebagai model Copula terbaik untuk menggambarkan kebergantungan antara kedua-dua stesen di kawasan Johor. Akhir sekali, parameter kebergantungan dari model taburan campuran dwi-pembolehubah dan taburan Copula dicadangkan sebagai kaedah pemberat reruang untuk menganggar nilai hujan di lokasi yang tiada rekod.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	-	Akaike Information Criterion
AMH	-	Ali-Mikhail-Haq
DID	-	Malaysian Irrigation and Drainage Department
CCWM	-	Modified Coefficient of Correlation Weighting Method
CDF	-	Cumulative Distribution Function
CIDW	-	Modified Correlation Coefficient with Inverse Distance Weighting Method
CV	-	Coefficient of Variation
GH	-	Gumbel-Hougaard
IFM	-	Inference Functions for Margins
MAE	-	Mean Absolute Error
MMD	-	Malaysian Meteorological Department
MLE	-	Maximum Likelihood Estimation
MPL	-	Maximum Pseudo-Likelihood
NRIDW	-	Modified Normal Ratio with Inverse Distance Method
pdf	-	Probability Density Function
SD	-	Standard Derivation
S-index	-	Similarity Index
UTM	-	Universiti Teknologi Malaysia

LIST OF SYMBOLS

λ	-	Continuous inverse scale parameter for Exponential distribution
μ	-	Continuous location parameter for Lognormal distribution
σ	-	Continuous scale parameter for Lognormal distribution
β	-	Continuous scale parameter for Gamma and Weibull distributi
α	-	Continuous shape parameter for Gamma and Weibull distributi
C(u,v)	-	Copula Function
θ	-	Copula parameter
$\emptyset(t)$	-	Generator of Archimedean Copula
$\emptyset^{-1}(s)$	-	Inverse generator
τ	-	Kendall's τ
ρ	-	Spearman's ρ
$A(t)$	-	Pickands dependence
Φ	-	Laplace integral
$\Gamma(\alpha)$	-	The Gamma function
R	-	Coefficient of Correlation
X_t	-	Estimated Value of the Missing Data at the Target Station
N	-	Number of Neigboring Stations Used
X_i	-	Observation at the ith Neighbouring Stations
W_i		Weight of the ith Neighbouring Stations

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

Rainfall is the most important variable in any kind of study such as climate, agriculture, hydrology, and water resource management which could provide essential information and its input is very beneficial to these studies (Srikanthan and McMahan, 2001; Serinaldi, 2009b). In analysing and modelling rainfall data, the characteristics of rainfall, such as amount, intensity, frequency, duration and seasonal distributions of rainfall are indispensable. Hence, a suitable model that can be proposed to describe rainfall patterns, while preserving most of the rainfall characteristics, is essential for related hydrological researches.

For instance, the rainfall pattern during dry and wet spells is vital to agriculture to obtain the maximum harvest. A long dry spell may occur when the plant is specifically sensitive, such as just after germination, in the season of flowering, or even the growing seasons. Thus, this provides enough supportive information for decision making in planning the plantation of the crops. Being able to estimate the probability of prolonged dry spell during these periods will be useful. Besides dry spells, the time at the end of the wet season is also prominent. Crops may not have adequate water to reach maturity if the wet season ends too soon. Not only that, immoderate wet weather may impede ripening or harvesting. Thus, understanding the characteristics of rainfall at the studied location is very important for providing information to the authorities, particularly those who are related to the water system management (Hoogmoed and Klaij, 1990; Adiku et al., 1997; Yemenu and Chemeda, 2013; Myers et al., 1998).

Modelling an appropriate distribution to a rain gauge station is one of the most critical parts in studying the characteristics of rainfall. Rainfall modelling can be separated into two categories: rainfall amount and occurrence. Rainfall amount takes into account the amount of rainfall on rainy days while the analysis of rainfall occurrence deals with the sequences of wet and dry days. Gamma distribution is the most popular option in fitting the rainfall amount. On top of that, Weibull, mixed Exponential and Lognormal are also used as the alternative distributions to model skewed rainfall data. On the other hand, rainfall occurrence is often analysed by applying Markov chain (Stern and Coe, 1984; Deni et al., 2009; Sadiq, 2014; Chowdhury et al., 2017).

Rainfall data is characterized by zero and non-zero values, where zero values indicate non-rainy days while non-zero values indicate rainy days. These values are represented by a combination of discrete (zero values) and continuous (non-zero values) parts in distribution fitting. Zero values, or non-rainy days, are important to explain the characteristics of drought and climate changes (Ha and Yoo, 2007; Pakoksung and Takagi, 2017). However, the importance of zero values is not evident and is often ignored by many researchers due to the difficulty in combining the zero values with continuous rainfall values. In rainfall amount analyses, mostly the rainfalls with zero values were excluded. Nevertheless, studies of rainfall occurrence using the Markov chain method takes into consideration the order of zero and non-zero values, but the model fit the zero and non-zero values separately.

The exclusion of zero values from the rainfall data might affect the modelling of rainfall distribution. Some vital rainfall information and characteristics may be neglected or lost if zero values are not taken into account in the analysis. Hence, researchers put an effort to combine the discrete and continuous part of rainfall data in fitting rainfall distribution (Dzupire, Ngare and Odongo, 2018); yet, the distributions created are complicated. Therefore, to preserve sufficient information for analysis, the effect of zero values toward rainfall data needs to be checked before being omitted.

An alternative way to perform the analysis which consists of continuous (non-zero values) and discrete (zero values) distribution is by using a mixed distribution. Generally, a probability distribution is either discrete or continuous. A mixed distribution is a distribution that is neither discrete nor continuous but rather combines the discrete and continuous elements in the distribution (Kedem, Chiu and

North, 1990). By applying a mixed distribution, the possibility of no rain and the skewness of real rain can be included in a probability distribution.

Mixed distribution can be applied to univariate and bivariate cases (Shimizu, 1993; Ha and Yoo, 2007; Yoo and Ha, 2007). A univariate mixed distribution model the rainfall data for one station, while bivariate mixed distribution describes the relation of rainfall between two stations. To apply the mixed bivariate case into an analysis, the data has to be restructured into three cases, which are: (a) considers only non-zero values for both locations, (b) taking data when either one of the locations recorded non-zero values, and (c) taking all values including zero values for both locations. By examining these three cases, the effect of zero measurements can be determined using the interstation correlation coefficients.

The rainfall characteristics of two stations that are close to each other are expected to show a degree of spatial association in terms of their rainfall behaviour since they tend to be wet or dry at the same time. However, the rainfall data of the two stations can be different due to different spatial conditions. One station may detect rain while the other nearby station may not receive rain at the same time. For the cases which involve two variables such as two rainfall stations with different marginal, a bivariate distribution cannot be used to find the joint distribution of the two variables, which is one of the limitations in applying bivariate distribution

Bivariate distribution requires all univariate marginal distributions to belong to the same type of distribution (Favre et al., 2004; Poulin et al., 2007). Given that some of the rainfall characteristics of two rain gauge stations are different, a bivariate distribution may not be appropriate in the modelling of rainfall data and may cause bias in the analysis. Therefore, the Copula method is proposed against such situations (Quinn, 2007).

A Copula model is designed to describe the dependency between two variables without considering the effect of the marginal distributions. The joint distribution can be modelled using one Copula function regardless of the univariate marginal distribution. Previous studies have shown that modelling based on Copulas could overcome the limitation of bivariate distribution (De Michele and

Salvadori, 2003; Favre et al., 2004; Grimaldi and Serinaldi, 2006; Zhang, Singh and Asce, 2006). Many Copula families are available, but the Copula that best describes the joint relation of rain gauge stations in Johor, Malaysia is yet to be determined. Therefore, several Copulas will be analysed to determine the best model for the joint relation of rain gauge stations in Johor area.

The study in detecting the importance of zero values has been proposed for quite some time, however, no attention has been placed by the researchers. The importance of zero values in rainfall data should be taken into account before being excluded from the analysis, especially in research regarding drought analysis and climate change. For the time being, no systematic study has been done in determining the importance of the zero value in the rainfall analysis before implementing the Copula to represent the joint relationship between the rain gauge stations in Johor. The best bivariate Copula model will be determined and can be further applied in performing spatial interpolation method.

1.2 Problem Statement

Rainfall data consists of two parts in distribution; continuous and discrete. Continuous rainfall distribution represents non-zero (rainy days) while discrete distribution represents zero values (non-rainy days). Researchers often model continuous and discrete rainfall amounts separately by ignoring the existence of zero values in their analysis. The removal of the zero values may cause the restructured data to be unable to describe the actual characteristics of the rainfall. Zero values are essential in specific analyses such as drought modelling, estimating the period of a dry spell as well as modelling of rainfall occurrences. Hence, there is a need to study the importance of zero rainfall values before deciding whether they need to be included or not in any univariate or bivariate rainfall analysis.

Rainfall data at two neighbouring stations are expected to have a high correlation and the rainfall behaviours of the two stations are assumed to be the same. However, if a location receives rainfall on a given day, there might be a possibility that nearby locations do not record any rain on that particular day. Hence, there might exist differences in rainfall distribution between two nearby stations. Without a suitable model to describe these relationships, a random field model would likely be missing some of the crucial details. Every event that occurs during the observation period (rain or no rain) is important, no matter at which geographic region because there is a possibility that these events can affect the study, especially in drought and climate change analysis. To model the rainfall data at two rainfall stations, a bivariate distribution is a possible option. Since zero and non-zero values are involved, a bivariate mixed distribution is employed for rainfall data at two rainfall stations. A bivariate mixed model has potential in describing the intermittent nature of rainfall and analysing the dependence structure of the rainfalls at two stations.

A bivariate model, which represents a joint distribution of rainfall data at two different stations, can be useful in spatial rainfall analysis. In conducting a bivariate model, two marginal distributions are assumed to be the same. In reality, the marginal distributions are not necessarily the same. Therefore, the Copula model is one of the alternative models that can handle this assumption. Besides, the mathematical formulations of bivariate distribution become complicated when the number of parameters increases, while the Copula model could be extended into multivariate analysis easily. For research that intends to add in more variables for further study, the Copula model is more suitable than a bivariate distribution. Not only that, bivariate distribution can't distinguish the marginal and joint behaviour of the variables used. The Copula function can investigate the marginal properties and dependence structure of variables separately and then can be further applied in estimating data at an unsampled location.

1.3 Research Objectives

This study embarks on the following objectives:

i. To carry out the preliminary analysis on the structure of bivariate rainfall data over rainfall stations in JohorTo determine the importance of zero values in rainfall analysis by using the bivariate mixed distribution.

ii. To estimate the parameters of Copula model families and determine the best fitted Copula.

To apply the estimated parameters derived from bivariate mixed distribution and Copula distribution for estimating the rainfall data at an unsampled location.

1.4 Scope of the Study

Daily rainfall for 28 rain gauge stations in Johor, Malaysia was obtained from the Malaysian Meteorological Department and the Malaysian Drainage and Irrigation Department. The period of rainfall data obtained is from the year 1980 to 2011. A homogeneity test had been applied to ensure the quality of the rainfall data. Standard normal homogeneity test, Buishand range test, Pettitt test and the Von Neumann ratio test proposed by Wijngaard, Klein Tank and Konnen, (2003) were among the test conducted in this study.

Mapping and kriging interpolation were generated using ArcGIS software, whereas the rainfall modelling was programmed and run under Microsoft Visual C++. On the other hand, for Copula analysis, R software programming was applied.

1.5 Significance of the Study

Rainfall data is frequently used in the study of hydrology and climate change, such as researches on rainfall-runoff, drought, spatial interpolation, and flood. In modelling the rainfall data, rainfall distribution comprised of discrete and continuous parts. However, researchers tend to exclude zero values in finding the best distribution. This problem can be solved by applying a mixed distribution that combines both discrete and continuous parts. Similarly, when dealing with the spatial correlation of rainfall, the importance of zero values of the rainfall data from two rain gauge stations also needs to be addressed. By detecting the significance of the zero values, researchers in hydrology and climate change can decide whether it is essential or not to include the zero values in their researches.

Two neighbouring rain gauge stations can have different rainfall patterns and characteristics on the same day. The situation such that one station may have rain recorded while another station has no rain recorded can occur. In this case, different marginal distributions of the joint relations between two stations might exist. When applying bivariate distribution to examine the joint relationship between the two stations, one of the limitations is that the univariate marginal distributions have to be of the same type. Biased estimation may occur if the bivariate approach does not follow the well-specified marginal distribution.

Copula function is introduced to avoid bias estimation for rainfall analysis of the relations between rain gauge stations in Johor, Malaysia. Copula function allows different marginal distributions to form a joint function. Besides, certain limitations of bivariate distribution can be overcome by using Copula distribution. Copula distribution is simpler compared to general bivariate distribution. By analysing rainfall at different rain gauge stations and their joint relationship using Copula function, their usage for further research can be extended in the study related to agriculture, hydrology, climate change or water resource management.

Issues of missing data in rainfall often happen due to unexpected human errors or instrument errors. The weighting method is one of the popular techniques in finding the missing data in rainfall. By using the estimated parameters in a bivariate mixed distribution and Copula distribution into the weighting method, the missing data for a rain gauge station could be estimated.

1.6 Organization of the Thesis

There are seven chapters comprised in this thesis as follows:

i. **Chapter 1** provides a brief discussion on the background of the research questions and problem statements. This chapter also outlines the objectives of this study, together with the research scopes. The significance of this study is also highlighted at the end of the chapter.

ii. **Chapter 2** discusses the literature review on mixed distribution and Copula function development in various fields.

iii. **Chapter 3** explains the research methodology, which interprets the fitting mixed univariate distribution, effect of zero measurements for two rain gauge stations and bivariate Copula modelling.

iv. **Chapter 4** presents the study area, data descriptive and discussion on the best fitted mixed univariate distribution.

v. **Chapter 5** presents the effect of zero measurements by using bivariate mixed Lognormal distribution.

vi. **Chapter 6** presents the best fitted bivariate Copula distribution.

vii. **Chapter 7** concludes the findings of this study. The contributions and limitations of this study are also discussed. Lastly, suggestions and recommendations for future research work are proposed.

1.7 Conclusion

Rainfall data may be formed by zero and non-zero values, where zero values indicate no rain, while non-zero values indicate the rainy days. However, in many rainfall amount analyses, the zero values are exempted from the analyses; otherwise, in rainfall occurrence analyses, the zero and non-zero values are fitted separately under different distributions. Mixed distribution can be applied to analyse the rainfall data without excluding or separating the zero values from the rainfall data. Due to the characteristic of bivariate mixed distribution, the importance of zero values for rainfall data of two rain gauge stations can be examined. Bivariate distribution requires the marginal to be of the same distribution, but the rainfall behaviour of two stations can be different due to spatial conditions. The requirements of using a bivariate Copula distribution are more flexible than traditional bivariate distribution. Therefore, Copula distribution is the best alternative model in finding the fitted model between two rain gauge stations. By using the estimated parameters, the missing rainfall data could be estimated under the spatial weighting method. The method suggested in this thesis can be applied in hydrology, climate change, and water resource management.

REFERENCES

Adiku, S. G. K. et al. (1997) ‘An analysis of the within-season rainfall characteristics and simulation of the daily rainfall in two savanna zones in Ghana’, Agricultural and Forest Meteorology, 86(1–2), pp. 51–62. doi: 10.1016/S0168-1923(96)02414-8.

Akaike, H. (1974) ‘A new look at the statistical model identification’, IEEE Transactions on Automatic Control, 19(6), pp. 716–723. doi: 10.1109/TAC.1974.1100705.

Aksoy, H. (2000) ‘Use of Gamma distribution in hydrological analysis’, Turkish Journal of Engineering and Environmental Sciences, 24(6), pp. 419–428.

Al-suhili, P. R. H. and Khanbilvardi, P. R. (2014) ‘Frequency analysis of the monthly rainfall data at Sulaimania’, American Journal of Engineering Research (AJER), (05), pp. 212–222.

Alexandersson, H. (1986) ‘A homogeneity test applied to precipitation’, 6, pp. 661–675.

Aranda, J. A. and Garc, R. (2018) ‘Synthetic Hydrographs Generation Downstream of a River Junction Using a Copula Approach for Hydrological Risk Assessment in Large Dams’. doi: 10.3390/w10111570.

Ariff, N. M. et al. (2012) ‘IDF relationships using bivariate Copula for storm events in Peninsular Malaysia’, Journal of Hydrology. Elsevier B.V., 470–471, pp. 158–171. doi: 10.1016/j.jhydrol.2012.08.045.

Babu, G. J. (2011) ‘Resampling methods for model fitting and model selection’, Journal of Biopharmaceutical Statistics, 21(6), pp. 1177–1186. doi: 10.1080/10543406.2011.607749.

Bárdossy, A. (2006) ‘Copula-based geostatistical models for groundwater quality parameters’, Water Resources Research, 42(11), pp. 1–12. doi: 10.1029/2005WR004754.

Bermudez, V. A. B. et al. (2017) ‘Probability Distribution of Philippine daily rainfall data’, (December), pp. 1–15. doi: 10.20944/preprints201712.0150.v1.

Buishand, T. A. (1978) ‘Some remarks on the use of daily rainfall models’, Journal of Hydrology, 36(3–4), pp. 295–308. doi: 10.1016/0022-1694(78)90150-6.

Buishand, T. A. (1982) ‘Some methods for testing the homogeneity of rainfall records’, Journal of Hydrology, 58(1), pp. 11–27. doi: https://doi.org/10.1016/0022-1694(82)90066-X.

Chapman, T. G. (1997) ‘Stochastic models for daily rainfall in the Western Pacific’, Mathematics and Computers in Simulation, 43(3–6), pp. 351–358. doi: 10.1016/S0378-4754(97)00019-0.

Chen, X. and Fan, Y. (2006) ‘Estimation of Copula-based semiparametric time series models’, Journal of Econometrics, 130(2), pp. 307–335. doi: 10.1016/j.jeconom.2005.03.004.

Chowdhury, A. F. M. K. et al. (2017) ‘Development and evaluation of a stochastic daily rainfall model with long term variability’, Hydrology and Earth System Sciences Discussions, pp. 1–30. doi: 10.5194/hess-2017-84.

Deni, S. M., Jemain, A. A. and Ibrahim, K. (2009) ‘Fitting optimum order of Markov chain models for daily rainfall occurrences in Peninsular Malaysia’, Theoretical and Applied Climatology, 97(1–2), pp. 109–121. doi: 10.1007/s00704-008-0051-3.

Dunn, P. K. (2004) ‘Occurrence and quantity of precipitation can be modelled simultaneously’, International Journal of Climatology, 24(10), pp. 1231–1239. doi: 10.1002/joc.1063.

Durante, F., Quesada-Molina, J. J. and Sempi, C. (2007) ‘A Generalization of the Archimedean Class of Bivariate Copulas’, Annals of the Institute of Statistical Mathematics, 59(3), pp. 487–498. doi: 10.1007/s10463-006-0061-9.

Dzupire, N. C., Ngare, P. and Odongo, L. (2018) ‘A Poisson-Gamma model for zero inflated rainfall data’, Journal of Probability and Statistics, 2018, pp. 1–12. doi: 10.1155/2018/1012647.

Embrechts, P., Lindskog, F. and Mcneil, A. (2003) ‘Modelling dependence with Copulas and applications to risk management’, Handbook of Heavy Tailed Distributions in Finance, pp. 329–384. doi: 10.1016/B978-044450896-6.50010-8.

Fadhilah, N., Radi, A. and Zakaria, R. (2017) ‘Generating monthly rainfall amount using multivariate Skew-t Copula Generating monthly rainfall amount using multivariate Skew- t Copula’.

Fang, Y., Madsen, L. and Liu, L. (2014) ‘Comparison of two methods to check Copula fitting’, IAENG International Journal of Applied Mathematics, 44(1), pp. 53–61.

Favre, A. C. et al. (2004) 'Multivariate hydrological frequency analysis using Copulas', Water Resources Research, 40(1), pp. 1–12. doi: 10.1029/2003WR002456.

Frees, E. W. and Wang, P. (2005) 'Credibility using Copulas', North American Actuarial Journal, 9(2), pp. 31–48. doi: 10.1080/10920277.2005.10596196.

Gabriel, K. R. (1959) 'The distribution of the number of successes in a sequence of dependent trials', Biometrika, 46(1), pp. 454–460. doi: 10.2307/2333542.

Galambos, J. (1975) 'Order statistics of samples from multivariate distributions', Journal of the American Statistical Association. [American Statistical Association, Taylor & Francis, Ltd.], 70(351), pp. 674–680. doi: 10.2307/2285954.

Genest, C. and Favre, A.-C. (2007) 'Everything you always wanted to know about Copula modeling but were afraid to ask', Journal of Hydrologic Engineering, 12(4), pp. 347–368. doi: 10.1061/(ASCE)1084-0699(2007)12:4(347).

Genest, C., Ghoudi, K. and Rivest, L. P. (1995) 'of dependence procedure A semiparametric estimation families of distributions in multivariate parameters', Biometrika, 82(3), pp. 543–552.

Genest, C. and Mackay, J. (1986) 'The joy of Copulas: Bivariate distributions with uniform marginals', American Statistician, 40(4), pp. 280–283. doi: 10.1080/00031305.1986.10475414.

González, J. and Valdés, J. B. (2003) 'Bivariate drought recurrence analysis using tree ring reconstructions', Journal of Hydrologic Engineering, 8(October), pp. 247–258. doi: 10.1061/(ASCE)1084-0699(2003)8:5(247).

Grimaldi, S. and Serinaldi, F. (2006) 'Asymmetric Copula in multivariate flood frequency analysis', Advances in Water Resources, 29(8), pp. 1155–1167. doi: 10.1016/j.advwatres.2005.09.005.

Gumbel, E. J. (1954) 'The maxima of the mean largest value and of the range', The Annals of Mathematical Statistics, 25(1), pp. 76–84. doi: 10.1214/aoms/1177728847.

Ha, E. and Yoo, C. (2007) 'Use of mixed bivariate distributions for deriving inter-station correlation coefficients of rain rate', Hydrological Processes, 21(22), pp. 3078–3086. doi: 10.1002/hyp.6526.

Hamdi, Y., Chebana, F. and Quarda, T. B. M. J. (2016) 'Bivariate drought frequency analysis in the Medjerda River Basin, Tunisia’, Journal of Civil & Environmental Engineering, 6(3). doi: 10.4172/2165-784x.1000227.

Herr, H. D. and Krzysztofowicz, R. (2005) ‘Generic probability distribution of rainfall in space : the bivariate model’, 306, pp. 234–263. doi: 10.1016/j.jhydrol.2004.09.011.

Hoogmoed, W. B. and Klaij, M. C. (1990) ‘Soil management for crop production in the West African Sahel. I. Soil and climate parameters’, Soil and Tillage Research, 16(1–2), pp. 85–103. doi: 10.1016/0167-1987(90)90023-7.

Jarušková, D. (1996) ‘Change-point detection in meteorological measurement’, Monthly Weather Review. American Meteorological Society, 124(7), pp. 1535–1543. doi: 10.1175/1520-0493(1996)124<1535:CPDIMM>2.0.CO;2.

Joe, H. (1997) Multivariate models and dependence concepts. Edited by and H. T. F. Bunea, V. Isham, N. Keiding, T. Louis, R. L. Smith. New York: Taylor & Francis Group. Available at: https://content.taylorfrancis.com/books/download?dac=C2010-0-32529-X&isbn=9781466581432&format=googlePreviewPdf.

Kao, S.-C. and Govindaraju, R. S. (2007) ‘A bivariate frequency analysis of extreme rainfall with implications for design’, Journal of Geophysical Research: Atmospheres, 112(D13), p. n/a-n/a. doi: 10.1029/2007JD008522.

Kao, S.-C. and Govindaraju, R. S. (2010) ‘A Copula-based joint deficit index for droughts’, Journal of Hydrology. Elsevier, 380(1–2), pp. 121–134. doi: 10.1016/J.JHYDROL.2009.10.029.

Karakas, A. M. and Dogan, M. (2017) ‘Archimedean Copula parameter estimation with kendall distribution function kendall’, Iğdır University Journal of the Institute of Science and Technology, 7(3), pp. 187–198.

Karmakar, S. and Simonovic, S. P. (2009) ‘Bivariate flood frequency analysis. Part 2: A Copula-based approach with mixed marginal distributions’, Journal of Flood Risk Management, 2(1), pp. 32–44. doi: 10.1111/j.1753-318X.2009.01020.x.

Katz, R. W. (1977) ‘Precipitation as a chain-dependent process’, Journal of Applied Meteorology, pp. 671–676. doi: 10.1175/1520-0450(1977)016<0671:PAACDP>2.0.CO;2.

Kedem, B., Chiu, L. S. and Karni, Z. (1990) ‘An analysis of the threshold method for measuring area-average rainfall’, Journal of Applied Meteorology, pp. 3–20.

doi: 10.1175/1520-0450(1990)029<0003:AAOTTM>2.0.CO;2.

Kedem, B., Chiu, L. S. and North, G. R. (1990) ‘Estimation of mean rain rate: application to satellite observations’, Journal of Geophysical Research, 95(D2), pp. 1965–1972. doi: 10.1029/JD095iD02p01965.

Khudri, M. and Sadia, F. (2013) ‘Determination of the Best Fit Probability Distribution for Annual Extreme Precipitation in Bangladesh’, European Journal of Scientific Research, 103(JUNE), pp. 391–404. doi: 10.5923/j.ajms.20160604.05.

Leonard, M., Metcalfe, A. and Lambert, M. (2008) ‘Frequency analysis of rainfall and streamflow extremes accounting for seasonal and climatic partitions’, Journal of Hydrology. Elsevier, 348(1–2), pp. 135–147. doi: 10.1016/J.JHYDROL.2007.09.045.

de Lima, J. L. M. P. (1990) ‘The effect of oblique rain on inclined surfaces: A nomograph for the rain-gauge correction factor’, Journal of Hydrology, 115(1–4), pp. 407–412. doi: 10.1016/0022-1694(90)90218-M.

Lokoman, R. M. and Yusof, F. (2019) ‘Parametric estimation methods for bivariate Copula in rainfall application’, Jurnal Teknologi, 1, pp. 1–10.

De Michele, C. et al. (2005) ‘Bivariate Statistical Approach to Check Adequacy of Dam Spillway’, Journal of Hydrologic Engineering, 10(1), pp. 50–57. doi: 10.1061/(ASCE)1084-0699(2005)10:1(50).

De Michele, C. and Salvadori, G. (2003) ‘A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas’, Journal of Geophysical Research, 108(D2), p. 4067. doi: 10.1029/2002JD002534.

Mirabbasi, R., Fakheri-Fard, A. and Dinpashoh, Y. (2012) ‘Bivariate drought frequency analysis using the Copula method’, Theoretical and Applied Climatology, 108(1), pp. 191–206. doi: 10.1007/s00704-011-0524-7.

Mohamadi, M. A. and Kavian, A. (2015) ‘Effects of rainfall patterns on runoff and soil erosion in field plots’, International Soil and Water Conservation Research. Elsevier, 3(4), pp. 273–281. doi: 10.1016/j.iswcr.2015.10.001.

Mohammed, N., Munassar, A. and Govardhan, A. (2010) ‘A comparison between five models of software engineering’, IJCSI International Journal of Computer Science Issues ISSN, 7(5), pp. 1694–814.

Mujiburrehman, K. (2017) ‘Fitting hourly rainfall using various distribution and evaluation of best fit for Mumbai region’, International Journal of Research in

Engineering and Innovation, 3(3), pp. 65–69.

Myers, B. A. et al. (1998) ‘Does irrigation affect leaf phenology in deciduous and evergreen trees of the savannas of northern Australia?’, Austral Ecology, 23(4), pp. 329–339. doi: 10.1111/j.1442-9993.1998.tb00738.x.

Naifar, N. (2011) ‘Modelling dependence structure with Archimedean Copulas and applications to the iTraxx CDS index’, Journal of Computational and Applied Mathematics. Elsevier B.V., 235(8), pp. 2459–2466. doi: 10.1016/j.cam.2010.10.047.

Nazemi, A. and Elshorbagy, A. (2012) ‘Application of Copula modelling to the performance assessment of reconstructed watersheds’, Stochastic Environmental Research and Risk Assessment, 26(2), pp. 189–205. doi: 10.1007/s00477-011-0467-7.

Nelsen, R. B. (1999) An introduction to Copulas. Springer-Verlag New York. Available at: https://www.springer.com/gp/book/9780387286594.

von Neumann, J. (1941) ‘Distribution of the ratio of the mean square successive difference to the variance’, Ann. Math. Statist. The Institute of Mathematical Statistics, 12(4), pp. 367–395. doi: 10.1214/aoms/1177731677.

Nguyen, V.-T.-V. and Mayabi, A. (1991) ‘Probabilistic analysis of summer daily rainfall for the Montreal Region’, Canadian Water Resources Journal, 16(1), pp. 65–80. doi: 10.4296/cwrj1601065.

Obaid, N. et al. (2014) ‘Fitting statistical distributions of monthly rainfall for some Iraqi Stations’, 6(6), pp. 40–47.

Oseni, B. and Ayoola, F. (2012) ‘Fitting the statistical distribution for daily rainfall in Ibadan, based on Chi-Square and Kolmogorov-Smirnov Goodness-of-Fit’, West African Journal of Industrial and Academic Research, 7(1), pp. 93–100. Available at: http://www.iiste.org/Journals/index.php/EJBM/article/view/3202.

Pakoksung, K. and Takagi, M. (2017) ‘Mixed zero-inflation method and probability distribution in fitting daily rainfall data’, Engineering Journal, 21(2), pp. 63–80. doi: 10.4186/ej.2017.21.2.63.

Papaioannou, G. et al. (2016) ‘Joint modelling of flood peaks and volumes : A Copula application for the Danube River’, (2006), pp. 382–392. doi: 10.1515/johh-2016-0049.

Park, M. et al. (2014) Bivariate frequency analysis of annual maximum rainfall event series in Seoul, Korea, Journal of Hydrologic Engineering. doi:

10.1061/(ASCE)HE.1943-5584.0000891.

Parzen, E. (1962) 'On estimation of a probability density function and mode', Ann. Math. Statist. The Institute of Mathematical Statistics, 33(3), pp. 1065–1076. doi: 10.1214/aoms/1177704472.

Peterson, T. C. et al. (1998) 'Homogeneity adjustments of in situ atmospheric climate data : A review', 1517, pp. 1493–1517.

Pettitt, A. N. (1979) 'A Non-parametric approach to the change-point problem', Journal of the Royal Statistical Society. Series C (Applied Statistics). [Wiley, Royal Statistical Society], 28(2), pp. 126–135. doi: 10.2307/2346729.

Pickands, J. (1981). 'Multivariate extreme value distributions (with a discussion).'InProceedings ofthe 43rd Session of the International Statistical Institute.Bull. Inst. Internat. Statist.49859–878,894–902.

Plackett, R. L. (1965) 'A class of bivariate distributions', Journal of the American Statistical Association, 60(310), pp. 516–522. doi: 10.1080/01621459.1965.10480807.

Poulin, A. et al. (2007) 'Importance of tail dependence in bivariate frequency analysis', Journal of Hydrologic Engineering, 12(4), pp. 394–403. doi: 10.1061/(ASCE)1084-0699(2007)12:4(394).

Qiao, H. and Tsokos, C. P. (1995) 'Estimation of the three parameter Weibull probability distribution', Mathematics and Computers in Simulation, 39(1), pp. 173–185. doi: https://doi.org/10.1016/0378-4754(95)95213-5.

Quinn, C. (2007) 'The health-economic applications of Copulas : methods in applied econometric research', October, (October).

Ran, Q. et al. (2012) 'Experimental study of the impact of rainfall characteristics on runoff generation and soil erosion', Journal of Hydrology. Elsevier B.V., 424–425, pp. 99–111. doi: 10.1016/j.jhydrol.2011.12.035.

Ray, L. I. P. et al. (2015) 'Rainfall characteristics, pattern and distribution at Cherapunjee, Meghalaya', Indian Journal of Hill Farming, 28, pp. 23–26.

Reddy, M. J. and Ganguli, P. (2012) 'Bivariate flood frequency analysis of upper Godavari river flows using Archimedean Copulas', Water Resources Management, 26(14), pp. 3995–4018. doi: 10.1007/s11269-012-0124-z.

Sadiq, N. (2014) 'Stochastic modelling of the daily rainfall frequency and amount', Arabian Journal for Science and Engineering, 39(7), pp. 5691–5702. doi: 10.1007/s13369-014-1132-5.

Salvadori, G. and De Michele, C. (2004) ‘Frequency analysis via Copulas: Theoretical aspects and applications to hydrological events’, Water Resources Research, 40(12), pp. 1–17. doi: 10.1029/2004WR003133.

Schweizer, B. and Wolff, E. F. (1981) ‘On nonparametric measures of dependence for random variables’, The Annals of Statistics. Institute of Mathematical Statistics, 9(4), pp. 879–885. Available at: http://www.jstor.org/stable/2240856.

Serinaldi, F. (2009a) ‘A multisite daily rainfall generator driven by bivariate Copula-based mixed distributions’, Journal of Geophysical Research Atmospheres, 114(10), pp. 1–13. doi: 10.1029/2008JD011258.

Serinaldi, F. (2009b) ‘Copula-based mixed models for bivariate rainfall data : an empirical study in regression perspective’, pp. 677–693. doi: 10.1007/s00477-008-0249-z.

Serinaldi, F. and Grimaldi, S. (2007) ‘Fully nested 3-Copula: Procedure and application on hydrological data’, Journal of Hydrologic Engineering, 12(4), pp. 420–430. doi: 10.1061/(ASCE)1084-0699(2007)12:4(420).

Sharda, V. N. and Das, P. K. (2005) ‘Modelling weekly rainfall data for crop planning in a sub-humid climate of India’, Agricultural Water Management, 76(2), pp. 120–138. doi: 10.1016/j.agwat.2005.01.010.

Sharma, M. A. and Singh, J. B. (2010) ‘Use of probability distribution in rainfall analysis’, New York Science Journal, 3(9), pp. 40–49.

Shiau, J. T. (2006) ‘Fitting drought duration and severity with two-dimensional Copulas’, Water Resources Management, 20(5), pp. 795–815. doi: 10.1007/s11269-005-9008-9.

Shimizu, K. (1993) ‘A bivariate mixed Lognormal distribution with an analysis of rainfall data’, Journal of Applied Meteorology, 32(2), pp. 161–171. doi: 10.1175/1520-0450(1993)032<0161:ABMLDW>2.0.CO;2.

Shimizu, K. and Sagae, M. (1990) ‘Modelling with bivariate an analysis data of daily containing rainfall zeroes , Data’, 19(1), pp. 19–31.

Singh, K. and Singh, V. P. (1991) ‘Derivation of bivariate probability density functions with Exponential marginals’, Stochastic Hydrology and Hydraulics, 5(1), pp. 55–68. doi: 10.1007/BF01544178.

Sklar, A. (1959) Fonctions de répartition à n dimensions et leurs marges. Publications de l’Institut Statistique de l’Université de Paris.

Sraj, M., Bezak, N. and Brilly, M. (2015) ‘Bivariate flood frequency analysis using the Copula function: A case study of the Litija station on the Sava River', Hydrological Processes, 29(2), pp. 225–238. doi: 10.1002/hyp.10145.

Srikanthan, R. and McMahon, T. A. (2001) 'Stochastic generation of annual, monthly and daily climate data: A review', Hydrology and Earth System Sciences, 5(4), pp. 653–670. doi: 10.5194/hess-5-653-2001.

Stern, R. D. and Coe, R. (1984) 'A model fitting analysis of daily rainfall data Published by : Blackwell Publishing for the Royal Statistical Society Stable URL : http://www.jstor.org/stable/2981736', 147(1), pp. 1–34. Available at: https://www.jstor.org/stable/pdf/2981736.pdf.

Stern, R. D., Dennett, M. D. and Dale, I. C. (1982) 'Analysing daily rainfall measurements to give agronomically useful results . I . direct methods' Departments of Applied Statistics and Agricultural Botany , University of Reading , England.

Suhaila, J., Deni, S. and Jemain, A. (2008) 'Revised spatial weighting methods for estimation of missing rainfall data', Asia-Pacific Journal of Atmospheric Sciences, 44, pp. 93–104.

Suhaila, J. and Jemain, A. A. (2007) 'Fitting daily rainfall amount in Malaysia using the Normal transform distribution', Journal of Applied Sciences, 7(14), pp. 1880–1886. Available at: https://ukm.pure.elsevier.com/en/publications/fitting-daily-rainfall-amount-in-malaysia-using-the-normal-transf.

Suhaila, J. and Jemain, A. A. (2008) 'Fitting the statistical distribution to the daily rainfall amount in Peninsular Malaysia based on AIC criterion', Journal of Applied Sciences Research, 4(12), pp. 1846–1857.

Suhaila, J. and Jemain, A. A. (2010) Analisis amaun hujan di Semenanjung Malaysia. Universiti Kebangsaan Malaysia. Available at: https://books.google.com.my/books?id=6dqJnQAACAAJ.

Tettey, M. et al. (2017) 'Markov chain analysis of the rainfall patterns of five geographical locations in the south eastern coast of Ghana', Earth Perspectives. Earth Perspectives, 4(1), p. 6. doi: 10.1186/s40322-017-0042-6.

Thanh, N. T. (2017) 'Fitting a probability distribution to extreme precipitation for a limited mountain area in Vietnam', Journal of Geoscience and Environment Protection, 05(05), pp. 92–107. doi: 10.4236/gep.2017.55007.

Todorovic, P. and Woolhiser, D. A. (1975) 'A stochastic model of n-day precipitation', Journal of Applied Meteorology. American Meteorological

Society, 14(1), pp. 17–24. doi: 10.1175/1520-0450(1975)014<0017:ASMODP>2.0.CO;2.

Utusan, M. (2004) Banjir di Johor makin teruk 7,661 mangsa dipindah. Available at: http://www.dartmouth.edu/~floods/Archives/2004sum.htm (Accessed: 30 August 2018).

Varikoden, H. et al. (2011) 'Seasonal variation of rainfall characteristics in different intensity classes over Peninsular Malaysia', Journal of Hydrology. Elsevier, 404(1–2), pp. 99–108. doi: 10.1016/J.JHYDROL.2011.04.021.

Verbist, K. et al. (2010) 'Seasonal predictability of daily rainfall characteristics in central northern Chile for dry-land management', Journal of Applied Meteorology and Climatology, 49(9), pp. 1938–1955. doi: 10.1175/2010JAMC2372.1.

Wijngaard, J. B., Klein Tank, A. M. G. and Können, G. P. (2003) 'Homogeneity of 20th century European daily temperature and precipitation series', International Journal of Climatology, 23(6), pp. 679–692. doi: 10.1002/joc.906.

Wilks, D. S. (2011) Statistical methods in the atmospheric sciences, International Geophysics Series. Available at: http://books.google.com/books?hl=en&lr=&id=IJuCVtQ0ySIC&oi=fnd&pg=PP2&dq=statistical+methods+in+the+atmospheric+sciences&ots=amHkqWwJHY&sig=90PZrnx5n8UXY95KClq4zdUZSLQ.

Willmott, C. J. (1981) 'On the validation of models', Physical Geography. Taylor & Francis, 2(2), pp. 184–194. doi: 10.1080/02723646.1981.10642213.

Wong, C. L. et al. (2016) 'Rainfall characteristics and regionalization in Peninsular Malaysia based on a high resolution gridded data set', Water. Elsevier B.V., 8(11), pp. 99–108. doi: 10.1016/j.jhydrol.2011.04.021.

Yan, J. (2007) 'Enjoy the joy of Copulas: With a package Copula', Journal of Statistical Software, 21(4). doi: 10.18637/jss.v021.i04.

Yemenu, F. and Chemeda, D. (2013) 'Dry and wet spell analysis of the two rainy seasons for decision support in agricultural water management for crop production in the Central Highlands of Ethiopia', Journal of Biology, Agriculture and Healthcare, 3(11), pp. 1–6. Available at: http://iiste.org/Journals/index.php/JBAH/article/view/7247.

Yoo, C. and Ha, E. (2007) 'Effect of zero measurements on the spatial correlation structure of rainfall', Stochastic Environmental Research and Risk Assessment, 21(3), pp. 287–297. doi: 10.1007/s00477-006-0064-3.

Yoo, C., Jung, K.-S. and Kim, T.-W. (2005) 'Rainfall frequency analysis using a mixed Gamma distribution: evaluation of the global warming effect on daily rainfall', Hydrological Processes, 19(19), pp. 3851–3861. doi: 10.1002/hyp.5985.

Yoo, C., Lee, J. and Ro, Y. (2016) 'Markov Chain decomposition of monthly rainfall into daily rainfall: Evaluation of climate change impact', Advances in Meteorology. Elsevier, 2016, pp. 1–10. doi: 10.1155/2016/7957490.

Yue, S. (2000) 'The Gumbel logistic model for representing a multivariate storm event', Advances in Water Resources, 24(2), pp. 179–185. doi: 10.1016/S0309-1708(00)00039-7.

Yue, S. (2002) 'The bivariate Lognormal distribution for describing joint statistical properties of a multivariate storm event', Environmetrics, 13(8), pp. 811–819. doi: 10.1002/env.483.

Yusof, F. et al. (2013) 'Characterisation of drought properties with bivariate Copula analysis', Water Resources Management, 27(12), pp. 4183–4207. doi: 10.1007/s11269-013-0402-4.

Yusuf, A. U. (2014) 'Markov Chain model and its application to annual rainfall distribution for crop production', American Journal of Theoretical and Applied Statistics, 3(2), p. 39. doi: 10.11648/j.ajtas.20140302.12.

Zakaria, R. et al. (2010) 'Using the Skew-t Copula to model bivariate rainfall distribution', 51, pp. 231–246.

Zhang, L. and Singh, V. P. (2007) 'Bivariate rainfall frequency distributions using Archimedean Copulas', Journal of Hydrology. Elsevier, 332(1–2), pp. 93–109. doi: 10.1016/J.JHYDROL.2006.06.033.

Zhang, L., Singh, V. P. and Asce, F. (2006) 'Using the Copula method', Water, 11(April), pp. 150–164. doi: 10.1061/(ASCE)1084-0699(2006)11:2(150).