AN INVESTIGATION INTO RAINFALL CHARACTERISTICS IN PENINSULAR MALAYSIA USING COPULAS METHOD

MOHD FARIS BIN FAUZI

MASTER OF SCIENCE (STATISTICS)

UNIVERSITI PERTAHANAN NASIONAL MALAYSIA

2021

AN INVESTIGATION INTO RAINFALL CHARACTERISTICS IN PENINSULAR MALAYSIA USING COPULAS METHOD

MOHD FARIS BIN FAUZI

Thesis submitted to the Centre for Graduate Studies, Universiti Pertahanan Nasional Malaysia, in fulfilment of the requirements for the Degree of Master of Science (Statistics)

2021

ABSTRACT

Extreme rain patterns are influenced by variations in daily rain. Floods, landslides, and crop failures are all caused by excessive rain. Physical damage to residential and other structures can be caused by floods. As a result, efforts must be made to mitigate the impact of loss caused by heavy rain. Estimating the huge damage to the floods arising from excessive rainfall is one of the efforts to limit this effect. Flood loss modeling can be used to estimate how much damage will occur when floods occur due to heavy rain. For modeling, extreme value theory and copula methods are used. Copula is a function that combines the multivariate distribution function to the marginal distribution function. The Extreme Value Theory focuses on modeling of the tail behaviour of a distribution using only extreme values rather than the whole dataset, it can potentially provide a more accurate estimate of tail risk. Based on the modeling results, from 30 stations, Frank Copula recorded 17 stations, so we chose Frank Copula is the best Copula model to explain the relationship between severity and duration of rainfall in Peninsular Malaysia. When the parameter value of both low or strong variables, Copula Frank explains close relationships.

ABSTRAK

Corak hujan yang melampau dipengaruhi oleh variasi hujan harian. Banjir, tanah runtuh, dan kegagalan tanaman semuanya disebabkan oleh hujan yang berlebihan. Kerosakan fizikal pada kediaman dan struktur lain boleh disebabkan oleh banjir. Akibatnya, usaha mesti dilakukan untuk mengurangkan kesan kerugian yang disebabkan oleh hujan lebat. Menganggarkan kerosakan besar banjir akibat hujan lebat adalah salah satu usaha untuk membatasi kesan ini. Pemodelan kerugian banjir dapat digunakan untuk menganggarkan berapa banyak kerosakan yang akan terjadi ketika banjir terjadi akibat hujan lebat. Untuk pemodelan, teori nilai ekstrem dan kaedah kopula digunakan. Copula adalah fungsi yang menggabungkan fungsi taburan multivariate dengan fungsi taburan marginal. Teori Nilai Ekstrim memfokuskan pada pemodelan tingkah laku ekor pengedaran menggunakan hanya nilai ekstrem dan bukannya keseluruhan kumpulan data, ia berpotensi memberikan anggaran risiko ekor yang lebih tepat. Berdasarkan hasil pemodelan, dari 30 stesen, Frank Copula mencatatkan 17 stesen, jadi kami memilih Frank Copula adalah model Copula terbaik untuk menjelaskan hubungan antara keparahan dan tempoh hujan di Semenanjung Malaysia. Apabila nilai parameter kedua-dua pemboleh ubah rendah atau kuat, Copula Frank menerangkan hubungan rapat.

ACKNOWLEDGEMENT

First and foremost, I'd like to express my gratitude to my thesis advisor, Associate Professor Dr Ummul Fahri Binti Abdul Rauf of the Universiti Pertahanan Nasional Malaysia's Centre for Defence Foundation Studies. When I ran into a problem or had a question about my research or writing, the door to my advisor's office was always open. She consistently let me do my work on this paper, but she always pointed me in the right direction when she thought I needed it.

I'd also like to thank Associate Professor Dr Wan Rozita Binti Wan Din and Dr Aisyah Fataf, who participated in the validation survey for this research project. The validation survey would not have been possible without their enthusiastic participation and input.

Finally, I want to express my heartfelt gratitude to my parents for their unwavering support and encouragement throughout my years of study, as well as during the research and writing of this thesis. Without them, this feat would not have been possible. Thank you very much.

APPROVAL

The Examination Committee has met on **15th September 2021** to conduct the final examination of **Mohd Faris bin Fauzi** on his degree thesis entitled **'An Investigation into Rainfall Characteristics in Peninsular Malaysia Using Copulas Method'**.

The committee recommends that the student be awarded the Master of Science (Statistics). Members of the Examination Committee were as follows.

Lt. Cdr. Associate Professor Ts. Dr. Mohd Norsyarizad bin Razali TLDM (Bersara)
Dean, Faculty of Sciences and Defence Technology
National Defence University of Malaysia
(Chairperson)

Dr. Mohd Syazwan bin Mohamad Anuar
Centre for Defence Foundation Studies
National Defence University of Malaysia
(Internal Examiner)

Associate Professor Dr Firdaus bin Mohd Hamzah
Head of Program, Faculty of Engineering,
National University of Malaysia (UKM)
(External Examiner)

APPROVAL

This thesis was submitted to the Senate of Universiti Pertahanan Nasional Malaysia and has been accepted as fulfilment of the requirements for the degree of **Master of Science (Statistics)**. The members of the Supervisory Committee were as follows.

Associate Professor Dr. Ummul Fahri binti Abdul Rauf
Centre for Defence Foundation Studies
National Defence University of Malaysia
(Main Supervisor)

Associate Professor Dr. Wan Rozita binti Wan Din
Centre for Defence Foundation Studies
National Defence University of Malaysia
(Co-Supervisor)

UNIVERSITI PERTAHANAN NASIONAL MALAYSIA

DECLARATION OF THESIS

Student's full name : MOHD FARIS BIN FAUZI

Date of birth : 20 NOVEMBER 1991

Title : AN INVESTIGATION INTO RAINFALL CHARACTERISTICS IN PENINSULAR MALAYSIA USING COPULAS METHOD

Academic session : 9

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

I further declare that this thesis is classified as:

☐ **CONFIDENTIAL** (Contains confidential information under the official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*

☐ **OPEN ACCESS** I agree that my thesis to be published as online open access (full text)

I acknowledge that Universiti Pertahanan Nasional Malaysia reserves the right as follows.

1. The thesis is the property of Universiti Pertahanan Nasional Malaysia.
2. The library of Universiti Pertahanan Nasional Malaysia has the right to make copies for the purpose of research only.
3. The library has the right to make copies of the thesis for academic exchange.

Prof. Madya Dr. Ummul Fahri Binti Abdul Rauf
Jabatan Matematik
Pusat Asasi Pertahanan
Universiti Pertahanan Nasional Malaysia

Signature | **Signature of Supervisor/Dean of CGS/ Chief Librarian

911120145675 | Assoc. Prof. Dr. Ummul Fahri Abdul Rauf

IC/Passport No. | **Name of Supervisor/Dean of CGS/ Chief Librarian

Date: 13/10/2021 | Date: 13 October 2021

*If the thesis is CONFIDENTAL OR RESTRICTED, please attach the letter from the organisation with period and reasons for confidentiality and restriction.
** Witness

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Research Background

Apart to landslides, forest fires, oil spills and so on, floods are also a natural disaster that often hit the country every year that leaving behind a devastating impact especially on the economy, environment, and human life. Quite a considerable number of properties were destroyed or damaged due to the flood. In Malaysia, natural disasters such as floods are categorised into flash floods and monsoon floods. The climate change phenomenon also contributes to the uncertainty of the intensity and frequency of a weather phenomenon that can record extreme conditions.

Several factors contribute to the occurrence of this catastrophic flood, including heavy rain, high tide, and water flow barriers in the drainage system. Floods also occur due to the tremendous water flows caused by rain or mud floods due to sudden changes in land use. The duration of a flood event is either long or fast for floodwater to recede depending on the condition of a river or the surface of the area. It can take several hours or even lasts for days.

The floods that hit the states across Malaysia, especially on the East Coast that took place in late 2014, were recorded as the worst floods since 1969. In general, Malaysia is often hit by floods. In 2014, all the states in Malaysia are experiencing floods from 2014 to February 2015. Selangor recorded 92 flood events that took place and made the country most affected by floods that year, followed by the state of Perak with 52 flood events and WP Labuan (Borneo Malaysia) is the least recorded state of a flood event with a record of 3 occurrences during the period. At the end of 2014 and early 2015 (14 December 2014 - 10 January 2015), the country has been shocked by the unexpected northeast monsoon incidents over 15 days involving six states in Peninsular Malaysia namely Kelantan, Terengganu, Pahang, Perak, Perlis and Johor and two states in Borneo Malaysia namely Sabah and Sarawak. The number of victims who had to be transferred was more than 300,000 peoples. Kelantan faces the worst situation when the floods bring similar muds to over 1,500 houses while the whole country is paralysed.

1.2 Problem Statement

Floods in Peninsula Malaysia are attracting more research attention. Floods usually occur when a region of the country receives heavy precipitation for an extended period of time. As a result of the urgent need for water resource management, natural disaster protection plans emphasise rainfall modelling studies. Due to the unpredictable climatic conditions in many parts of the world, rainfall statistical analysis is becoming increasingly important. However, traditional methods of building multivariate distribution models have some limitations, particularly in hydrology. This is because hydrological variables are highly dependent and do not always meet traditional modelling assumptions. Copulas are one of the best tools for helping statisticians build a joint distribution of multivariate random variables with varying types of marginal distributions.

In statistical analysis, it is usually necessary to assume that the random variable is normally distributed. In fact, often also found data that is not normally distributed. This resulted in some problems during statistical testing. In the multivariate case, the most common method of analysis is assuming that the random variable is the multivariate normal distribution.

In research in the field of meteorology or climatology is often encountered abnormalities in the data. For example, rainfall data, temperature, humidity, air pressure, wind speed. According to Schölzel (2008), climatic data often produce non-Gaussian, e.g., rainfall, wind speed, cloud coverage, or relative humidity, which has a finite or skew distribution.

The normalised assumption of distribution in the method of statistical analysis is very necessary because to facilitate in the calculation of estimation methods. For example, Pearson correlation is used to find the relationship between two variables that are assumed to be a normal distribution. Whereas in many climate data that has no normal distribution, so it takes an approach method that can explain the relationship between 2 variables that do not have a normal distribution.

1.3 Objectives

The objective of this research is to model bivariate joint distributions of precipitation severity and duration in Peninsular Malaysia, and then apply the findings to various aspects of recent flooding in Peninsular Malaysia. Several studies estimate flood return periods based on merely on flood characteristics such as duration and severity. However, these characteristics commonly are correlated with those that may not be accurate with the results. To achieve the following research objectives, this thesis addresses the following research goals:

1.3.1 Objective 1

Analyse the dependence measure

The first objective of this research is to estimate the correlation coefficient of the two variables with 3 types of correlation, namely Pearson, Spearman and Kendall. To identify factors that influence flooding, a statistical analysis using these three correlations was applied to rainfall data. In addition, this correlation coefficient is calculated with the target to determine the relationship between these two variables, rainfall duration and rainfall severity.

1.3.2 Objective 2

Monitor and comparing the extreme wet condition at different time scale

The second objective of this research was to identify extreme wet conditions until the flood occurred using the Standard Precipitation Index (SPI). The SPI index is used in hydrology. Comparison is made using different time scales, short- and long-term scales, where short-term scales represent 3-month and 6-month time scales and for the long-term represent 24 months and 48-months.

1.3.3 Objective 3

Finding the best of copula

Before implement to fit the copula, the marginal distribution will be discuss to find the best distribution for every catchment using the Akaike Information Criterion (AIC). In this study, the family of copula, Archimedean and Extreme Value copula will be use in this study to determine the best dependency structure model on copula by maximum loglikelihood and calculation of parameters, θ.

1.3.4 Objective 4:

Analyse tail dependence

In analysis of tail dependence, the bivariate copulas are used to estimate the degree of tail dependence displayed by rainfall characteristics. Another possibility is that tail dependence helps explain the extent of dependence in the tail of a bivariate distribution.

1.4 Research Scope

Peninsular Malaysia (Figure 1) is located in the western part of Malaysia with an area of 130,590 square kilometres (50,420 square meters). Peninsular Malaysia is divided into two, east and west coast (Britannica, T, 2011). The states on the East Coast consist of Pahang, Terengganu and Kelantan states and are facing the south China sea. As for the west coast, it meets the Melaka Straits, the states on the West Coast are divided into three regions: northern, southern, and central. In the north region, the states are Perak, Kedah, Penang and Perlis, while in the south region are the states of Melaka, Johor and Negeri Sembilan and in the central region is Selangor. Malaysia has 189 water basins and the average rainfall in a year is over 2000-4000mm. Malaysia is exposed to riverine, mud floods that occur from long hours flash floods to prolonged flooding, lowland along the main river (Haynes, K., 2011).

There are several factors that are emphasized in the selection of rain-gauge stations in this study and they are closely related to the objectives to be achieved. In determining the severity of rainfall through SPI estimation, we need to understand what SPI is, where there are limitations in the use of SPI, where long term (30-50 + years) is required. The reason for this is that not all stations have a long enough period of time to be considered, so stations with a long enough period are chosen, but rainfall records (severity) in the area are also taken into account.

Thirty rain stations spread across Peninsular Malaysia (except for the state of Melaka and the Federal Territory) with a registration period of 45 years were included in the study.

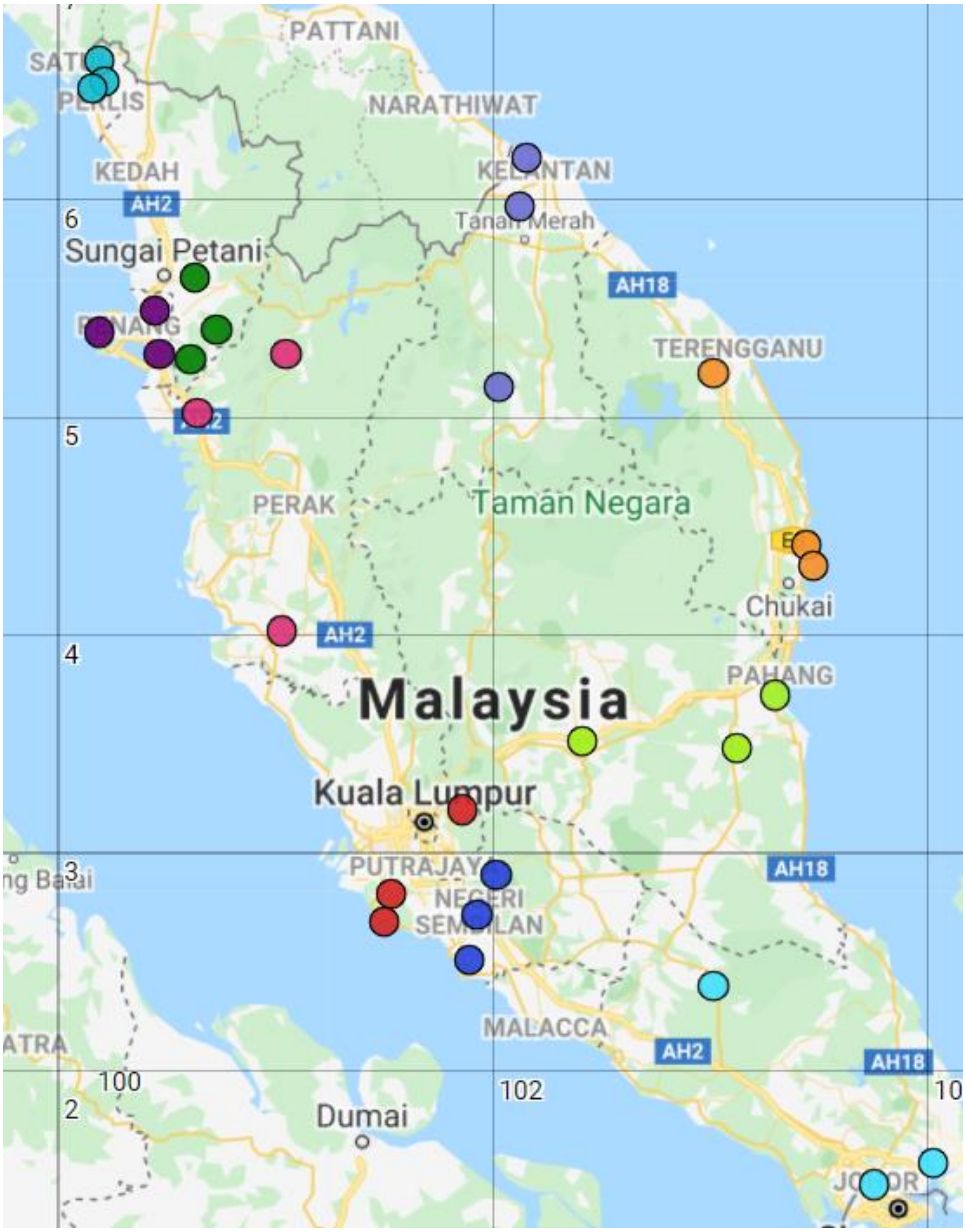

Figure 1 Study areas: Peninsular Malaysia

In Peninsular Malaysia, in December, humid weather conditions with the expected amount of rainfall received more than 400mm (Shafie, A.,2009). Generally, throughout 2014 to February 2015, all the states in Malaysia flooded. The State most often affected is Selangor, with the highest number of flood events were recorded and followed by Perak. From December 14, 2014, until January 10, 2015 an incident flood northeast monsoon tremendous hit the country involving six (6) states in Peninsular Malaysia namely Kelantan, Terengganu, Pahang, Perak, Perlis, Johor and two (2) states in Borneo Malaysia, Sabah and Sarawak which resulted in the flooding of more than 15 days. The location of rain gauge stations can be seen in Table 1 and Figure 1.

The study area for this research work involves eleven states in Peninsular Malaysia, these includes state of Selangor, Melaka, Negeri Sembilan, Johor, Perak, Pulau Pinang, Kedah, Perlis, Kelantan, Pahang and Terengganu. The South China Sea separates Malaysia into two similarly sized regions, Peninsular Malaysia (West Malaysia) with 130,598 square kilometres (50,424 sq. mi) area and Malaysian Borneo, also known as East Malaysia.

Table 1 Summary selected stations in Peninsular Malaysia

Station Code	Station Number	Station Name	States	Latitude	Longitude
JHR-1	1437116	Stor JPS Johor Bahru	Johor	01°28'N	103°45'E
JHR-2	2330009	Ldg. Sg. Labis	Johor	02°23'N	103°01'E
JHR-3	1540135	Ldg. Telok Sengat	Johor	01°34'N	104°02'E
NS-1	2719043	Astana Hinggap	Negeri Sembilan	02°43'N	101°56'E
NS-2	2519046	Ladang Sua Betong	Negeri Sembilan	02°30'N	101°54'E
NS-3	2820011	Kg. Ulu Klawang	Negeri Sembilan	02°54'N	102°01'E
SEL-1	2615131	Ldg. Batu Untong	Selangor	02°41'N	101°30'E
SEL-2	2815001	Pejabat JPS Sg. Manggis	Selangor	02°49'N	101°32'E
SEL-3	3218101	Stsesen Jenaletrik Lln	Selangor	03°12'N	101°52'E
PHG-1	3524080	Kg. Tebing Tinggi	Pahang	03°31'N	102°25'E
PHG-2	3431099	Kg. Serambi	Pahang	03°29'N	103°08'E
PHG-3	3732021	Kg. Sg. Soi	Pahang	03°43'N	103°18'E
TRG-1	5230041	SK Kuala Telemong	Terengganu	05°12'N	103°01'E
TRG-2	4334094	SK Kijal	Terengganu	04°19'N	103°29'E
TRG-3	4434093	SK Kemasek	Terengganu	04°25'N	103°27'E
KEL-1	5921009	Ibu Bekalan Tok' Uban	Kelantan	05°58'N	102°08'E
KEL-2	6121067	Stesen Keretapi Tumpat	Kelantan	06°11'N	102°10'E
KEL-3	5120025	Balai Polis Bertam	Kelantan	05°08'N	102°02'E
PRK-1	4010001	JPS Telok Intan	Perak	04°01'N	101°02'E
PRK-2	5210069	Stsesen Pemeriksaan Hutan Lawin	Perak	05°17'N	101°03'E
PRK-3	5006021	Kolam Air Bkt. Merah	Perak	05°01'N	100°39'E

PEN-1	5204048	Sg. Simpang Ampat	Pulau Pinang	05°17'N	100°28'E
PEN-2	5302001	Tali Air Besar Sg. Pinang	Pulau Pinang	05°23'N	100°12'E
PEN-3	5404041	Ladang Malakoff	Pulau Pinang	05°29'N	100°27'E
KED-1	5407080	Ladang Dublin	Kedah	05°24'N	100°44'E
KED-2	5606077	Ladang Lubok Segintah	Kedah	05°38'N	100°38'E
KED-3	5206102	Terap	Kedah	05°16'N	100°37'E
PLS-1	6501005	Abi Kg. Bahru	Perlis	06°30'N	100°10'E
PLS-2	6502010	Bukit Temiang	Perlis	06°32'N	100°13'E
PLS-3	6602002	Kaki Bukit	Perlis	06°38'N	100°12'E

1.5 Thesis Outline

The rest chapter of the thesis is organised as follows. The next chapter presents an analytical review of the literature related to this research study. The review describes the previous studies on flood analysis using Standard Precipitation Index (SPI) method, multivariate copula, tail dependence and conditional of the return period. It relates to the previous researchers experience in the use of the method study.

Chapter 3 describes the methodology of the approach. It describes how SPI works to estimates the occurrence of the flood and its rainfall characteristics. It is necessary to understand the inner structure of the model before conducting further analysis. Each characteristic can influence the result differently. This chapter focuses on the difference of the rainfall characteristics and describes the point of severity of wet or dry conditions.

Chapter 4 discusses the results and analysis conducted in this study. These results are calculated and determined by the model. The analysis is based on the identified results that have been. The final chapter the entire results of the study that have been analysed and some recommendations based on analytical results will be described.

CHAPTER 2

LITERATURE REVIEW

2.1 Flood History

Hydrologist are concerned about extreme complexity and to make decisions on water resources planning and implementation of flood mitigation strategies. This is attributed to the floods are among the most devastating natural disasters that cause mortalities and financial losses in the affected areas. They are essential not only for their destructive effects at the time of occurrence but also for their post-disaster damaging effects persisting for a long time.

Floods are a natural disaster because with excessive quantities of water can cause losses to the local economy and the destroy of an area and property and involve loss of life. Flood disasters can be associated with several types such as coastal floods, flash floods, river floods, groundwater floods, and sewerage floods (Rees, J. A., & Handmer, J, 1988; Shawn, 1989). Coastal floods occur when storms or severe weather joint with high tides. Although many factors contribute to coastal flooding, heavy rain from a storm causes sea levels to rise above usual, forcing seawater to overflow to the ground surface. For causing coastal flooding; flash floods occur when heavy rains occurred an area where the quantity of water accumulates in large quantities and is