MARKOV CHAIN-MIXED EXPONENTIAL MODEL FOR DAILY RAINFALL IN HONG KONG

XU YUCHEN

A project report submitted in partial fulfilment of the requirements for the award of Master of Mathematics

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

December 2023

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature : 

Name : Xu Yuchen

ID No. : 22UEM0154

Date : 5 Dec 2023

APPROVAL FOR SUBMISSION

I certify that this project report entitled **"MARKOV CHAIN-MIXED EXPONENTIAL MODEL FOR DAILY RAINFALL IN HONG KONG"** was prepared by **XU YUCHEN** has met the required standard for submission in partial fulfilment of the requirements for the award of Master of Mathematics at Universiti Tunku Abdul Rahman.

Approved by,

Signature	:	
Supervisor	:	Dr. Tan Wei Lun
Date	:	14 December 2023

Signature	:	
Co-Supervisor	:	Dr. Ng Kooi Huat
Date	:	14 December 2023

ACKNOWLEDGEMENTS

I would first like to thank my parents for their unquestioning support in continuing my education, not only mentally but also materially. I am very grateful to my supervisor Dr. Tan Wei Lun for his great help, academic and spiritual support and encouragement, and I would also like to thank my co-supervisor Dr. Ng Kooi Huat for giving me many ideas and reminding me to make less detours. Of course, I would like to thank my head of program, Dr. Wong Wai Kuan, who has been very kind and caring to me since I started my study in Universiti Tunku Abdul Rahman.

I would like to thank the Hong Kong Observatory for providing me with detailed meteorological data so that I can do my research well. All data used in this paper belongs to the Hong Kong Government.

I would also like to thank myself for finally persevering to finish my studies, and at this moment, it may be time to start a new path, when finish this study, other things should be simpler for me.

ABSTRACT

In this study, we applied a stochastic rainfall model which is capable in generating synthetic daily rainfall sequences that exhibit similar characteristics to observed data, thereby assessing the amount of rainfall over a specific period. The model utilized for this purpose is the Markov Chain Mixing Index (MCME). This model integrates both rainfall occurrence, represented by a first-order two-state Markov chain, and rainfall distribution, described by a mixture index distribution. The feasibility of the MCME model was evaluated using daily rainfall data collected from 15 stations in Hong Kong over a 20-year record period (2003-2022). The evaluation revealed that the proposed MCME model adequately captures both the occurrence and quantity of rainfall across all stations. Various statistical analysis were implemented to analyze the rainfall data. In conclusion, the validation results indicate that while the model effectively describes the characteristics of rainfall and able to simulate the rainfall based on the parameters estimated.

TABLE OF CONTENTS

DECLARATION			**ii**
APPROVAL FOR SUBMISSION			**iii**
ACKNOWLEDGEMENTS			**v**
ABSTRACT vi			
TABLE OF CONTENTS			**vii**
LIST OF TABLES			**ix**
LIST OF FIGURES			**x**
LIST OF SYMBOLS			**xi**
LIST OF ABBREVIATIONS			**xii**
CHAPTER			**1**
1	**INTRODUCTION**		**1**
	1.1	Background of the study	1
	1.2	Research Objective	2
	1.3	Study Scope	3
	1.4	Problem Statement	3
2	**LITERATURE REVIEW**		**5**
	2.1	Introduction	5
	2.2	Analysis of time series data	5
	2.3	Rainfall fitting models	5
	2.4	Stochastic Rainfall Simulator	6
	2.5	Evaluation of models	7
3	**METHODOLOGY**		**9**
	3.1	Rainfall data processing	9
	3.2	The Markov Chain-Mixed Exponential Model	10
	3.3	Simulation: A Rainfall Generator	12
	3.4	MCME Model Evaluation	13

4 RESULTS AND DISCUSSIONS 15

4.1 Data Description 15

4.2 Performance of mixed exponential distributions 26

4.3 Performance of the Month MCME Model in Calibration Period (2003-2022) 40

4.4 Assessment of the MCME Model 48

5 CONCLUSIONS AND RECOMMENDATIONS 57

5.1 Conclusions 57

5.2 Recommendations for future work 57

REFERENCES 58

APPENDICES 60

LIST OF TABLES

Table 2.5.1: Different assessment methods used by by past researchers 8

Table 4.1.1: List of Rainfall Stations with Geographical Coordinates and Characteristics 15

Table 4.1.2: Descriptive statistics for data from all rainfall stations 16

Table 4.1.3: Comparison of total values for February with and without leap years 25

Table 4.2.1: Comparison of the parameters of the mixing exponential distribution 27

Table 4.3.1: Summary of MCME parameters estimation for all rainfall stations 40

Table 4.3.2: Mean and Standard deviation of all MCME parameters estimation 44

Table 4.4.1: Observed and simulated monthly rainfall intensities 49

Table 4.4.2: Monthly RE average 49

LIST OF FIGURES

Figure 3.1.1: Characteristics of a Boxplot 10

Figure 3.3.1: Simulation process 13

Figure 4.1.1: The location of the rainfall stations sites 16

Figure 4.1.2: Box plots of annual rainfall at 15 rainfall stations 17

Figure 4.1.3: Box plot of monthly rainfall at each rainfall station over the last 20 years (2003-2022) 24

Figure 4.2.1: Comparison of observed and modeled rainfall fits for each rainfall station 39

Figure 4.3.1: Line chart of p for different rainfall Stations across Months 45

Figure 4.3.2: Line chart of $\mu 1$ for different rainfall stations across months 46

Figure 4.3.3: Line chart of $\mu 2$ for different rainfall stations across months 47

Figure 4.3.4: Line chart of $p01$ for different rainfall stations across months 47

Figure 4.3.5: Line chart of $p11$ for different rainfall stations across months 48

Figure 4.4.1: Box plot illustrating monthly rainfall at each rainfall stations 56

LIST OF SYMBOLS

δ	Actual relative error
Δ	Absolute error
L	True value
p	Mixing probability of the Mixed Exponential distribution
μ_1, μ_2	Mean parameters of the Mixed Exponential distribution
p_{00}	Probability of a day to be dry given that the previous day was dry
p_{01}	Probability of a day to be wet given that the previous day was dry
p_{10}	Probability of a day to be dry given that the previous day was wet
p_{11}	Probability of a day to be wet given that the previous day was wet

LIST OF ABBREVIATIONS

ME	Mixed Exponential model
MCME	Markov Chain Mixed Exponential model
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
AE	Absolute Error
RE	Relative Error
PDF	Probability Density Function
IQR	The interquartile range
Q1	Splits off the lowest 25% of data from the highest 75%
Q3	Splits off the highest 25% of data from the lowest 75%

LIST OF APPENDICES

APPENDIX A: Computer Programme Core Code 61

CHAPTER 1

INTRODUCTION

1.1 Background of the study

Rainfall significantly impacts human life and affects activities in agriculture and the economy. In urban areas, rainfall exceeding drainage capacity can cause urban flooding, directly affecting the lives of residents and leading to property losses for both the state and individuals. It also triggers secondary disasters, such as landslides and mudslides, further increasing overall losses. In rural areas, rainfall greatly influences the growth of various types of crops. Moderate rainfall promotes crop growth, while both insufficient and excessive rainfall can result in crop losses.

The advantage of stochastic rainfall simulators is that they not only compensate for the lack of historical data but also provide the ability to synthesize rainfall events from different time periods. This is critical for modeling extreme rainfall events that may occur under various meteorological conditions and for assessing the flood resilience of urban sewer systems. Additionally, these generators are more flexible in adapting to different spatial requirements, thereby offering a broader range of more accurate simulation results. This flexibility is valuable for scientists and engineers in their applications.

Due to the uncertainty of climate change and the stochastic nature of rainfall, scientists have endeavored to predict rainfall using various methods. However, there is currently no highly reliable and effective method for predicting rainfall in a timely and accurate manner.

In recent years, many researchers have turned to applying Markov chains to model rainfall, with a common approach involving two key steps. In the first step, a Markov chain generates the time series of rainfall occurrence (i.e., wet day or dry day), providing a simple and effective means of preserving wet and dry periods. In the second step, the amount of rainfall on a wet day is typically estimated using various probability distributions. A significant number of researchers continue to utilize this approach in an ongoing effort to enhance the effectiveness of the model.

A Markov chain represents a network of concepts and propositions utilized to organize, describe, and explain our experiences through mathematical analysis of natural processes. In the early 20th century, Markov discovered, through numerous observational experiments, that the state obtained by the nth transition in the process of state transitions of a system is related to the state of the previous (n-1) transition. After extensive experiments and generalizations, it was concluded that for a system, there exists a transfer probability during the transition from one state to another. This transfer probability can be deduced based on one of its previous states, independent of the original state of the system and the Markov process before this transfer.

Predicting rainfall serves as a crucial foundation for government decision-making and provides essential information for disaster prevention departments. This, in turn, helps mitigate threats to people's lives and properties, fostering social stability. Additionally, it aids water conservancy departments in making informed decisions to address flooding and manage water conservancy facilities effectively. The data generated supports municipal construction departments in developing urban drainage facilities. Social enterprises, including insurance companies, can enhance the accuracy of insurance rate calculations by leveraging weather disaster compensation data. Furthermore, accurate rainfall prediction contributes to facilitating daily activities, such as commuting, thereby playing a vital role in promoting regional economic prosperity and environmental protection.

Forecasting rainfall is essential for government decision-making, especially in disaster prevention. It provides crucial information that helps in reducing risks to people's lives and properties, thereby promoting stability in society. Additionally, it supports water conservancy departments in managing floods by guiding the operation of water facilities. This data also assists municipal construction departments in creating effective urban drainage systems. Moreover, rainfall forecasts aid social enterprises like insurance companies in improving their calculations for weather-related compensation. These accurate forecasts also contribute to everyday activities such as commuting, ultimately boosting regional economic growth and protecting the environment.

1.2 Research Objective

The research objectives for this study are:

1. To investigate the rainfall characteristics in Hong Kong.
2. To apply Markov chain-mixed exponential to model the rainfall in Hong Kong.

3. To assess the performance of the Markov chain-mixed exponential for its rainfall.

1.3 Study Scope

Hong Kong, officially known as the Hong Kong Special Administrative Region, is a special administrative region of the People's Republic of China. Situated in the southern part of China, it is an integral part of the Guangdong-Hong Kong-Macao Greater Bay Area and is internationally recognized as a global city. Positioned on the north coast of the South China Sea and on the east side of the Pearl River Estuary, Hong Kong experiences a subtropical climate characterized by four distinct seasons. The period from May to November is considered the windy season, while July to September is particularly vulnerable to tropical cyclones.

On average, about 30 tropical cyclones impact the western part of the North Pacific Ocean and the South China Sea annually, with approximately 15 of them reaching typhoon strength or above. Hong Kong serves as a crucial corridor for both the summer southwest monsoon and winter northeast monsoon currents in Asia, and the region's rainfall is significantly influenced by these two types of monsoon air masses.

Adverse weather conditions affecting Hong Kong include tropical cyclones, strong winter and summer monsoon winds, monsoon troughs, and severe thunderstorms, which often occur between April and September. The rainfall data used in this study were obtained from the Hong Kong Observatory, specifically from 15 rainfall stations spanning the years 1993 to 2022.

In this study, we aim to analyze Hong Kong's rainfall patterns and develop a rainfall model using the Markov Chain-Mixed Exponential Distribution. This effort seeks to create a more advanced and accurate method for simulating rainfall, specifically designed for practical applications. The model's development is anticipated to provide more dependable hydrological information, ultimately supporting better decision-making and planning in relevant fields.

1.4 Problem Statement

Accurate rainfall prediction can significantly benefit human life by providing sufficient time to minimize negative impacts through measures such as technical interventions for artificial rainfall. It can also be leveraged to bring positive benefits to people's lives. For instance, in the design and management of urban sewers, crucial hydrological information—such as rainfall amount and hourly rate—is essential for modeling and determining the design parameters of the sewer system. Typically, these models rely on historical data, especially daily rainfall data, for accurate hydrologic simulations. However, in practice, historical rainfall records often encounter various issues, such as insufficient length, incomplete data, and inadequate spatial coverage, making it challenging to obtain reliable simulation results.

To address these challenges, rainfall stochastic simulation or stochastic rainfall simulators become essential. These tools are widely used to generate a large number of synthetic rainfall time series capable of accurately characterizing the physical and statistical properties of rainfall processes observed at a specific location. By employing these generators, engineers and water resource professionals can create more comprehensive and reliable hydrologic models to support the design and management of urban sewer systems.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Rainfall modeling, a crucial research area in hydrology and meteorology, has witnessed extensive development and evolution over the years. In recent decades, researchers have made significant strides in enhancing the application and accuracy of statistical methods and techniques. In this chapter, we review the various models in rainfall modelling.

2.2 Analysis of time series data

The analysis of time series data serves as the cornerstone of rainfall modeling. This data comprises three main components: trend, seasonality, and residuals or white noise. The trend reflects the slow change or direction of the record over time, providing the foundation for subsequent modeling. On the other hand, seasonality examines cyclical patterns occurring over a fixed period, such as the frequency of events in a given month or season. Finally, residuals or white noise represent random variations that cannot be explained by trends or seasonal components.

Commonly used models in the statistical modeling of time series data include autoregressive (AR) models, moving average (MA) models, autoregressive moving average (ARMA) models, and autoregressive integrated moving average (ARIMA) models.

Zhang and Xia (2012), Arumugam and Karthik (2018), and Jale et al. (2019) utilized Markov chains to generate time series in their respective studies.

2.3 Rainfall fitting models

Fitting a distribution to rainfall is a challenge due to the complexity of atmospheric processes and the hydrologic cycle.

The common probability distribution used by the researchers included the exponential, Gamma, lognormal, Weibull, and Poisson distributions. The exponential distribution is often employed to describe the time between equally spaced independent events, while the Gamma distribution is suitable for characterizing the duration and intensity distribution of rainfall events. The lognormal distribution accommodates the right skewness observed in rainfall data, and the Weibull distribution is commonly applied when studying extreme rainfall events. Additionally, for larger rainfall events, the lognormal distribution is a frequent choice, while the Poisson distribution is appropriate for describing the counts of rainfall events occurring over a specific period.

In a study by Shibabaw et al. (2022), a stochastic daily rainfall model was developed, focusing on the spatial and temporal distribution of rainfall in Ethiopia. The research utilized the Markov chain along with a combination of the Weibull distribution, lognormal distribution, mixed exponential distribution, and Gamma distribution to construct the rainfall model.

In certain cases, opting for a mixture of distributions is a reasonable choice as it enables the simultaneous consideration of the impacts of multiple probability distributions on the rainfall process. The selection of an appropriate distribution involves model testing and parameter fitting, ensuring that the probability density function of the chosen distribution closely aligns with the observed rainfall data.

2.4 Stochastic Rainfall Simulator

A stochastic rainfall simulator is a tool that combines time series and rainfall distributions to synthesize stochastic rainfall events with trends and characteristics similar to the observed data. By assuming that the synthesized time series shares the same trend as the observed series and that the rainfall distribution is a highly fitting representation, the simulator can generate synthesized data closely resembling actual rainfall events.

2.4.1 Markov Chain-Mixed Exponential Model (MCME)

The Markov Chain-Mixed Exponential Model (MCME) is a composite model that integrates Markov chains and mixed exponential distributions to comprehensively characterize two key aspects of the rainfall process: the occurrence of rainy days and the distribution of daily rainfall. Through in-depth analysis of historical rainfall data, the model reveals the state transfer law between dry and wet days using Markov chains. Simultaneously, it more accurately portrays the probability density function of daily rainfall by employing the mixed exponential distribution.

Successive studies by Hussain (2008), Yusof et al. (2015), Senthamarai Kannan and Jawahar Farook (2015), El Outayek (2020), and Berhane et al. (2020) have employed MCME models to simulate rainfall in distinct regions, all yielding improved results.

2.4.2 Advantages of Stochastic Rainfall Simulator

The stochastic rainfall simulator, as a simulation tool, offers numerous advantages in the fields of hydrology and meteorology. It not only generates a substantial amount of rich data in a short period, thereby enhancing the estimation of the probability of extreme events, but also proves useful in simulating scenarios with missing data. Furthermore, the stochastic rainfall simulator can be applied for spatial interpolation, enabling the simulation of rainfall events at various geographic locations with the added flexibility of temporal downscaling. Owing to its capability to generate diverse data, the stochastic rainfall simulator finds wide applications in areas such as water resource management, flood risk assessment, urban planning, and meteorological studies.

In summary, academic modeling of rainfall not only enhances the comprehension of rainfall processes but also establishes a scientific foundation for the fields of hydrology and meteorology. The selection and parameter estimation of these models hold significant practical implications for water resource management, flood risk assessment, and climate change impact studies. Thus, the interdisciplinary application of statistics, probability theory, and hydrology remains particularly crucial in this research domain.

2.5 Evaluation of models

In modeling daily and extreme rainfall events, researchers have extensively employed various stochastic weather generators, comparing their performance. However, a consistent consensus on the superior type of stochastic weather generator across different locations is yet to be established. These variations may be influenced by the meteorological and hydrological characteristics specific to each study site.

Many scholars often resort to mathematical methods to discern differences between observed and model-generated data. Table 2.5.1 below outlines some of the methods employed by researchers for model evaluation at different sites.

Table 2.5.1: Different assessment methods used by by past researchers

Author	Mathematical methods	Plot
Hussain, A. (2008).	MAE (Mean Absolute Error) RMSE (Root Mean Squared Error)	Chart box plots Line plots Histograms
El Outayek, S. (2020).	RMSE(Root Mean Squared Error)	Box plots
Berhane, T., Shibabaw, N., Awgichew, G., & Kebede, T. (2020).	AE (Absolute Error)	Bar charts
Shibabaw, N., Berhane, T., Kebede, T., & Walelign, A. (2022).	AE (Absolute Error)	Line plots

The application of these methods at various locations has equipped researchers with a diverse set of intuitive and comprehensive tools for assessing the simulation effectiveness of stochastic weather generators across different geographic environments. Future studies can explore more advanced assessment methods to further enhance our understanding of the performance of random weather generators.

CHAPTER 3

METHODOLOGY

3.1 Rainfall data processing

In academic research, the dataset's quality is fundamental in shaping the reliability of the final model and the validity of research outcomes. A dataset characterized by reliability, validity, quality, and consistency serves as a cornerstone supporting academic research endeavors and constructing robust models. This ensures that the conclusions drawn from the research hold credibility both within academic realms and practical applications. Before cleaning the data, it is essential to thoroughly understand the dataset's context, structure, and features through extensive literature review and dataset characterization. Handling missing values is a critical step in data preprocessing, necessitating a thoughtful strategy to manage them, which may involve removal, interpolation, or imputation with mean or median values. Eliminating duplicate records from the dataset ensures data uniqueness, while addressing outliers using statistical methods or domain knowledge maintains data consistency and quality. Standardizing data formats and normalizing numerical values enhance data consistency, making subsequent analysis and modeling more reliable. Error correction, involving rectifying spelling mistakes, standardizing naming conventions, and resolving other accuracy-affecting issues, is imperative for maintaining dataset accuracy, particularly for scholarly research. Documenting each step of the data cleaning process and its effects enables others to replicate and understand the procedure. Visual analysis through visualization tools aids in comprehending data distributions, relationships, and trends, providing additional support for academic research endeavors.

3.1.1 Boxplot

Figure 3.1.1 provides a detailed characterization of the boxplot diagram, offering insights into how it represents accuracy, robustness, and variability in estimating parameters from the generated rainfall series. This intuitive representation enables a swift understanding of central tendencies and potential outliers, allowing for a concise assessment of the reliability of the parameter estimates.

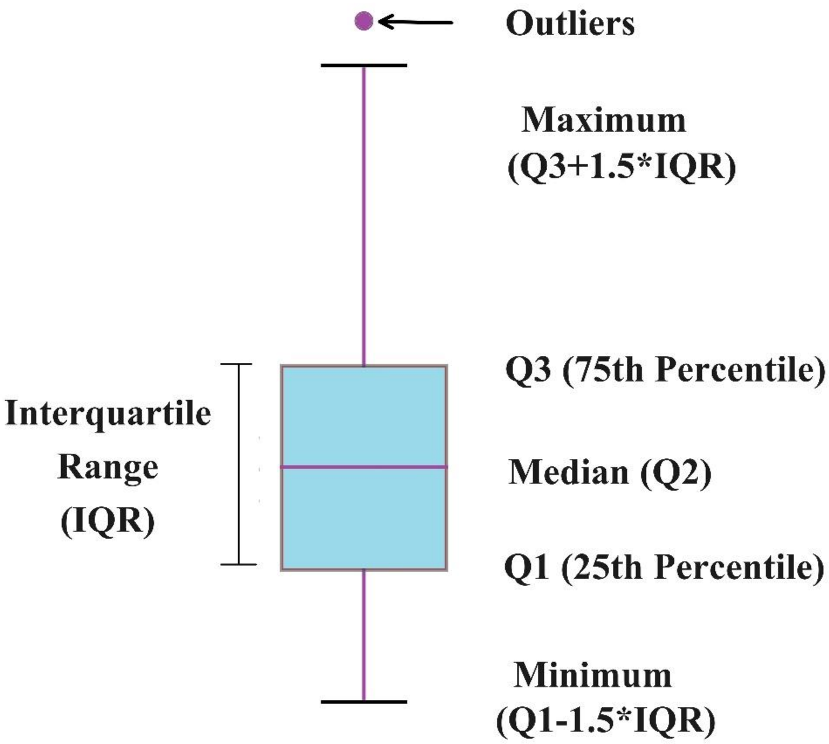


Figure 3.1.1: Characteristics of a Boxplot

3.2 The Markov Chain-Mixed Exponential Model

3.2.1 The Occurrence Process

In previous studies, researchers have often recommended the use of Markov chains for modeling daily rainfall (Chin, 1977; Roldan and Woolhiser, 1982). These studies simplify the observed rainfall data series as a sequence of two states: dry or wet. The correlation between wet and dry days on consecutive days is modeled using a first-order Markov Chain denoted by 0 or 1, representing dry and wet days, respectively. We distinguish at 2.5 mm; when the rainfall in a day is less than 2.5 mm, it is defined as a dry day (denoted by 0), and when it is greater than or equal to 2.5 mm, it is defined as a wet day (denoted by 1). (Note: According to China's meteorological operations, light rain is defined as rainfall less than or equal to 2.5 mm in 1 hour.)

$$X_n = \begin{cases} 0 \ if \ the \ n^{th} \ day \ is \ dry \\ 1 \ if \ the \ n^{th} \ day \ is \ wet \end{cases} \tag{3.1}$$

Hence, the transition probabilities of the first-order Markov chain are defined as follows:

$$P\,(X_{n+1} = j\,|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = p\,(i, j) \tag{3.2}$$

3.2.2 The Rainfall Amount

Mixed exponential distribution is commonly used to fit the rainfall data (citation). The mixed exponential distribution is described as follows,

$$f(x) = \frac{p}{\mu_1} e^{\frac{-x}{\mu_1}} + \frac{1-p}{\mu_2} e^{\frac{-x}{\mu_2}} \tag{3.3}$$

for

$x > 0, 1 > p > 0, \mu_1 > 0, \mu_2 > 0.$

Where μ_1 and μ_2 are the means of two exponential distributions, p represents the mixing probability, dictating the allocation of weights to the two exponential distributions.

3.2.2.1 Estimation of Parameters through the Method of Maximum Likelihood

The parameters of the mixed exponential distribution are determined using the maximum likelihood method. To solve for these parameters, the log-likelihood function is defined as follows:

$$logL = \sum_{i=1}^{n} log\left(\frac{p}{\mu_1} e^{\frac{-x_i}{\mu_1}} + \frac{1-p}{\mu_2} e^{\frac{-x_i}{\mu_2}}\right) \tag{3.4}$$

3.2.3 Parameter Optimization Techniques

Optimal solutions for maximizing the log-likelihood function can be achieved through an iterative optimization technique. The parameter estimates are obtained by solving the log-likelihood equation, as described by:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} \frac{\frac{p}{\mu_1} e^{\frac{-x_i}{\mu_1}}}{\frac{p}{\mu_1} e^{\frac{-x_i}{\mu_1}} + \frac{1-p}{\mu_2} e^{\frac{-x_i}{\mu_2}}} \tag{3.5}$$

$$\widehat{\mu_1} = \frac{1}{n\hat{p}} \sum_{i=1}^{n} \left(\frac{\frac{p}{\mu_1} e^{\frac{-x_i}{\mu_1}}}{\frac{p}{\mu_1} e^{\frac{-x_i}{\mu_1}} + \frac{1-p}{\mu_2} e^{\frac{-x_i}{\mu_2}}} \right) x_i \tag{3.6}$$

$$\widehat{\mu_2} = \frac{1}{n(1-\hat{p})} \sum_{i=1}^{n} \left(\frac{\frac{1-p}{\mu_2} e^{\frac{-x_i}{\mu_2}}}{\frac{p}{\mu_1} e^{\frac{-x_i}{\mu_1}} + \frac{1-p}{\mu_2} e^{\frac{-x_i}{\mu_2}}} \right) x_i \tag{3.7}$$

The optimal solution for these iterative equations was obtained using a method recommended by Nguyen and Mayabi (1990), known for its fast convergence rate. Initial values can be assigned using the method of moments. However, to achieve a fast convergence rate, seven initial estimates for the three parameters are selected. The initial values for p and μ_1 range from 0.01 to 0.99 at intervals of 0.01, and from 0.01 $\hat{x}$ to 0.99 $\hat{x}$ at intervals of 0.01 $\hat{x}$, respectively. For a given pair of p and μ_1, the corresponding μ_2 is calculated as $\mu_2 = \frac{(\hat{x} - p/\mu_1)}{(1-p)}$. The optimal solution for the parameters is determined as the one providing the highest value among the iterations of all likelihood functions.

3.3 Simulation: A Rainfall Generator

An R language program was developed to implement our rainfall simulation. The simulation involves generating daily rainfall using the MCME model for each station and each month of data. The rainfall simulation model operates as follows:

For a specific month at any site, the initial step of the model involves constructing a Markov transfer matrix for dry and wet days within that month. Subsequently, a time series of the same length is generated based on this matrix.

The second part of the model consists of constructing a mixed exponential model for the rainfall data of the same month. The time series generated in the first part is then input into the mixed exponential model to simulate rainfall data.

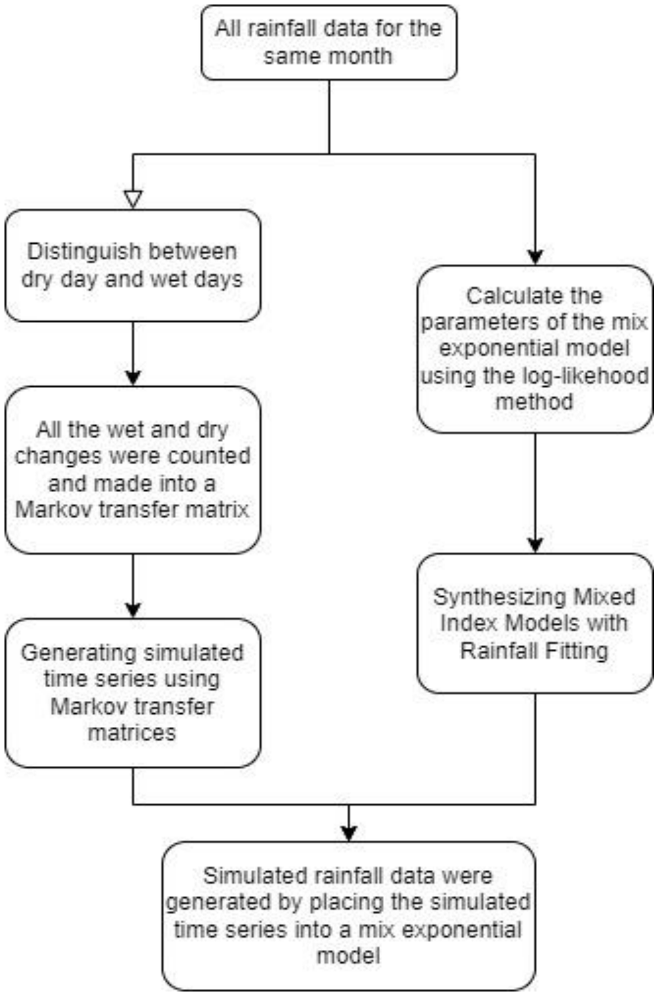


Figure 3.3.1: Simulation process

3.4 MCME Model Evaluation

Monthly rainfall sequences from 15 sites were employed to assess the performance of the MCME model. The complete 20-year daily rainfall series was utilized to calculate MCME model parameters, while the existing model parameters were employed to generate daily rainfall for the same 20-year duration. The observed rainfall will be compared with the generated rainfall. This comparative analysis was conducted on monthly basis. However, for the month of February, the analysis was not further processed as the rainfall totals for February were found to be unaffected by leap year versus non-leap year.

3.4.1 Relative Error

Relative Error (RE) is a metric widely employed in measurement science and engineering to quantify the disparity between actual observations and theoretical or true values. It was introduced to address the inherent bias of absolute error for datasets of varying magnitudes, offering a standardized method for evaluating accuracy in diverse contexts.

The mathematical expression for the relative error is given below:

$$\delta = \frac{\Delta}{L} \times 100\% \tag{3.8}$$

where δ is actual relative error, usually given in percentage, Δ is absolute error, L is true value.

Among other considerations, the use of absolute values ensures that the relative error is non-negative and is expressed as a percentage, allowing for comparisons across different problem domains. This metric serves not only to evaluate the accuracy of an experiment or measurement but also to facilitate a quantitative analysis of whether the measurement tends to overestimate or underestimate the true value.

Relative error holds significance in scientific research, laboratory testing, and engineering design as it plays a crucial role in determining the reliability and accuracy of measurements. In both academic and industrial research, assessing the confidence and accuracy of results is paramount for ensuring experimental reproducibility and robust engineering design. Therefore, as a standardized assessment tool, relative error contributes to elevating confidence levels in the interpretation and application of measured data.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Data Description

The rainfall data were obtained from the Hong Kong Observatory. Fifteen rainfall observation sites with over 20 years of data were selected, and rainfall data spanning a total of 20 years (from 2003 to 2022) were utilized. Table 4.1.1 shows the geographic coordinates of the sites and Characteristics. All rainfall observation sites experienced data loss within 5%, with 12 sites having data loss within 2%. Therefore, we considered the data to have minimal impact on the results. Figure 4.1.1 shows the location of each rainfall station from the map. Table 4.1.2 presents descriptive statistics for data from all rainfall stations, providing a comprehensive overview of the dataset.

We defined a day with less than 2.5 mm of rainfall as a dry day. Given the overall high percentage of dry days, we uniformly treated all days with missing data as having 0 rainfall.

Table 4.1.1: List of Rainfall Stations with Geographical Coordinates and Characteristics

S	Automatic Weather Station	Location		Missing data day)	Missing data(%)	Wet day	Dry day
		Latitude N	Longitude E				
1	Cheung Chau	22°12'04"	114°01'36"	55	0.753%	1437	5867
2	Ching Pak House(Tsing Yi)	22°20'53"	114°06'33"	10	0.137%	1582	5722
3	King's Park	22°18'43"	114°10'22"	7	0.096%	1682	5622
4	Lau Fau Shan	22°28'08"	113°59'01"	33	0.452%	1479	5825
5	Pak Tam Chung (Tsak Yue Wu)	22°24'10"	114°19'23"	75	1.027%	1698	5606
6	Sha Tin	22°24'09"	114°12'36"	48	0.657%	1712	5592
7	Shek Kong	22°26'10"	114°05'05"	124	1.698%	1572	5732
8	Ta Kwu Ling	22°31'43"	114°09'24"	21	0.288%	1629	5675
9	Tai Mei Tuk	22°28'31"	114°14'15"	328	4.491%	1472	5832
10	Tai Mo Shan	22°24'38"	114°07'28"	202	2.766%	1979	5323
11	Tate's Cairn	22°21'28"	114°13'04"	117	1.602%	1910	5394
12	Tseung Kwan O	22°18'57"	114°15'20"	85	1.164%	1767	5537
13	Waglan Island	22°10'56"	114°18'12"	235	3.217%	1236	6068
14	Hong Kong Observatory	22°18'07"	114°10'27"	0	0.000%	1687	5617
15	Hong Kong International Airport	22°18'34"	113°55'19"	0	0.000%	1478	5826

Figure 4.1.1: The location of the rainfall stations sites

Table 4.1.2: Descriptive statistics for data from all rainfall stations

Automatic Weather Station	mean	std	max	Q1	median	Q3
Cheung Chau	4.5918	15.5028	215	0	0	0.5
Ching Pak House(Tsing Yi)	5.2769	17.4925	362	0	0	1
King's Park	6.2783	20.3638	324	0	0	1.5
Lau Fau Shan	4.2885	13.9991	226.5	0	0	1
Pak Tam Chung (Tsak Yue Wu)	6.1835	20.2235	339	0	0	1.5
Sha Tin	6.3830	20.7378	347	0	0	1.5
Shek Kong	5.4140	19.0490	414.5	0	0	1
Ta Kwu Ling	5.3723	17.3029	340.5	0	0	1.5
Tai Mei Tuk	4.6816	16.2394	307.5	0	0	1
Tai Mo Shan	6.4370	19.2266	375	0	0.5	3
Tate's Cairn	6.8782	20.9357	362	0	0	2.5
Tseung Kwan O	6.2073	19.3370	329	0	0	2
Waglan Island	3.3102	11.7302	198	0	0	0.5
Hong Kong Observatory	6.3912	20.5920	329.7	0	0	1.7
Hong Kong International Airport	5.0893	17.3044	337.5	0	0	0.8

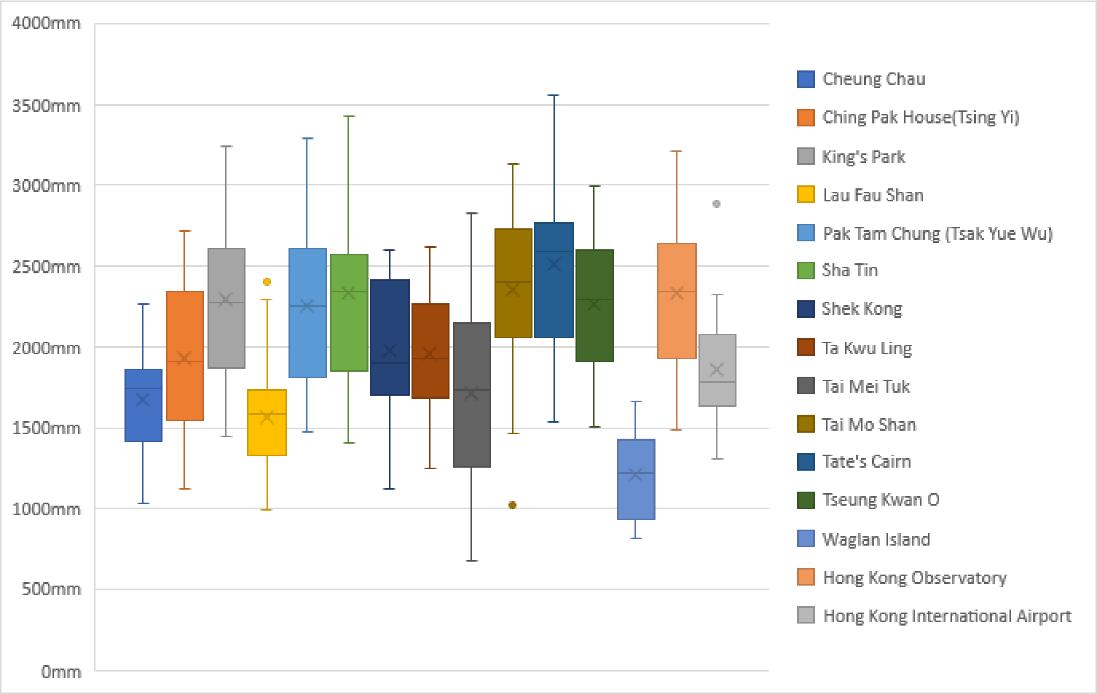
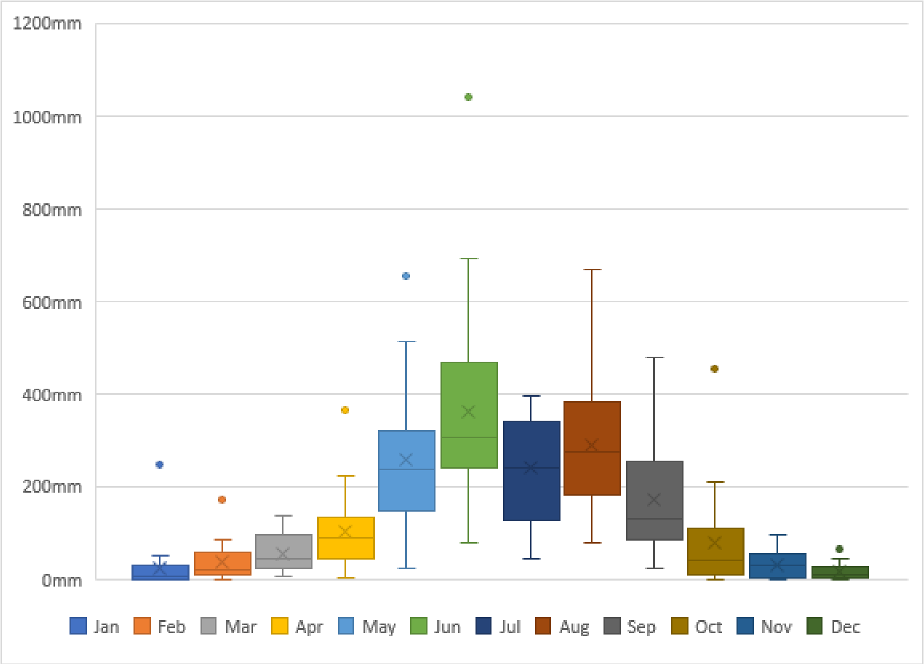
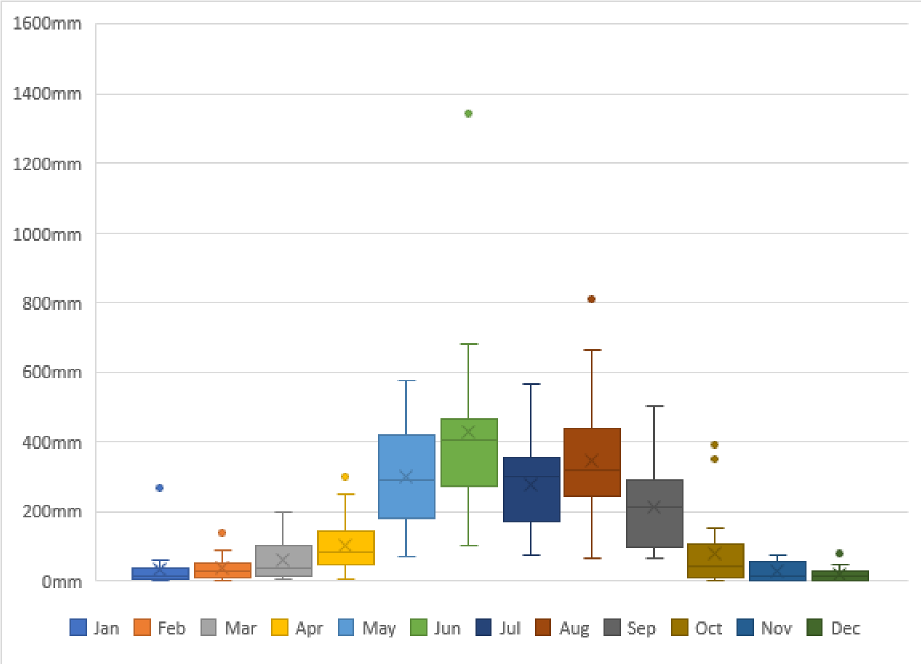


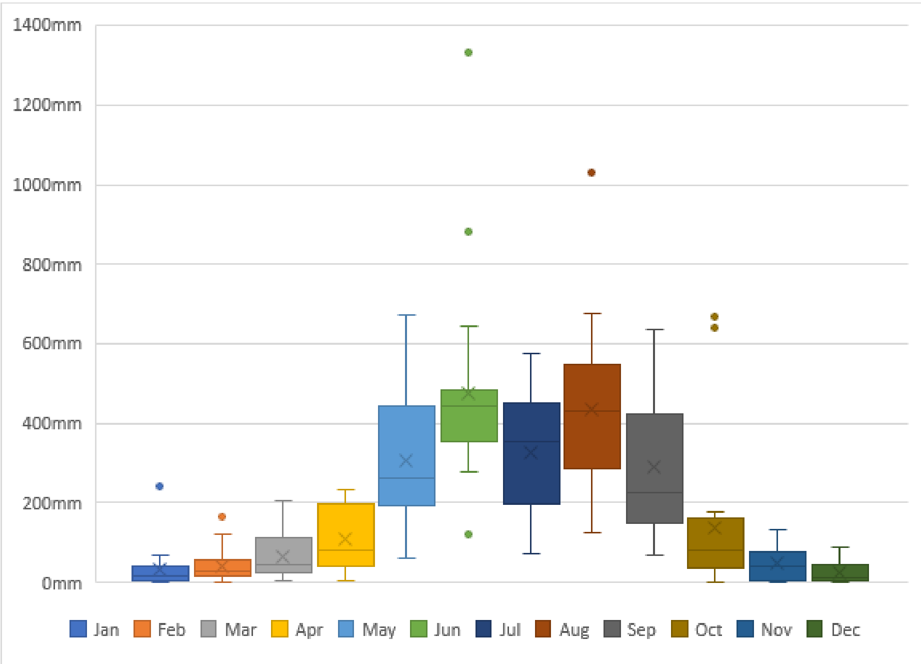
Figure 4.1.2: Box plots of annual rainfall at 15 rainfall stations



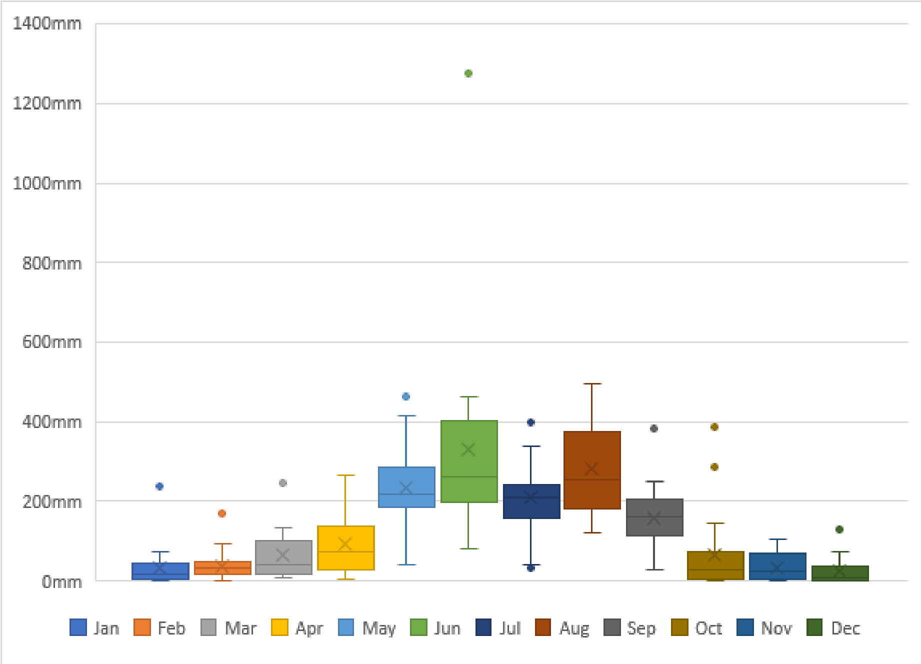
(a) Cheung Chau



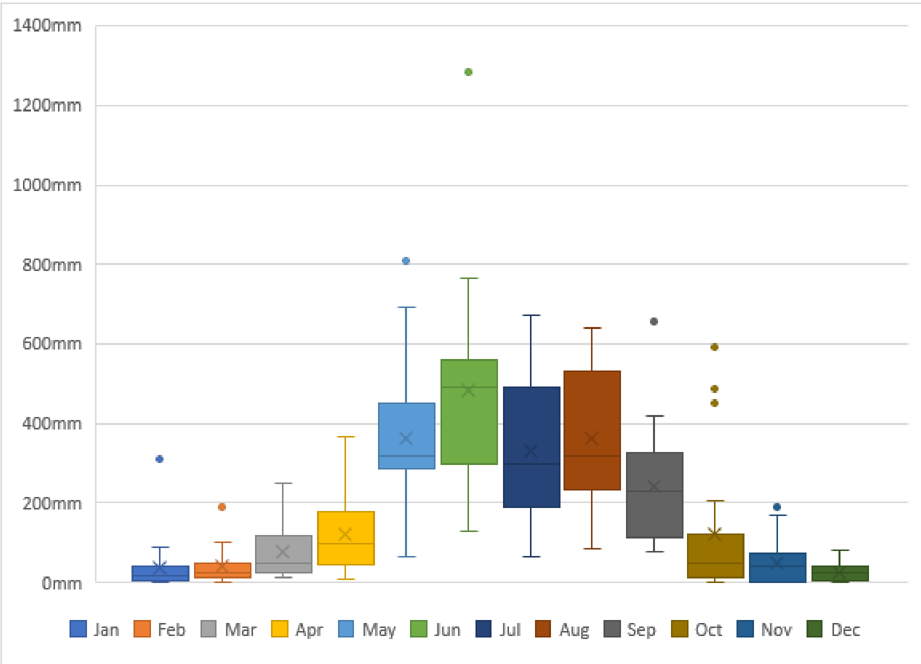
(b) Ching Pak House (Tsing Yi)



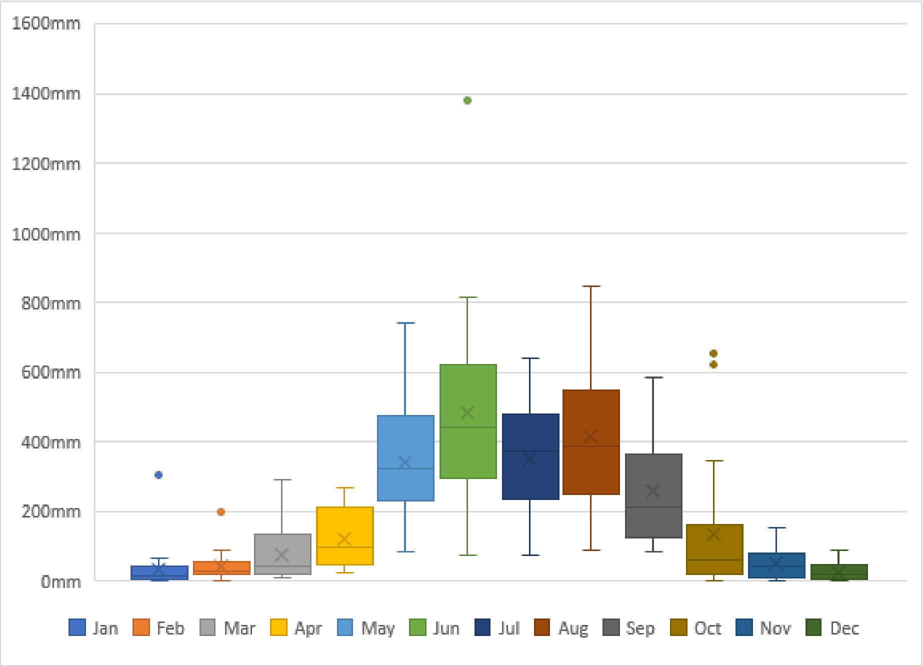
(c) King's Park



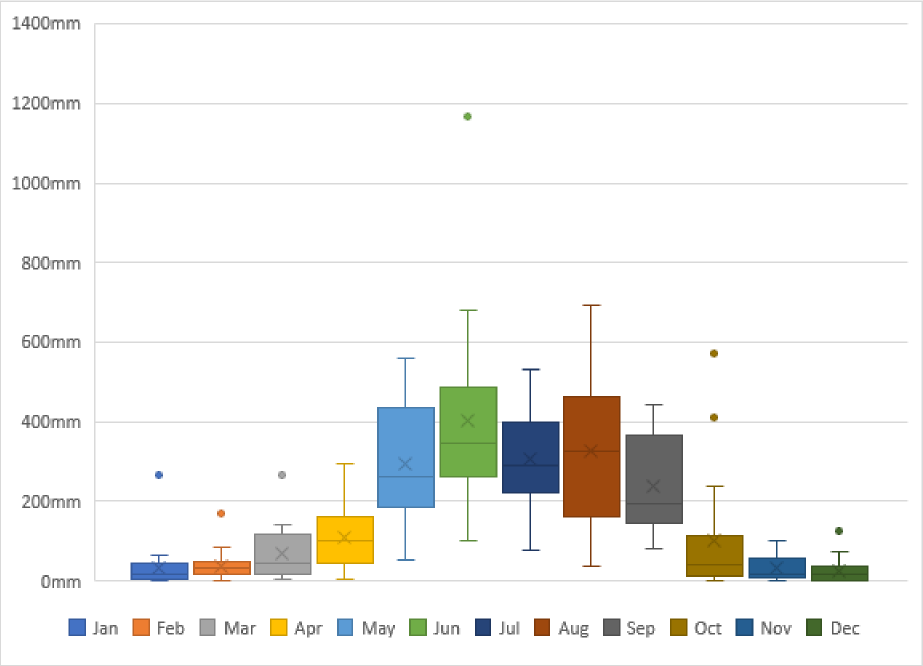
(d) Lau Fau Shan



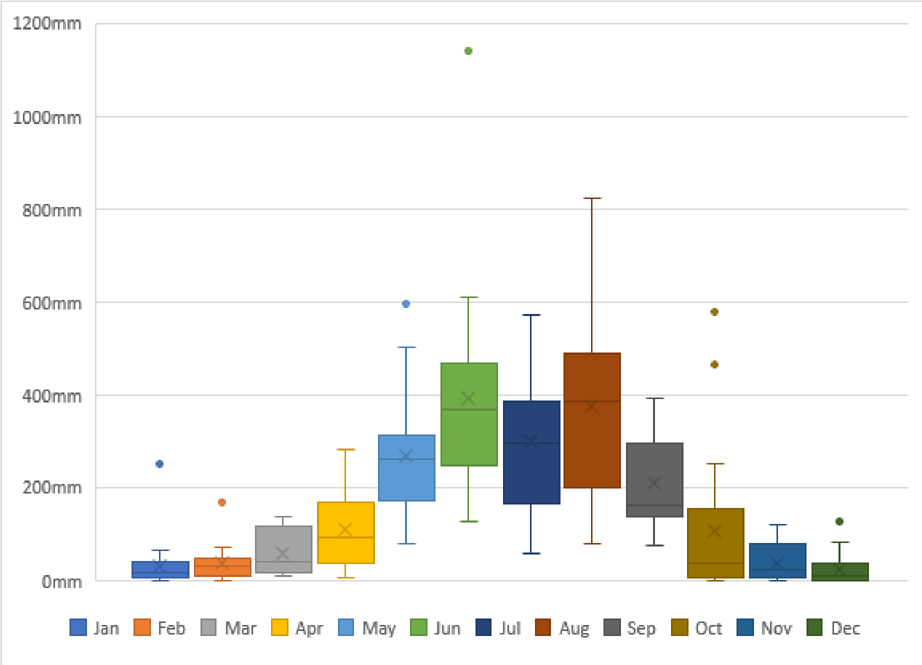
(e) Pak Tam Chung (Tsak Yue Wu)



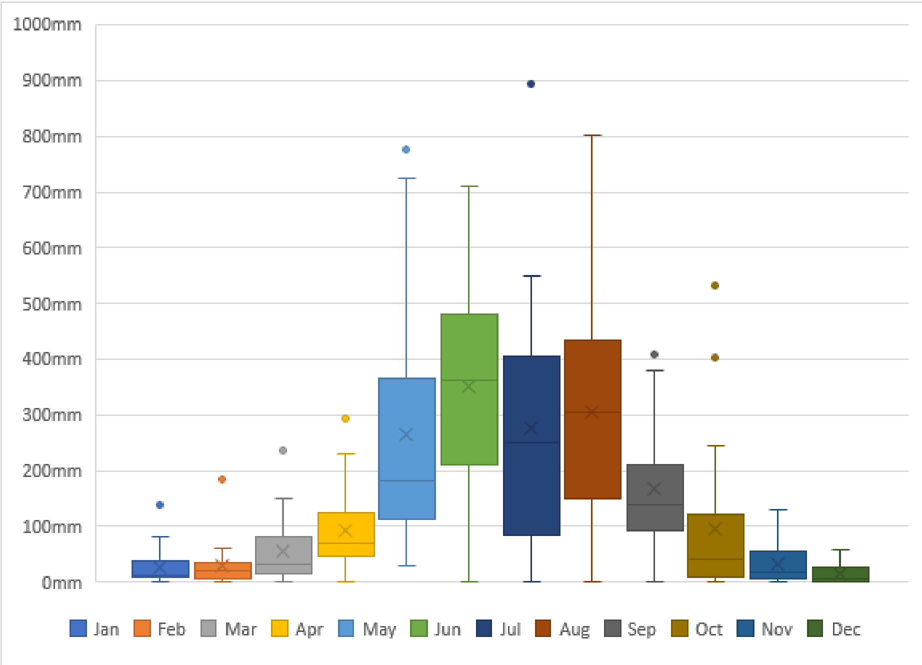
(f) Sha Tin



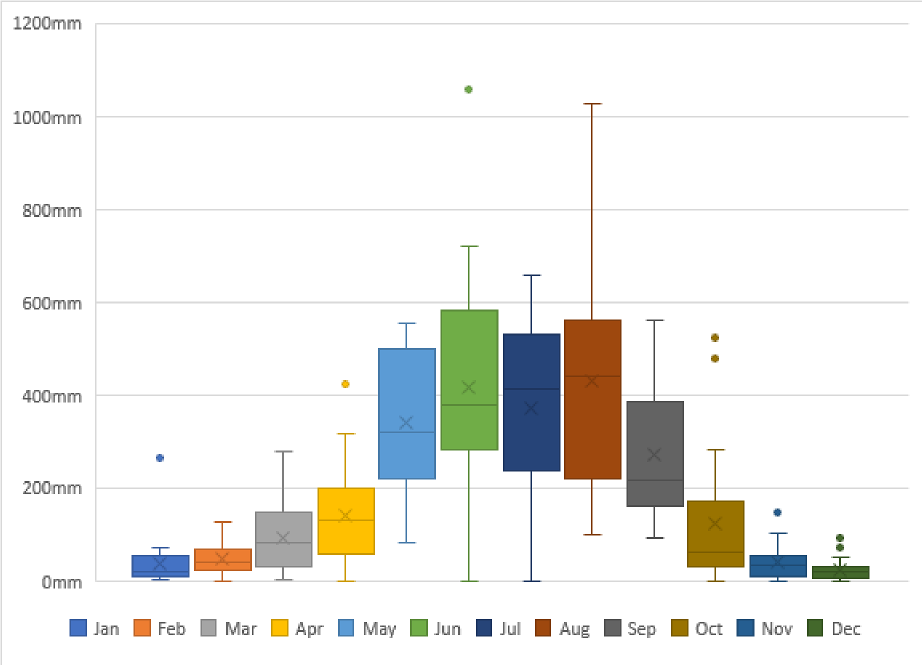
(g) Shek Kong



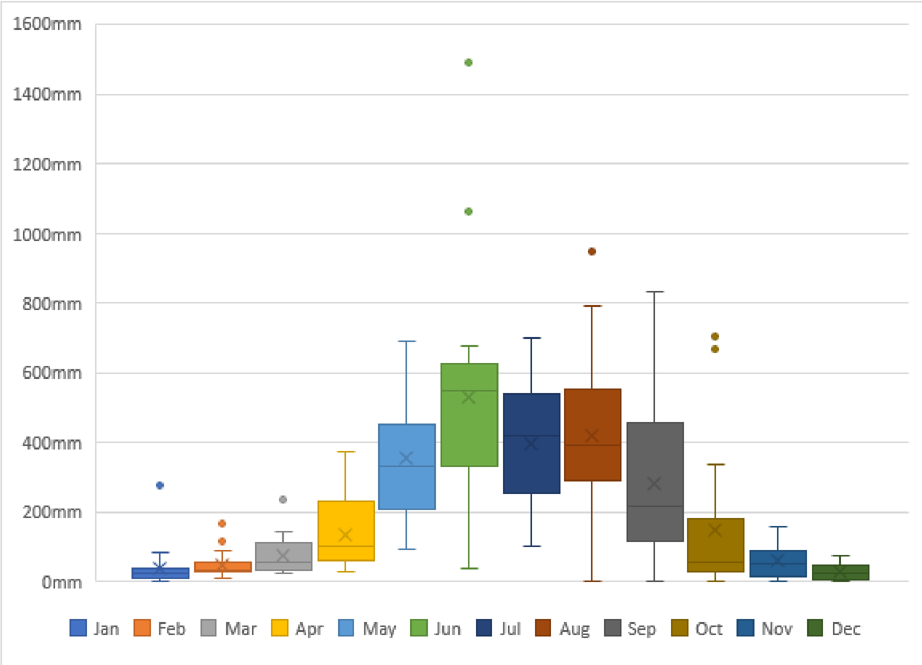
(h) Ta Kwu Ling



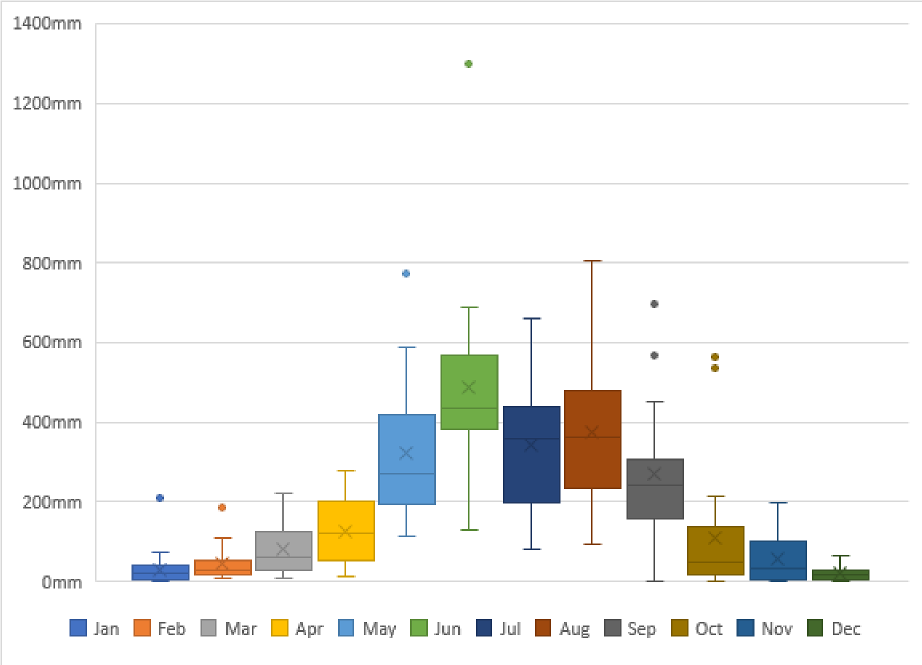
(i) Tai Mei Tuk



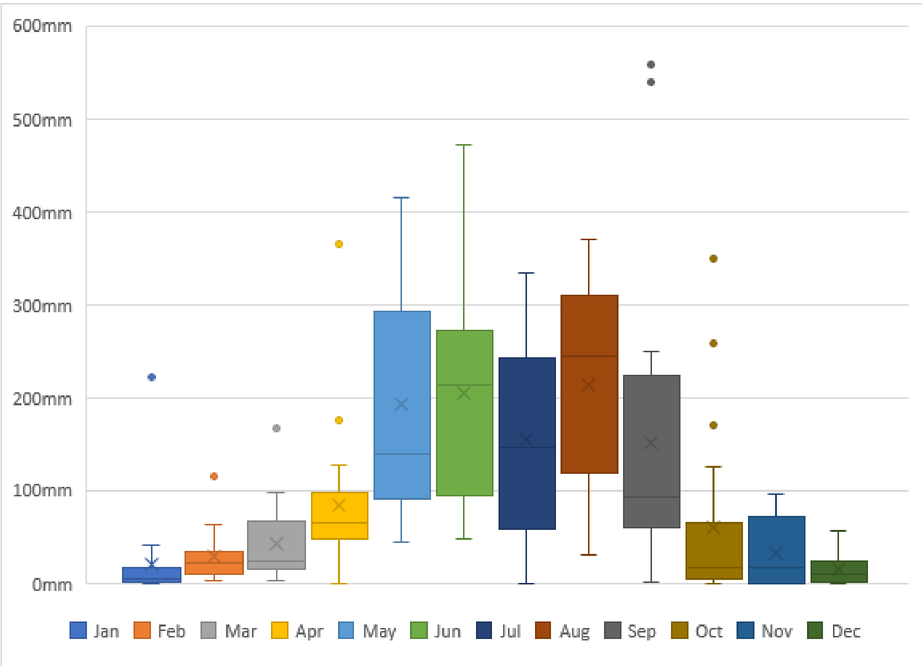
(j) Tai Mo Shan



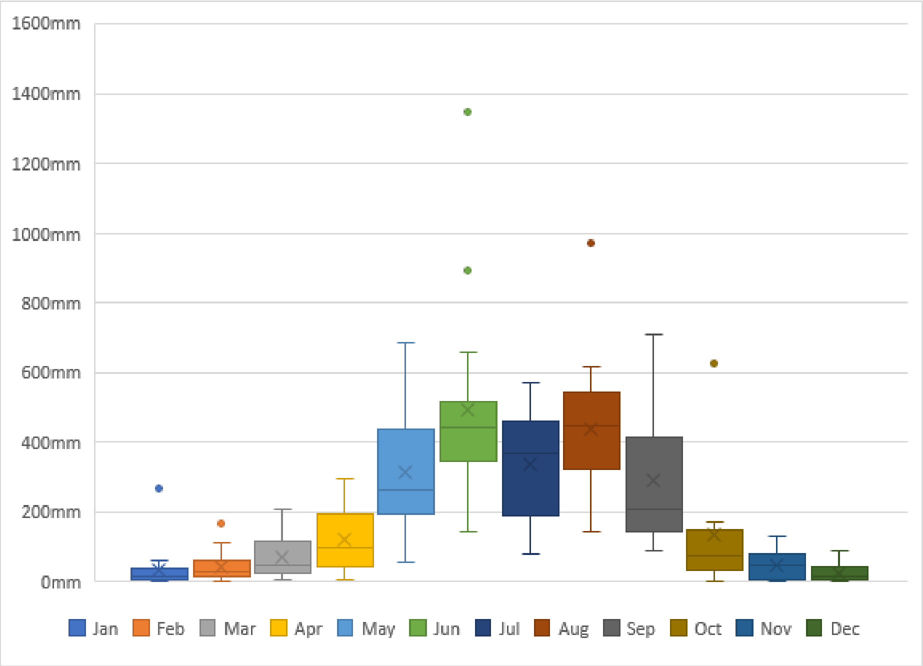
(k) Tate's Cairn



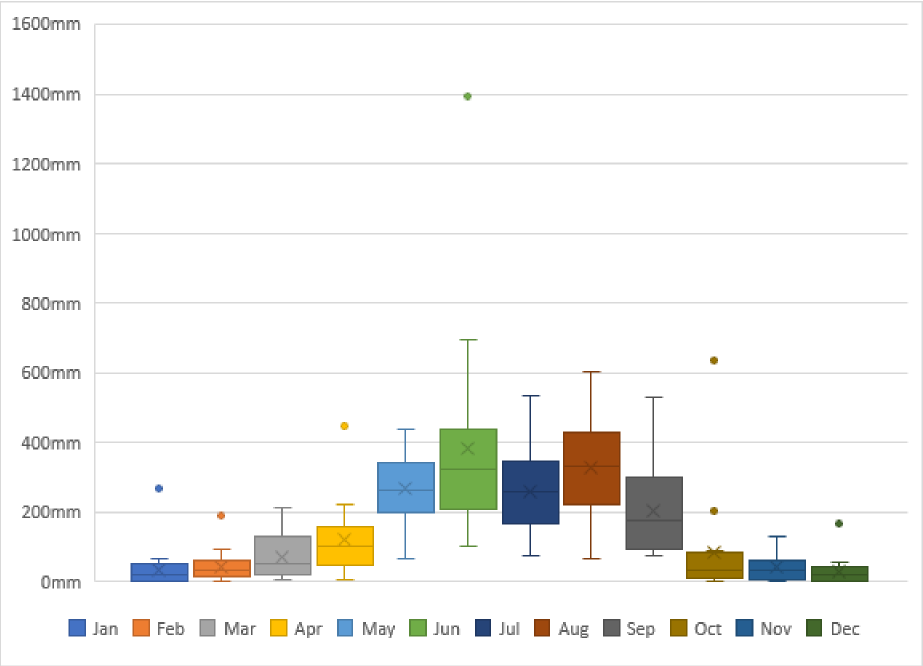
(l) Tseung Kwan O



(m) Waglan Island



(n) Hong Kong Observatory



(o) Hong Kong International Airport

Figure 4.1.3: Box plot of monthly rainfall at each rainfall station over the last 20 years (2003-2022)

.

From the monthly rainfall box plots of the fifteen stations, it is evident that Hong Kong experiences heavy rainfall from May to September, with less rainfall in the remaining months. This pattern aligns well with the description of Hong Kong's climate provided by the Hong Kong Observatory.

Table 4.1.3: Comparison of total values for February with and without leap years

	Cheung Chau			Ching Pak House(Tsing Yi)			King's Park		
	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error
mean	38.8000	39.4667	1.72%	35.3250	34.0667	3.56%	41.8300	41.3333	1.19%
std	40.6950	44.7882		33.6598	37.2773		39.8426	44.7120	
	Lau Fau Shan			Pak Tam Chung (Tsak Yue Wu)			Sha Tin		
	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error
mean	38.8750	38.3000	1.48%	39.8100	36.5800	8.11%	41.2500	39.6333	3.92%
std	37.7177	43.0855		42.2463	47.3814		42.3658	48.3964	
	Shek Kong			Ta Kwu Ling			Tai Mei Tuk		
	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error
mean	38.2250	37.9667	0.68%	37.3250	36.2667	2.84%	29.1500	27.6667	5.09%
std	37.1398	42.2589		36.2823	41.5238		39.0391	43.2646	
	Tai Mo Shan			Tate's Cairn			Tseung Kwan O		
	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error
mean	50.2000	48.4000	3.59%	47.4500	47.0333	0.88%	44.0000	43.7667	0.53%
std	34.9723	38.5557		37.6400	42.0632		40.5370	45.8828	
	Waglan Island			Hong Kong Observatory			Hong Kong International Airport		
	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error	ALL	Excluding leap years	Relative Error
mean	29.1250	30.8333	5.87%	41.3450	40.9000	1.08%	43.7500	44.0400	0.66%
std	26.3063	29.6775		40.3087	44.9437		42.2036	48.0013	

In Table 4.1.3, when comparing the totals for February in leap years to February in non-leap years, the difference is not significant, and therefore, no special treatment is applied.

Overall, the bins are relatively long, signifying a substantial interquartile range (IQR). A lengthy box indicates a large variation in the middle 50% of the range, suggesting that the data is widely spread.

Similarly, the whiskers are relatively long, depicting the overall spread of the data. Longer whiskers, reaching the maximum and minimum values, suggest increased dispersion and a wider range of numerical variation.

The presence of outlier in a box plot refers to data points that significantly deviate from the box and whiskers. If outlier are observed, it indicates the presence of extreme values in the data, contributing to an overall increase in dispersion.

4.2 Performance of mixed exponential distributions

After parameter fitting, we employed Relative Error (RE) to assess the consistency of the fitted parameters with the observed parameters of the mixed exponential distribution model. This assessment aims to compare the relative difference between the fitted parameters and the actual observed parameters, and is quantified by the RE. The calculation of RE is presented as a percentage, and fitting results are deemed excellent when the RE is less than 10%, highlighted in bold to emphasize this superiority.

Referring to the Table 4.2.1, the *p*-value, μ_1 and μ_2 in the fitted parameters represent the percentage of rainfall and the mean of the rainfall sizes, respectively. The results indicate that the fitted model satisfactorily describes the proportion of rainfall (*p*-value). However, there is a degree of underfitting in revealing the mean values of large and small rainfalls (μ_1 and μ_2). Specifically, the mixed exponential distribution model exhibits relatively low fit in terms of rainfall size, suggesting limitations in modeling the rainfall characteristics of the study area.

This finding poses challenges to the applicability of the model, especially in accurately simulating rainfall intensity. It provides valuable insights for future model improvements or consideration of alternative distribution forms to better capture the observed data characteristics. This academic assessment serves as useful guidance for further model refinement and optimization.

Table 4.2.1: Comparison of the parameters of the mixing exponential distribution

(a) Cheung Chau

Month	op	p	re_p	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9435	0.38	59.73%	0.0701	0.8144	1061.95%	13.4000	0.5741	95.72%
Feb	0.9062	0.98	**8.14%**	0.1104	0.7966	621.88%	13.5755	7.1613	47.25%
Mar	0.8677	0.96	10.63%	0.1403	0.6277	347.32%	12.6402	6.6064	47.73%
Apr	0.8217	0.92	11.97%	0.1379	0.3453	150.31%	18.7243	9.8471	47.41%
May	0.7161	0.81	13.11%	0.1351	0.1674	23.90%	29.1506	18.6005	36.19%
Jun	0.5917	0.63	**6.48%**	0.2113	0.2412	14.18%	29.2327	25.5408	12.63%
Jul	0.6694	0.73	**9.06%**	0.1819	0.2343	28.81%	23.2561	17.3931	25.21%
Aug	0.6177	0.68	10.08%	0.1841	0.1868	**1.49%**	24.1392	17.8166	26.19%
Sep	0.7500	0.85	13.33%	0.1389	0.2303	65.84%	22.6167	13.7869	39.04%
Oct	0.9016	0.98	**8.69%**	0.0680	0.4312	534.27%	25.1557	13.1701	47.65%
Nov	0.9167	0.98	**6.91%**	0.0618	0.9957	1510.68%	12.4500	5.4963	55.85%
Dec	0.9419	0.12	87.26%	0.0591	0.5229	785.21%	8.1389	0.3395	95.83%

(b) Ching Pak House (Tsing Yi)

Month	op	p	re_p	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9306	0.99	**6.38%**	0.0676	1.0059	1388.26%	14.3605	7.4705	47.98%
Feb	0.9062	0.98	**8.14%**	0.1113	0.8753	686.24%	12.2547	6.5419	46.62%
Mar	0.8758	0.96	**9.61%**	0.1335	0.5737	329.67%	14.9870	7.6209	49.15%
Apr	0.8133	0.92	13.11%	0.1393	0.3457	148.07%	17.9107	9.9393	44.51%
May	0.6968	0.76	**9.07%**	0.2153	0.1928	10.42%	31.3032	23.7534	24.12%
Jun	0.5567	0.58	**4.19%**	0.2560	0.2866	11.95%	31.9981	29.2961	**8.44%**
Jul	0.6161	0.66	**7.12%**	0.2212	0.2688	21.50%	22.9832	19.1271	16.78%
Aug	0.5597	0.6	**7.20%**	0.1758	0.2238	27.31%	25.1905	21.2736	15.55%
Sep	0.7100	0.79	11.27%	0.1491	0.2135	43.20%	24.1695	16.2567	32.74%
Oct	0.8935	0.98	**9.68%**	0.0442	0.4190	847.38%	24.2273	13.9733	42.32%
Nov	0.9217	0.48	47.92%	0.0669	0.8753	1208.26%	10.5000	0.6458	93.85%
Dec	0.9274	0.19	79.51%	0.0539	0.6283	1065.45%	8.0556	0.4102	94.91%

(c) King's Park

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9258	0.99	**6.93%**	0.0740	1.0034	1255.20%	12.7370	2.6916	78.87%
Feb	0.8991	0.97	**7.88%**	0.1368	0.7552	451.97%	13.4579	6.5405	51.40%
Mar	0.8661	0.95	**9.68%**	0.1538	0.5346	247.53%	14.9771	7.2210	51.79%
Apr	0.8150	0.91	11.66%	0.1798	0.3326	85.02%	19.1829	10.6572	44.44%
May	0.6726	0.73	**8.54%**	0.2182	0.1973	**9.57%**	29.6882	22.8450	23.05%
Jun	0.5350	0.55	**2.80%**	0.2673	0.3179	18.93%	33.8738	31.4759	**7.08%**
Jul	0.5968	0.64	**7.24%**	0.2414	0.2101	12.95%	25.6944	20.7178	19.37%
Aug	0.5452	0.6	10.06%	0.1828	0.1405	23.18%	30.6599	24.4325	20.31%
Sep	0.6750	0.74	**9.63%**	0.2007	0.1928	**3.94%**	29.2503	22.3246	23.68%
Oct	0.8823	0.96	**8.81%**	0.1247	0.2686	115.44%	37.0890	22.5766	39.13%
Nov	0.9050	0.98	**8.29%**	0.1013	0.6817	573.02%	15.7228	7.3868	53.02%
Dec	0.9258	0.3	67.60%	0.0676	0.7475	1005.76%	9.3326	0.5052	94.59%

(d) Lau Fau Shan

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9274	0.99	**6.75%**	0.0696	0.9968	1332.83%	13.8778	7.8546	43.40%
Feb	0.8885	0.97	**9.17%**	0.1155	0.8119	602.72%	11.4206	6.0461	47.06%
Mar	0.8694	0.96	10.43%	0.1429	0.5431	280.15%	15.0370	8.0241	46.64%
Apr	0.8100	0.92	13.58%	0.1389	0.3773	171.66%	15.9561	8.8224	44.71%
May	0.7129	0.78	**9.41%**	0.1561	0.2267	45.25%	25.9382	18.7183	27.83%
Jun	0.5950	0.64	**7.56%**	0.1751	0.2192	25.21%	26.8045	22.3341	16.68%
Jul	0.6694	0.73	**9.06%**	0.2024	0.2686	32.72%	19.9024	14.8103	25.59%
Aug	0.6097	0.65	**6.61%**	0.2103	0.2713	29.01%	22.8430	18.9964	16.84%
Sep	0.7367	0.82	11.31%	0.1697	0.2655	56.49%	19.6930	12.3489	37.29%
Oct	0.9129	0.98	**7.35%**	0.0424	0.5256	1139.55%	23.6944	11.8950	49.80%
Nov	0.9100	0.98	**7.69%**	0.0723	0.9478	1210.10%	12.1111	6.0921	49.70%
Dec	0.9323	0.33	64.60%	0.0554	0.7776	1304.59%	10.8333	0.5390	95.02%

(e) Pak Tam Chung (Tsak Yue Wu)

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9210	0.98	**6.41%**	0.0841	0.9126	985.63%	14.2143	6.3483	55.34%
Feb	0.8850	0.97	**9.61%**	0.1230	0.7892	541.59%	11.3031	6.0013	46.91%
Mar	0.8403	0.93	10.67%	0.2054	0.4799	133.67%	14.7374	8.3988	43.01%
Apr	0.7800	0.87	11.54%	0.1774	0.3285	85.25%	18.0379	11.2195	37.80%
May	0.6258	0.66	**5.46%**	0.2075	0.2343	12.93%	30.9591	26.1691	15.47%
Jun	0.5300	0.55	**3.77%**	0.2248	0.3226	43.49%	34.0674	32.0578	**5.90%**
Jul	0.6048	0.65	**7.47%**	0.1747	0.2141	22.58%	26.8245	21.9139	18.31%
Aug	0.6065	0.64	**5.53%**	0.2194	0.2343	**6.79%**	29.4324	24.9579	15.20%
Sep	0.7067	0.81	14.62%	0.1509	0.1607	**6.43%**	27.0199	15.7394	41.75%
Oct	0.8855	0.96	**8.42%**	0.1084	0.3103	186.27%	33.0282	19.6007	40.65%
Nov	0.9000	0.98	**8.89%**	0.0796	0.6457	710.86%	16.2750	9.0697	44.27%
Dec	0.9306	0.39	58.09%	0.0849	0.8191	864.59%	10.7907	0.5759	94.66%

(f) Sha Tin

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9242	0.99	**7.12%**	0.0785	0.9769	1143.96%	13.3617	7.2111	46.03%
Feb	0.8885	0.97	**9.17%**	0.1404	0.7593	440.66%	11.9762	6.0890	49.16%
Mar	0.8516	0.94	10.38%	0.1742	0.4850	178.35%	15.3424	8.1143	47.11%
Apr	0.7983	0.88	10.23%	0.1962	0.3266	66.43%	19.4669	11.5673	40.58%
May	0.6468	0.69	**6.68%**	0.2244	0.2206	**1.70%**	30.8196	25.4969	17.27%
Jun	0.5250	0.54	**2.86%**	0.3206	0.3215	**0.26%**	33.4860	31.2923	**6.55%**
Jul	0.6016	0.64	**6.38%**	0.2520	0.2271	**9.87%**	28.1275	23.7213	15.67%
Aug	0.5823	0.61	**4.76%**	0.2299	0.2670	16.14%	31.6390	28.3751	10.32%
Sep	0.6950	0.78	12.23%	0.1667	0.1711	**2.63%**	27.6612	18.1474	34.39%
Oct	0.8710	0.96	10.22%	0.0880	0.2606	196.28%	33.0688	16.4981	50.11%
Nov	0.8883	0.97	**9.19%**	0.1107	0.6527	489.60%	14.5000	7.7084	46.84%
Dec	0.9210	0.4	56.57%	0.0674	0.8287	1129.10%	9.8061	0.5907	93.98%

(g) Shek Kong

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9242	0.99	**7.12%**	0.0663	0.9769	1373.12%	13.5106	7.2111	46.63%
Feb	0.9027	0.98	**8.57%**	0.1176	0.8119	590.08%	12.8091	7.2995	43.01%
Mar	0.8677	0.96	10.63%	0.1487	0.5220	251.05%	15.4695	8.3980	45.71%
Apr	0.8150	0.91	11.66%	0.1892	0.3293	74.10%	18.9459	9.9549	47.46%
May	0.6758	0.74	**9.50%**	0.1874	0.1909	**1.91%**	29.0572	21.8120	24.93%
Jun	0.5467	0.57	**4.27%**	0.2561	0.2673	**4.37%**	29.1728	26.1222	10.46%
Jul	0.6097	0.66	**8.25%**	0.2063	0.1976	**4.25%**	24.9876	19.2312	23.04%
Aug	0.6306	0.68	**7.83%**	0.1829	0.2108	15.26%	28.2205	22.8516	19.02%
Sep	0.7167	0.77	**7.44%**	0.1512	0.2397	58.54%	27.8118	20.7622	25.35%
Oct	0.8952	0.97	**8.36%**	0.0613	0.3554	480.09%	30.2923	16.7034	44.86%
Nov	0.9083	0.98	**7.89%**	0.0596	0.9857	1552.90%	11.4909	5.6629	50.72%
Dec	0.9306	0.38	59.17%	0.0537	0.8144	1415.75%	11.1395	0.5741	94.85%

(h) Ta Kwu Ling

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9306	0.99	**6.38%**	0.0763	1.0146	1230.49%	14.0581	7.0203	50.06%
Feb	0.8991	0.97	**7.88%**	0.1309	0.8456	545.96%	11.9298	5.8038	51.35%
Mar	0.8677	0.96	10.63%	0.1524	0.5802	280.64%	13.6220	6.9790	48.77%
Apr	0.8117	0.91	12.11%	0.1674	0.3323	98.58%	18.8850	10.6024	43.86%
May	0.6645	0.74	11.36%	0.1699	0.1744	**2.63%**	25.6514	17.2105	32.91%
Jun	0.5633	0.59	**4.73%**	0.2944	0.2622	10.94%	29.6412	26.4850	10.65%
Jul	0.5887	0.64	**8.71%**	0.2151	0.1938	**9.87%**	23.2569	17.7506	23.68%
Aug	0.5758	0.61	**5.94%**	0.2045	0.2430	18.81%	28.3593	24.7097	12.87%
Sep	0.7100	0.79	11.27%	0.1491	0.2107	41.32%	23.8477	15.5780	34.68%
Oct	0.8919	0.98	**9.87%**	0.0778	0.3109	299.78%	31.3209	15.0759	51.87%
Nov	0.9017	0.98	**8.69%**	0.0804	0.8357	939.38%	12.5424	6.6605	46.90%
Dec	0.9258	0.33	64.36%	0.0497	0.7760	1462.96%	9.9457	0.5353	94.62%

(i) Tai Mei Tuk

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9419	0.37	60.72%	0.0745	0.8056	981.50%	12.8056	0.5626	95.61%
Feb	0.9221	0.99	**7.36%**	0.1468	1.0215	595.72%	11.5114	6.2733	45.50%
Mar	0.8839	0.97	**9.74%**	0.1679	0.6241	271.73%	14.0764	7.6253	45.83%
Apr	0.8083	0.9	11.34%	0.1629	0.4024	147.01%	15.4609	8.5814	44.50%
May	0.6903	0.78	12.99%	0.1717	0.1704	**0.76%**	27.1328	17.9274	33.93%
Jun	0.5817	0.62	**6.59%**	0.1777	0.2335	31.41%	27.6554	23.7281	14.20%
Jul	0.6500	0.72	10.77%	0.1935	0.1780	**8.05%**	25.0645	17.3311	30.85%
Aug	0.6323	0.69	**9.13%**	0.2143	0.1976	**7.80%**	26.4934	20.5990	22.25%
Sep	0.7350	0.85	15.65%	0.1156	0.2222	92.17%	20.6447	11.5402	44.10%
Oct	0.8935	0.97	**8.56%**	0.0749	0.3723	396.99%	28.5152	16.5642	41.91%
Nov	0.9050	0.98	**8.29%**	0.0783	0.9849	1158.40%	10.9035	5.5838	48.79%
Dec	0.9435	0.1	89.40%	0.0650	0.4790	637.46%	7.4857	0.3057	95.92%

(j) Tai Mo Shan

Month	**op**	**p**	**re_op**	**o μ_1**	μ_1	**re_μ_1**	**o μ_2**	μ_2	**re_μ_2**
Jan	0.9048	0.97	**7.20%**	0.2959	0.8852	199.14%	10.4746	5.6224	46.32%
Feb	0.8177	0.89	**8.84%**	0.4491	0.7108	58.26%	7.7330	4.7716	38.30%
Mar	0.7774	0.82	**5.48%**	0.5290	0.5806	**9.74%**	11.8804	9.1291	23.16%
Apr	0.7083	0.74	**4.47%**	0.4894	0.5198	**6.22%**	15.0143	12.7012	15.41%
May	0.5919	0.6	**1.36%**	0.5054	0.5496	**8.73%**	26.2016	24.7483	**5.55%**
Jun	0.5300	0.54	**1.89%**	0.4796	0.5557	15.88%	29.0195	28.0905	**3.20%**
Jul	0.5629	0.58	**3.04%**	0.4040	0.3619	10.42%	27.0793	24.9074	**8.02%**
Aug	0.5419	0.56	**3.33%**	0.3720	0.4179	12.34%	29.9718	28.6152	**4.53%**
Sep	0.6583	0.7	**6.33%**	0.2987	0.2715	**9.12%**	25.9098	21.5689	16.75%
Oct	0.8661	0.94	**8.53%**	0.2477	0.3194	28.94%	28.2169	17.4750	38.07%
Nov	0.8767	0.95	**8.37%**	0.2795	0.8332	198.12%	9.0878	4.5119	50.35%
Dec	0.9161	0.4	56.34%	0.1831	0.8199	347.82%	7.8750	0.5673	92.80%

(k) Tate's Cairn

Month	**op**	**p**	**re_op**	**o μ_1**	μ_1	**re_μ_1**	**o μ_2**	μ_2	**re_μ_2**
Jan	0.9210	0.98	**6.41%**	0.1848	0.9042	389.40%	13.1020	6.0927	53.50%
Feb	0.8531	0.93	**9.01%**	0.3122	0.7055	125.93%	9.6205	5.1620	46.34%
Mar	0.8113	0.89	**9.70%**	0.3559	0.5335	49.92%	11.3205	6.8797	39.23%
Apr	0.7500	0.81	**8.00%**	0.2811	0.3634	29.27%	17.3267	12.1766	29.72%
May	0.5952	0.62	**4.17%**	0.2995	0.3419	14.16%	27.7072	25.2145	**9.00%**
Jun	0.5200	0.53	**1.92%**	0.3205	0.3518	**9.77%**	36.3003	34.2221	**5.73%**
Jul	0.5661	0.59	**4.22%**	0.2650	0.2546	**3.89%**	29.0000	25.4032	12.40%
Aug	0.5565	0.58	**4.23%**	0.2290	0.2713	18.46%	30.2909	27.2017	10.20%
Sep	0.6783	0.74	**9.09%**	0.1978	0.1887	**4.61%**	28.9093	21.1964	26.68%
Oct	0.8403	0.94	11.86%	0.1296	0.2422	86.96%	29.6566	16.0592	45.85%
Nov	0.8617	0.95	10.25%	0.1615	0.5759	256.57%	13.3494	6.7244	49.63%
Dec	0.9145	0.47	48.61%	0.1243	0.8663	596.69%	8.9057	0.6272	92.96%

(l) Tseung Kwan O

Month	**op**	**p**	**re_op**	**o μ_1**	μ_1	**re_μ_1**	**o μ_2**	μ_2	**re_μ_2**
Jan	0.9226	0.75	18.71%	0.0857	0.9716	1034.24%	11.6563	0.8382	92.81%
Feb	0.8903	0.97	**8.96%**	0.2018	0.7165	255.05%	12.5565	6.7881	45.94%
Mar	0.8242	0.91	10.41%	0.2270	0.4738	108.72%	13.9083	7.9071	43.15%
Apr	0.7683	0.85	10.63%	0.1855	0.3335	79.83%	17.3813	10.8046	37.84%
May	0.6419	0.69	**7.49%**	0.2073	0.2070	**0.12%**	28.5383	22.6413	20.66%
Jun	0.5117	0.53	**3.58%**	0.2524	0.3248	28.67%	32.9932	31.0833	**5.79%**
Jul	0.6065	0.65	**7.18%**	0.2194	0.2215	**0.97%**	27.8074	23.2643	16.34%
Aug	0.5581	0.59	**5.72%**	0.1734	0.2420	39.53%	27.1551	23.5587	13.24%
Sep	0.6983	0.77	10.26%	0.1635	0.1799	10.05%	29.4420	20.5047	30.36%
Oct	0.8613	0.95	10.30%	0.0946	0.3225	240.99%	25.2442	12.7421	49.52%
Nov	0.8933	0.97	**8.58%**	0.0849	0.5895	594.46%	17.1172	8.5417	50.10%
Dec	0.9274	0.24	74.12%	0.0748	0.6898	822.42%	8.6444	0.4590	94.69%

(m) Waglan Island

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9500	0.21	77.89%	0.0662	0.6531	886.32%	11.9355	0.4280	96.41%
Feb	0.9062	0.99	**9.25%**	0.1299	1.0207	685.83%	9.7358	6.1016	37.33%
Mar	0.8790	0.96	**9.21%**	0.1844	0.7944	330.79%	10.3867	5.2517	49.44%
Apr	0.8350	0.93	11.38%	0.1397	0.4201	200.69%	16.2677	8.3887	48.43%
May	0.7419	0.82	10.52%	0.1696	0.2499	47.40%	23.7250	16.4863	30.51%
Jun	0.6817	0.77	12.96%	0.1369	0.2055	50.11%	21.2277	13.4971	36.42%
Jul	0.7371	0.84	13.96%	0.1740	0.2515	44.55%	18.6411	10.5527	43.39%
Aug	0.6581	0.74	12.45%	0.1961	0.2079	6.01%	19.8868	12.9582	34.84%
Sep	0.7917	0.89	12.42%	0.1368	0.2538	85.43%	23.8400	14.2510	40.22%
Oct	0.9129	0.98	**7.35%**	0.0680	0.5613	725.17%	21.5093	9.4753	55.95%
Nov	0.9283	0.99	**6.64%**	0.0548	0.9464	1628.34%	15.0116	8.0597	46.31%
Dec	0.9516	0.11	88.44%	0.0746	0.4950	563.75%	8.8667	0.3121	96.48%

(n) Hong Kong Observatory

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9323	0.99	**6.19%**	0.0882	1.0148	1050.05%	13.9167	4.9390	64.51%
Feb	0.9097	0.98	**7.72%**	0.1307	0.7464	470.91%	14.8961	7.5290	49.46%
Mar	0.8629	0.95	10.09%	0.1521	0.5037	231.03%	15.7412	8.0631	48.78%
Apr	0.8150	0.9	10.43%	0.1859	0.3214	72.92%	20.9000	12.1810	41.72%
May	0.6758	0.73	**8.02%**	0.2198	0.2012	**8.45%**	30.5791	23.8319	22.06%
Jun	0.5283	0.54	**2.21%**	0.2845	0.3278	15.21%	34.4332	32.0524	**6.91%**
Jul	0.5903	0.62	**5.03%**	0.2549	0.3243	27.20%	26.0161	23.4123	10.01%
Aug	0.5419	0.56	**3.33%**	0.1964	0.2817	43.39%	30.5116	27.4875	**9.91%**
Sep	0.6800	0.74	**8.82%**	0.1900	0.1930	**1.62%**	29.7583	22.3783	24.80%
Oct	0.8726	0.97	11.16%	0.1092	0.2568	135.03%	32.8354	16.7074	49.12%
Nov	0.8967	0.97	**8.18%**	0.0950	0.6840	620.11%	14.9355	7.0103	53.06%
Dec	0.9306	0.3	67.76%	0.0773	0.7465	865.75%	9.8349	0.5031	94.88%

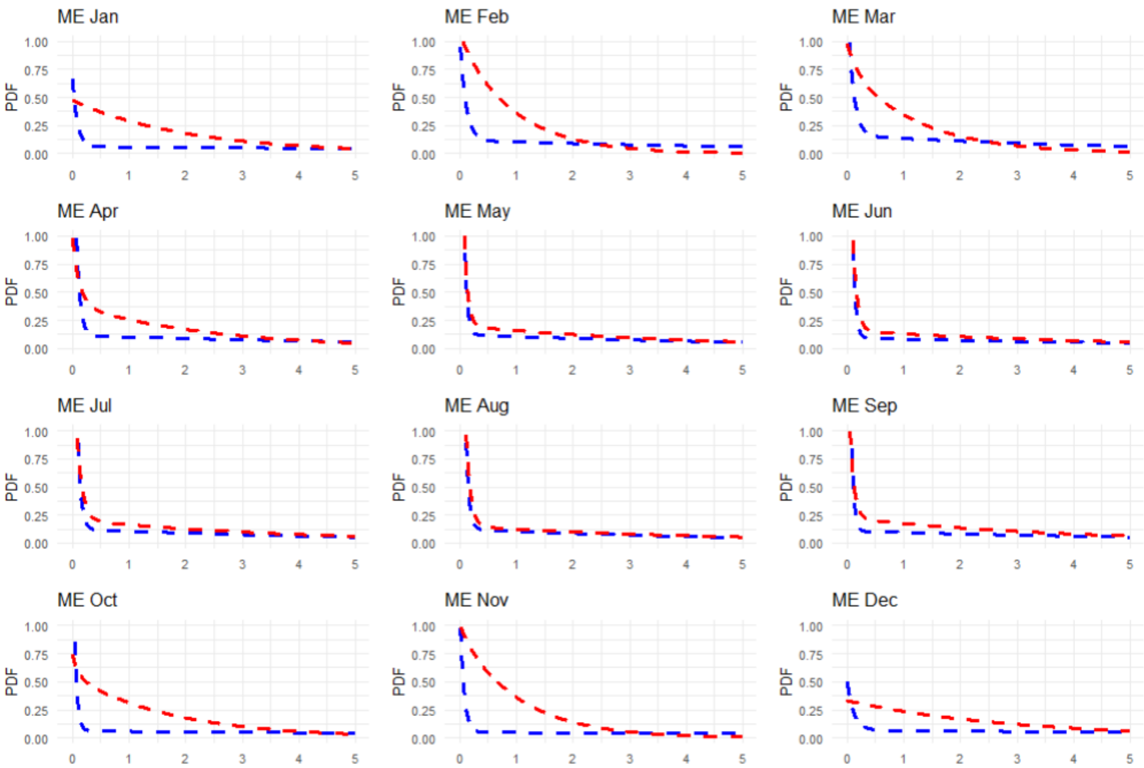
(o) Hong Kong International Airport

Month	op	p	re_op	o μ_1	μ_1	re_μ_1	o μ_2	μ_2	re_μ_2
Jan	0.9387	0.99	**5.46%**	0.0780	0.9425	1108.16%	17.3316	8.5032	50.94%
Feb	0.9080	0.98	**7.93%**	0.1193	0.7124	497.15%	15.6500	8.6510	44.72%
Mar	0.8710	0.96	10.22%	0.1583	0.5034	217.92%	16.6638	9.5234	42.85%
Apr	0.8100	0.9	11.11%	0.1381	0.3170	129.62%	20.2684	11.2396	44.55%
May	0.7097	0.79	11.32%	0.1682	0.1713	**1.87%**	29.0956	18.8354	35.26%
Jun	0.5917	0.62	**4.79%**	0.2620	0.2537	**3.18%**	30.6796	26.9426	12.18%
Jul	0.6419	0.69	**7.49%**	0.2324	0.2515	**8.20%**	22.9928	18.1876	20.90%
Aug	0.5984	0.64	**6.95%**	0.1814	0.2124	17.10%	26.1759	21.1341	19.26%
Sep	0.7567	0.85	12.33%	0.1667	0.2038	22.25%	27.4048	17.4982	36.15%
Oct	0.9177	0.99	**7.87%**	0.0673	0.3990	492.84%	31.5902	17.9420	43.20%
Nov	0.9133	0.98	**7.30%**	0.0639	0.8258	1193.03%	14.6962	7.2665	50.56%
Dec	0.9210	0.57	38.11%	0.0347	0.9169	2544.10%	11.3143	0.7080	93.74%

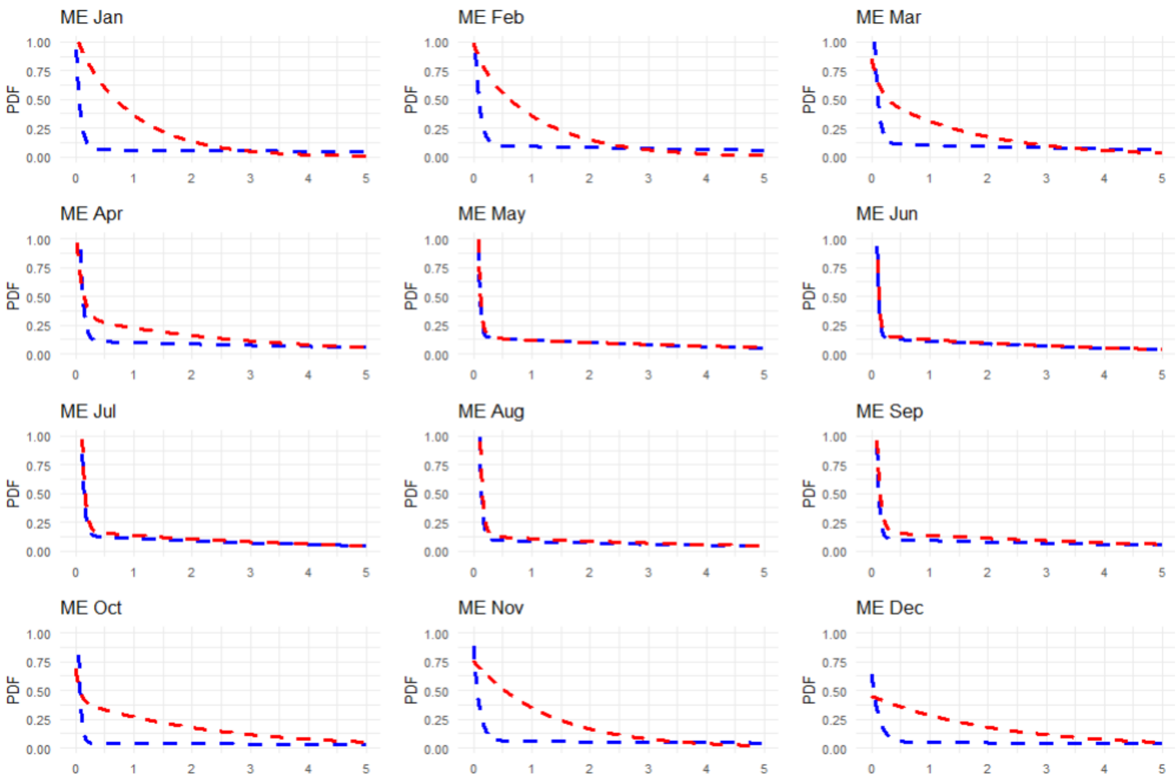
Note: Those with o represent those of observed data, those without o are fitted, and re is relative error.

From the Figure 4.2.1, we can see that during the monsoon season (May-September), the comparison between the observed data (blue line) and the model fitted data (red line) shows a high degree of fit, which suggests that the model has a better performance in simulating the rainfall events during the monsoon. However, during the non-monsoon season, there is a large discrepancy between the observed data and the model-fitted data, revealing a relatively poor fit of the model during these periods.

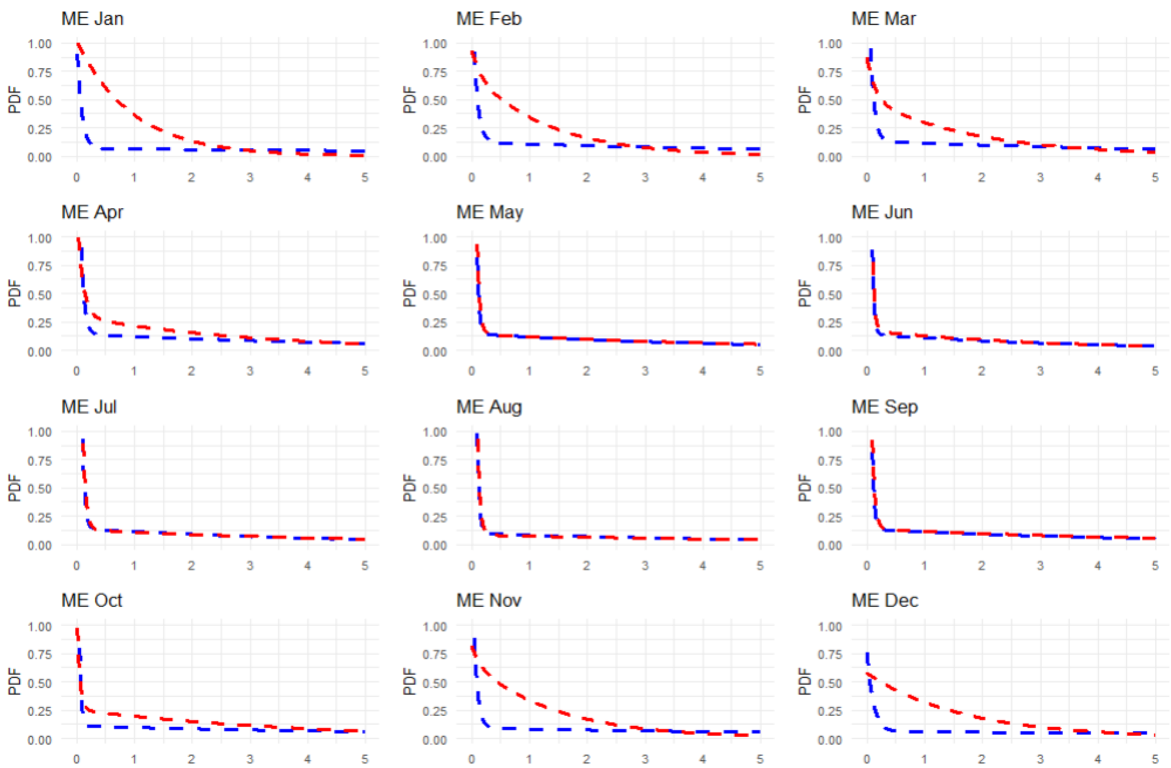
For the combined results from 15 rainfall stations, this trend is validated across the region, further emphasizing the sensitivity of model performance to seasonal variations.



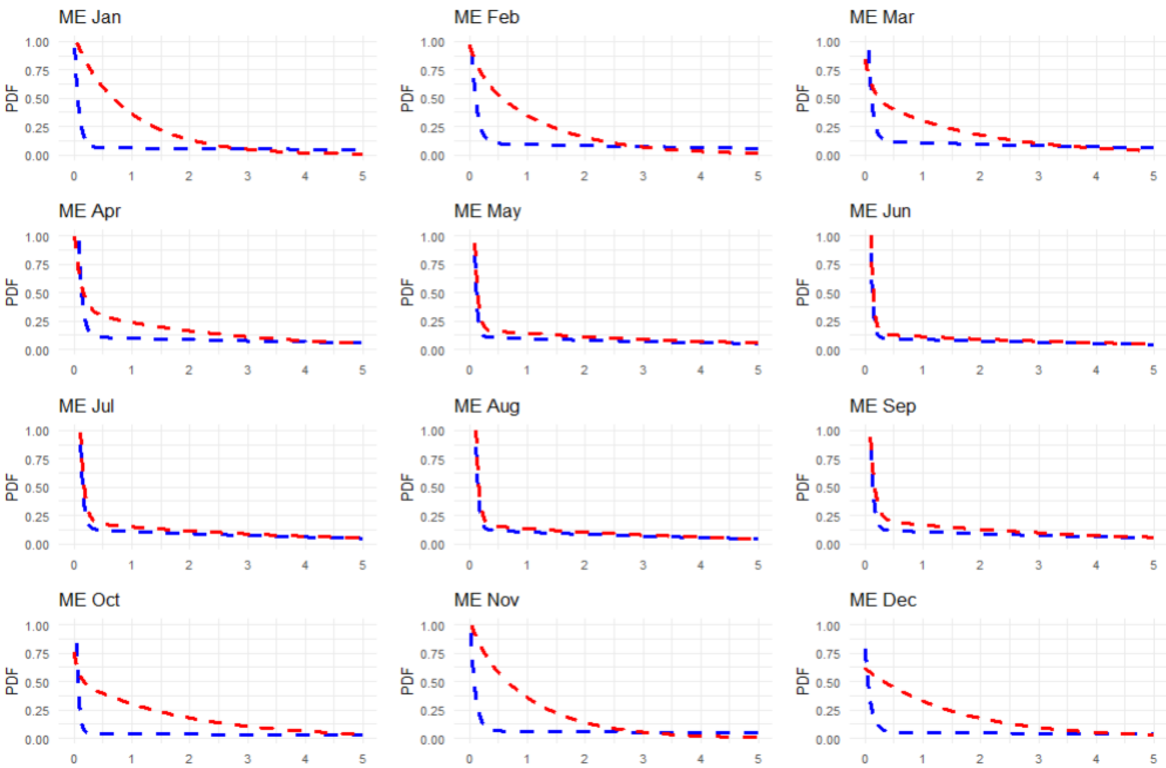
(a) Cheung Chau



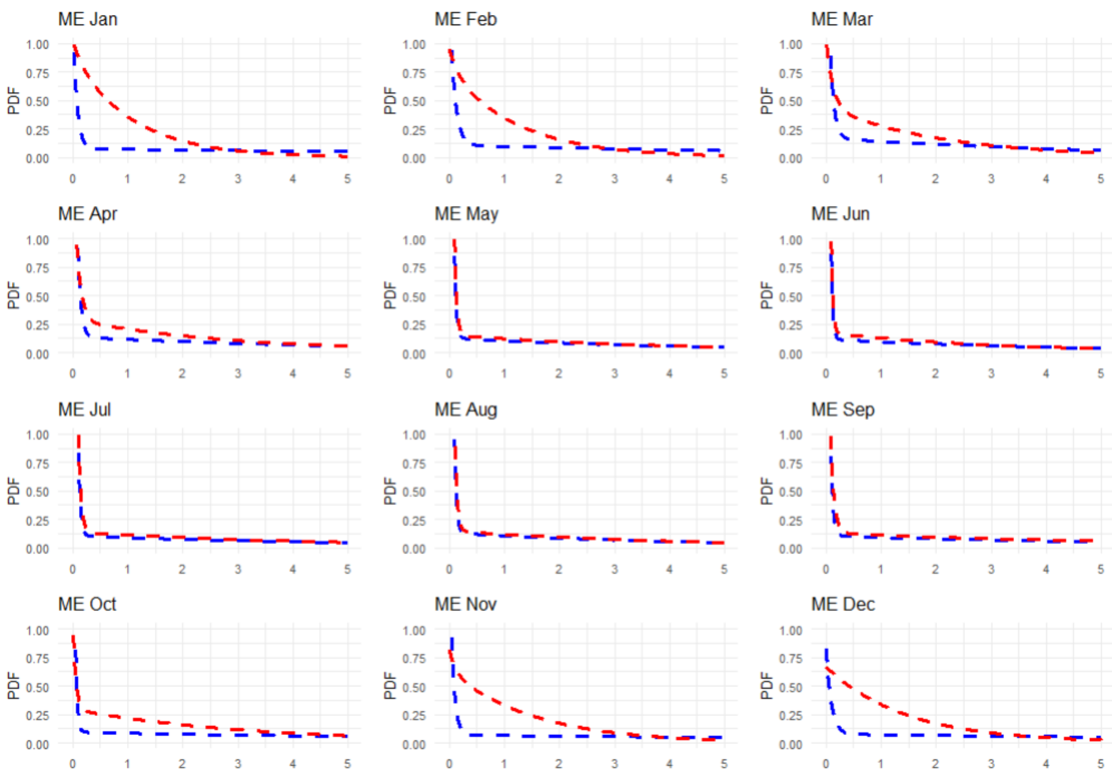
(b) Ching Pak House (Tsing Yi)



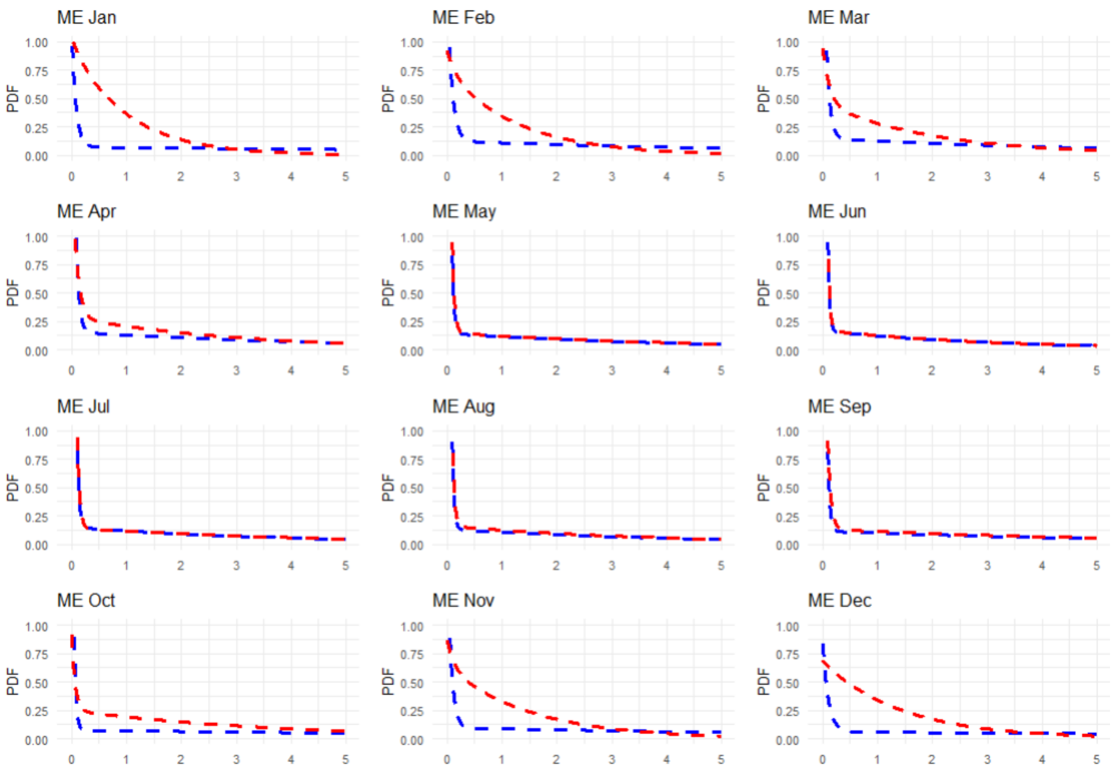
(c) King's Park



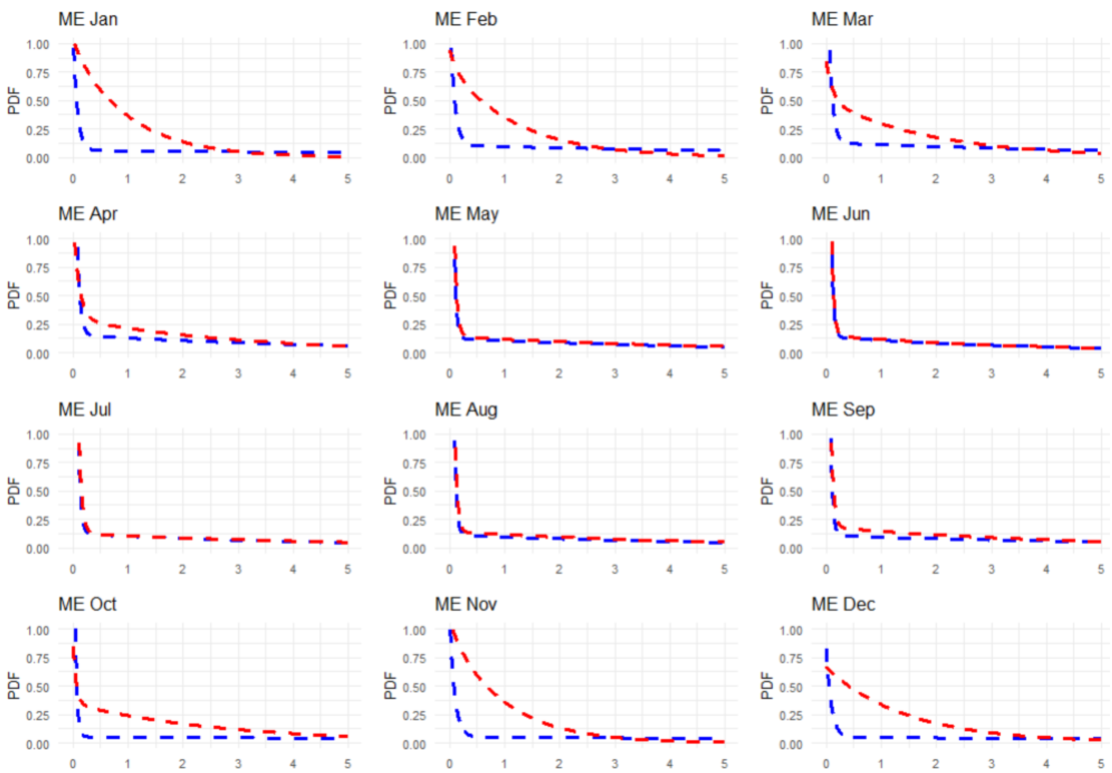
(d) Lau Fau Shan



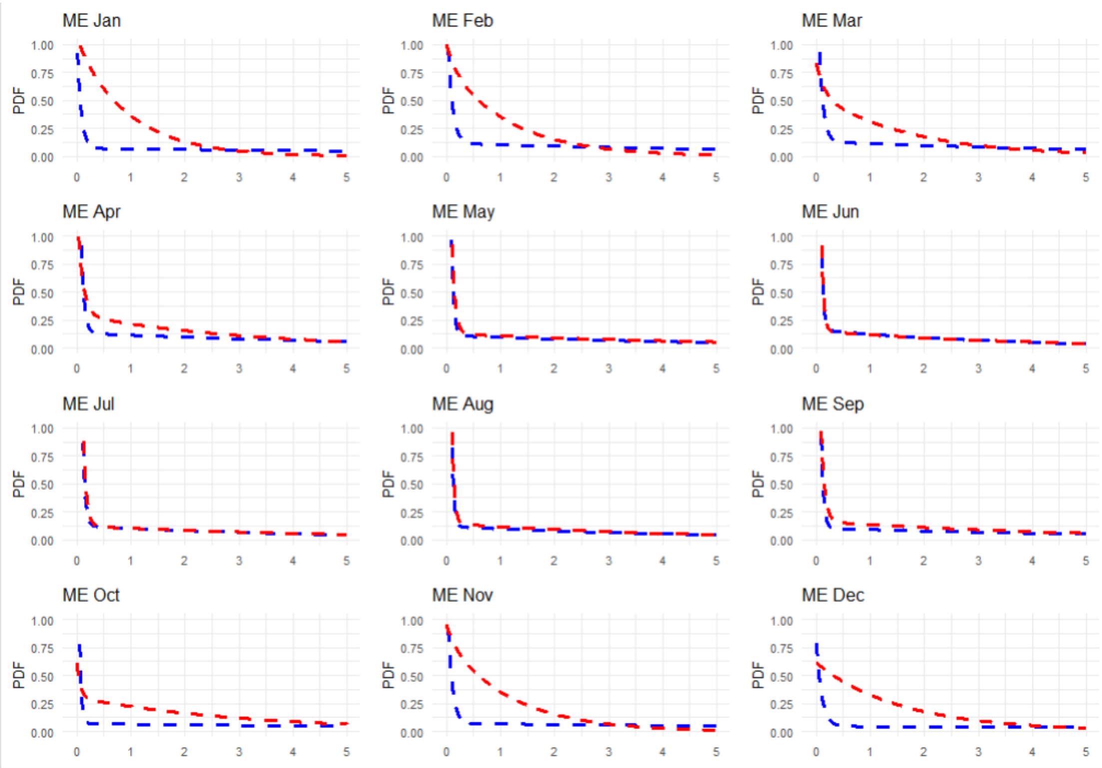
(e) Pak Tam Chung (Tsak Yue Wu)



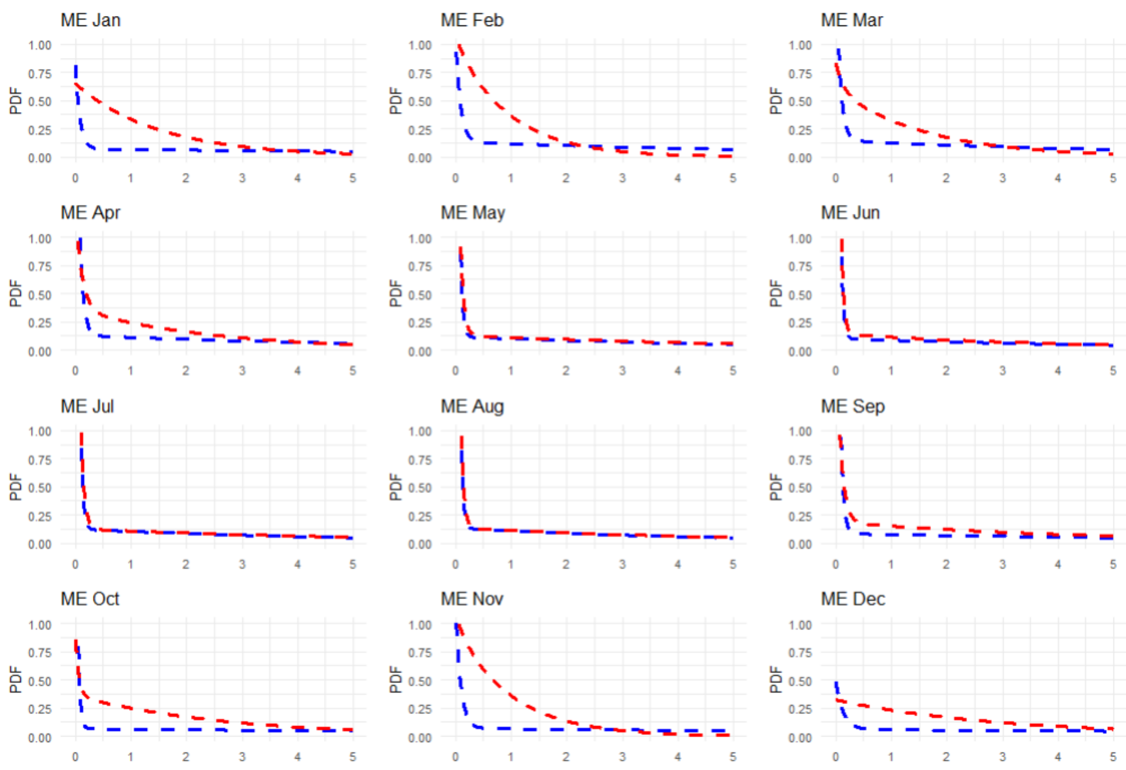
(f) Sha Tin



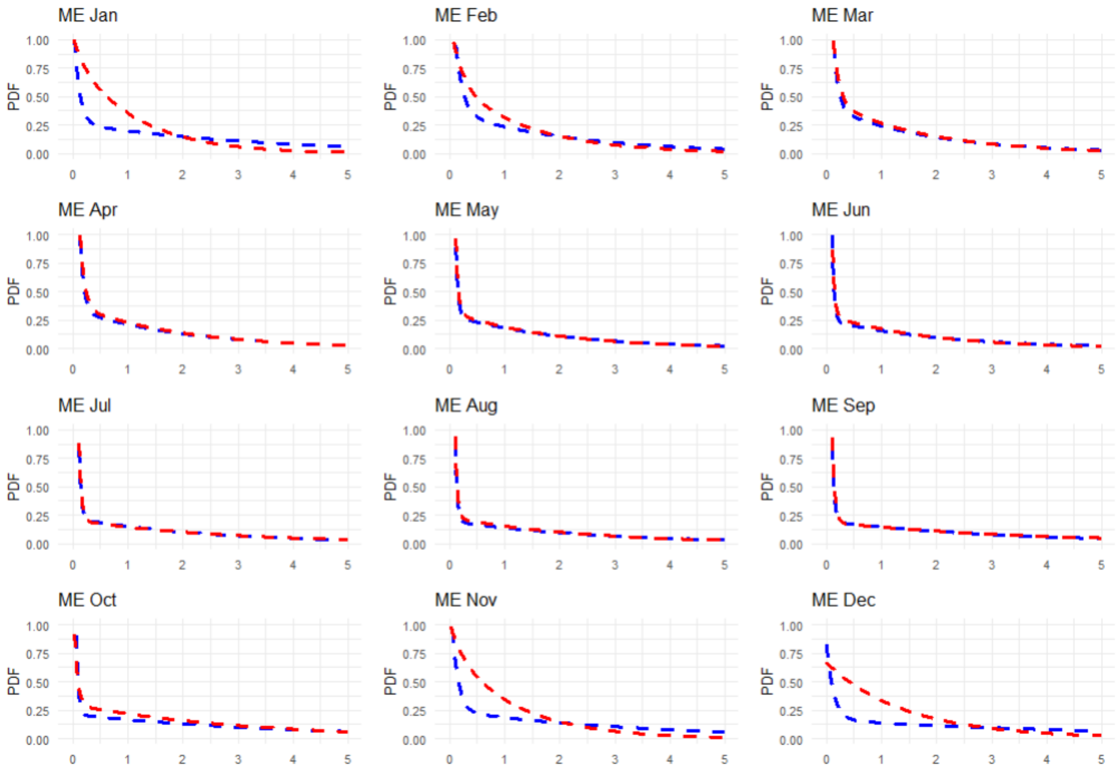
(g) Shek Kong



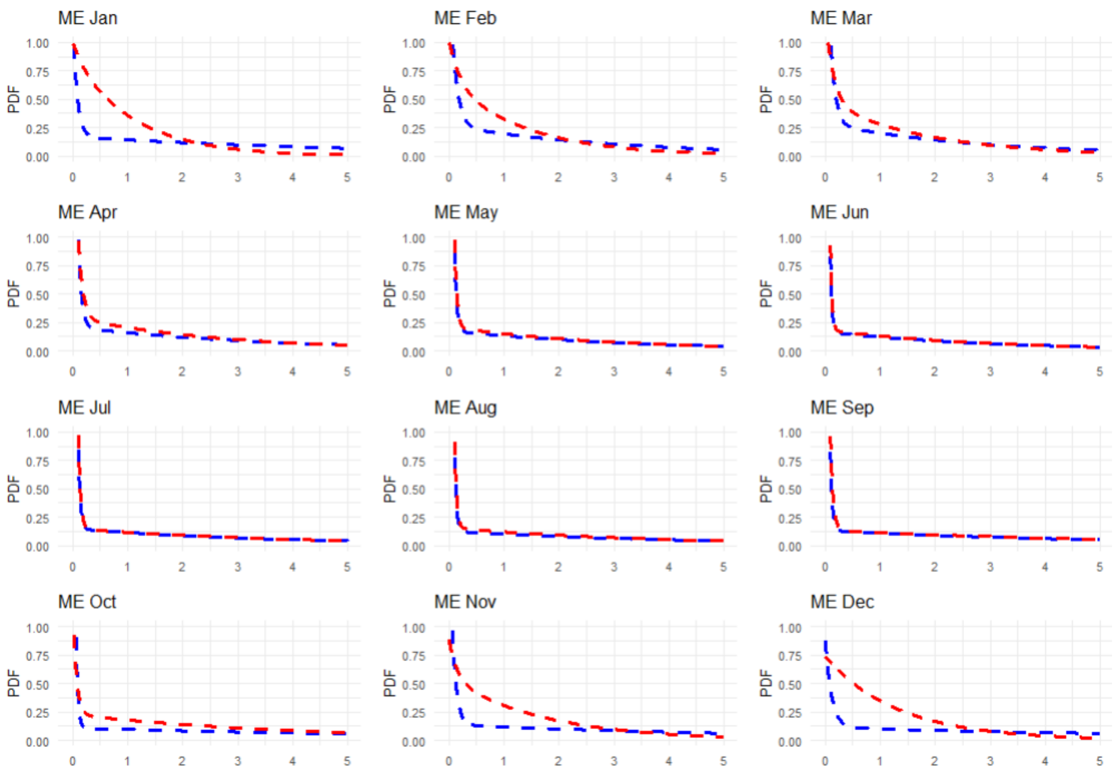
(h) Ta Kwu Ling



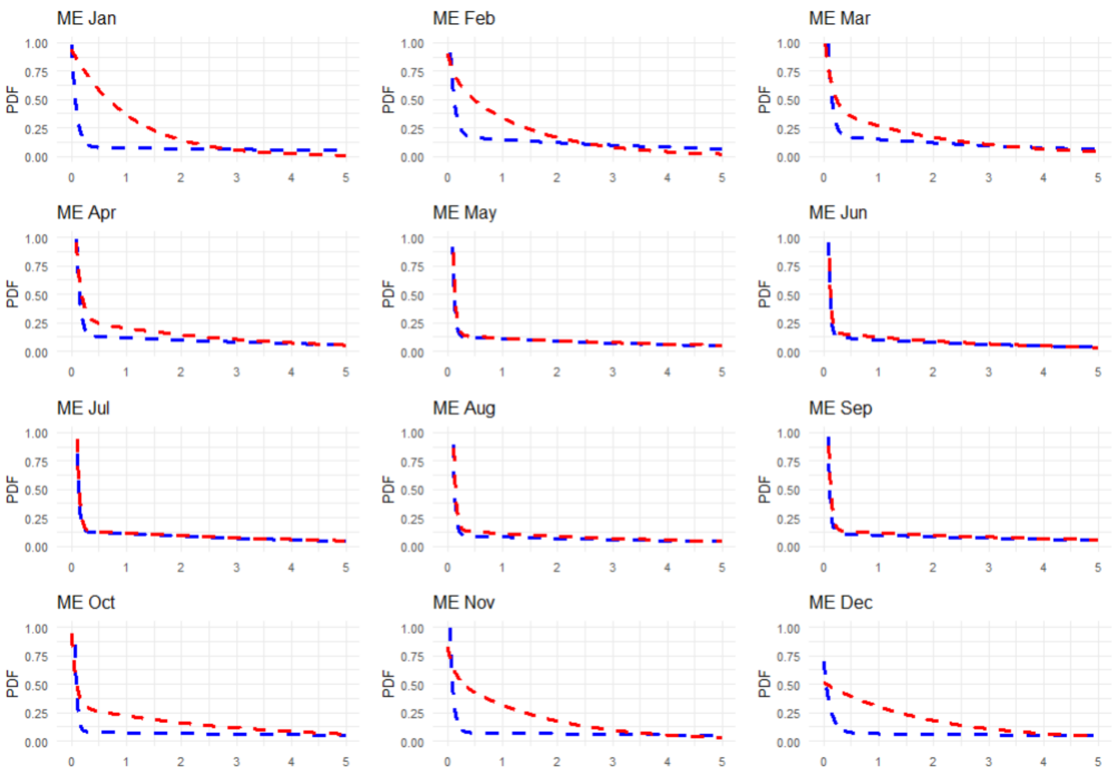
(i) Tai Mei Tuk



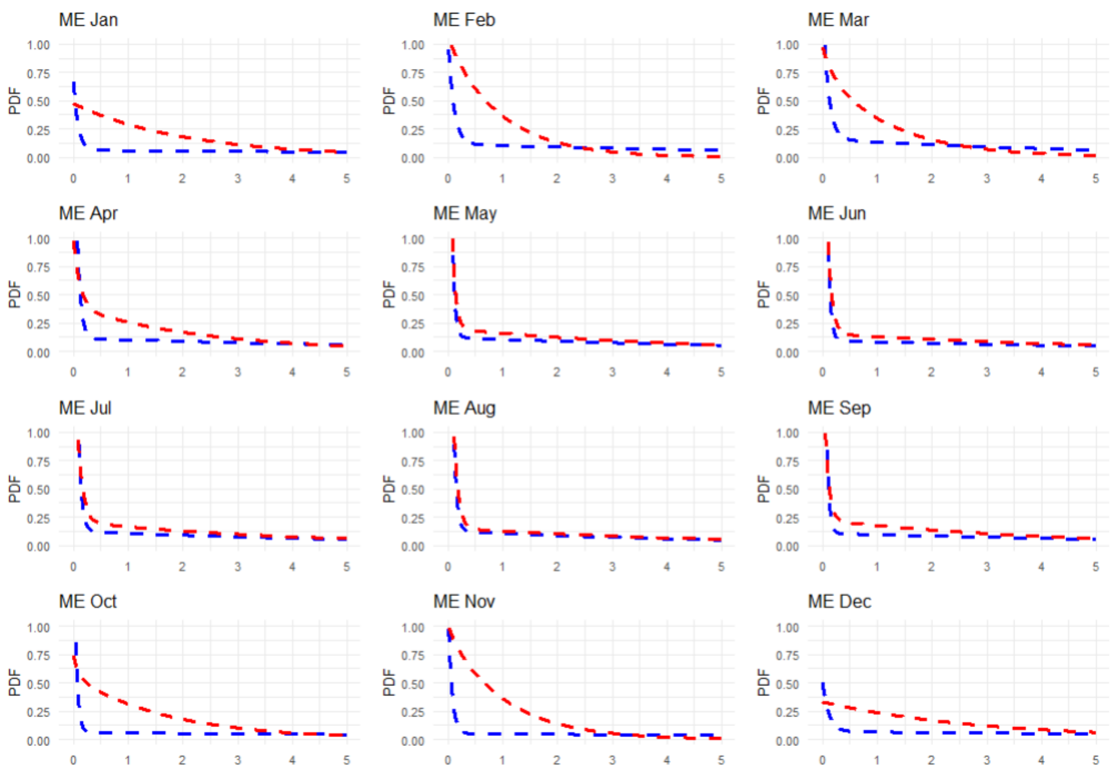
(k) Tai Mo Shan



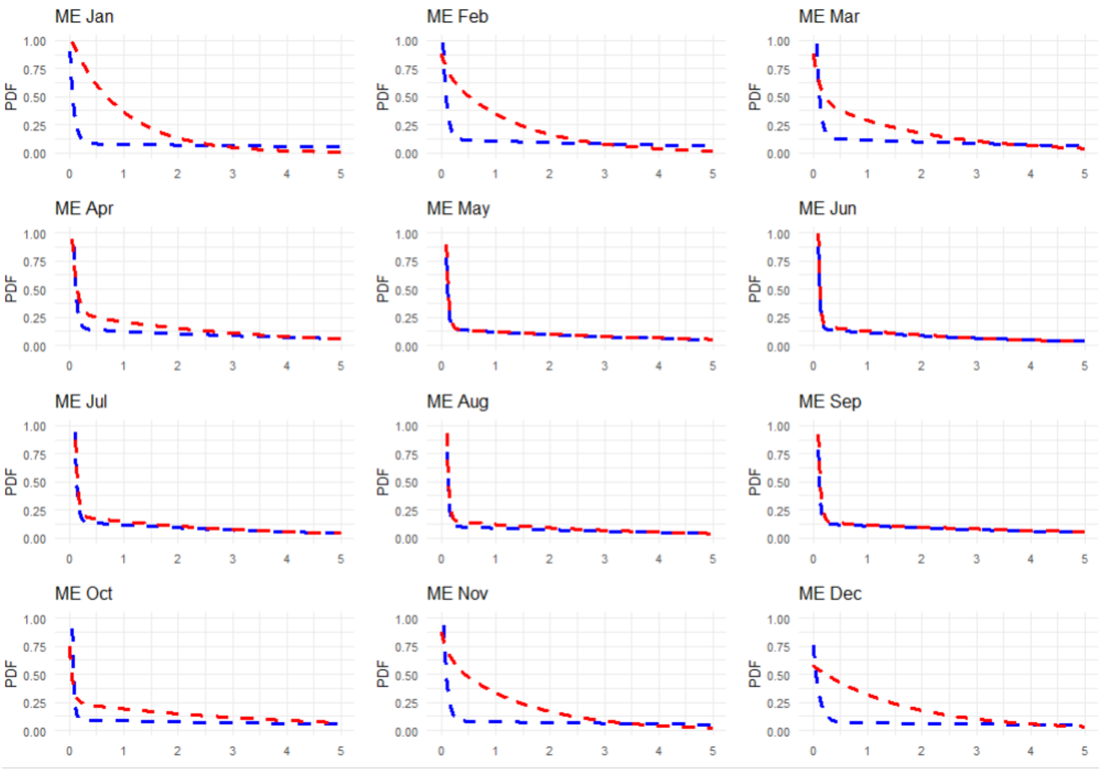
(l) Tate's Cairn



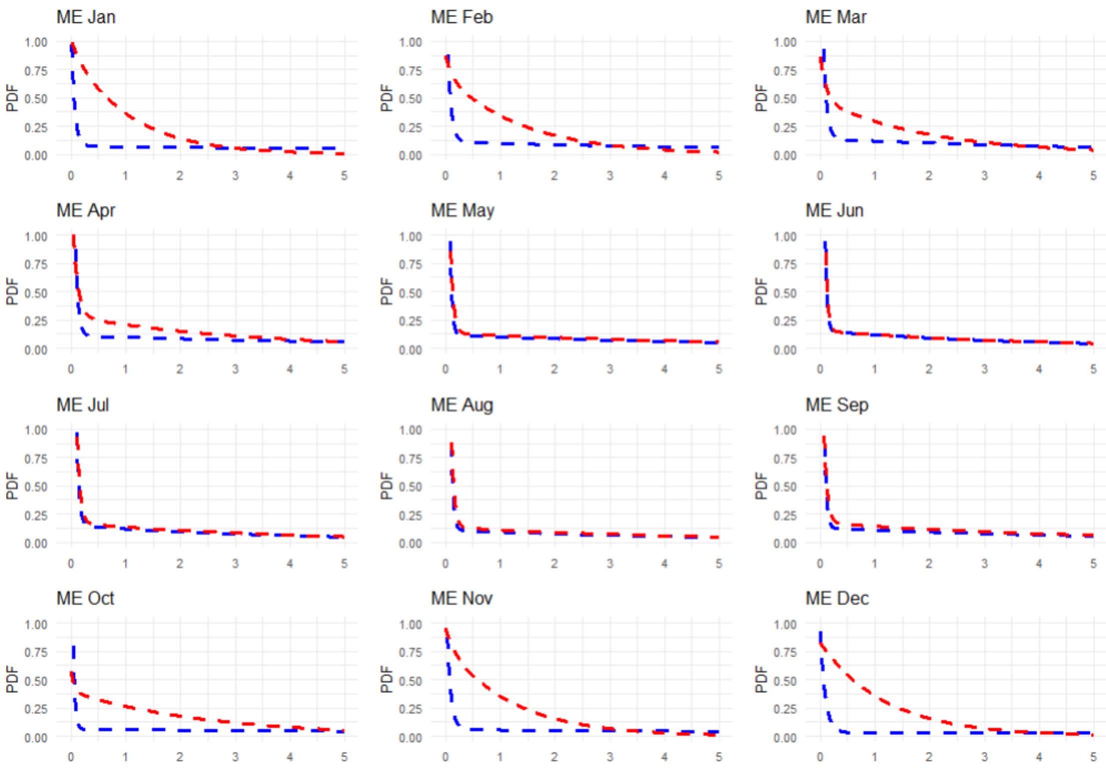
(m) Tseung Kwan O



(n) Waglan Island



(n) Hong Kong Observatory



(o) Hong Kong International Airport

Figure 4.2.1:Comparison of observed and modeled rainfall fits for each rainfall station

4.3 Performance of the Month MCME Model in Calibration Period (2003-2022)

In evaluating the descriptive capability of the MCME model, the calibration process utilized daily rainfall series from 2003-2022. Based on these historical records, two Markov chain transition probabilities (p_{01} and p_{11}) were obtained. The Maximum Likelihood Estimation (MLE) was then applied to estimate the monthly calibration cycle MCME parameters, totaling 60 for a 12-month rainfall process at one site and a cumulative total of 900 parameters for 15 sites. The parameter results are presented in Table 4.3.1. Using these calibrated parameters, a synthetic time series of the same length (20 years) was simulated.

Table 4.3.1: Summary of MCME parameters estimation for all rainfall stations

Station	Parameters	Month					
		Jan	Feb	Mar	Apr	May	Jun
Cheung Chau	p	0.38	0.98	0.96	0.92	0.81	0.63
	μ_1	0.8144	0.7966	0.6277	0.3453	0.1674	0.2412
	μ_2	0.5741	7.1613	6.6064	9.8471	18.6005	25.5408
	p_{01}	0.0411	0.0665	0.0968	0.1443	0.2005	0.2704
	p_{11}	0.3143	0.3585	0.3659	0.3364	0.4971	0.6107
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.73	0.68	0.85	0.98	0.98	0.12
	μ_1	0.2343	0.1868	0.2303	0.4312	0.9957	0.5229
	μ_2	17.3931	17.8166	13.7869	13.1701	5.4963	0.3395
	p_{01}	0.2174	0.2461	0.1556	0.0663	0.0583	0.0360
	p_{11}	0.5610	0.6034	0.5369	0.3934	0.3600	0.4167
Station	Parameters	Month					
		Jan	Feb	Mar	Apr	May	Jun
Ching Pak House(Tsing Yi)	p	0.99	0.98	0.96	0.92	0.76	0.58
	μ_1	1.0059	0.8753	0.5737	0.3457	0.1928	0.2866
	μ_2	7.4705	6.5419	7.6209	9.9393	23.7534	29.2961
	p_{01}	0.0486	0.0626	0.0941	0.1499	0.2088	0.2575
	p_{11}	0.3488	0.3962	0.3377	0.3482	0.5213	0.6792
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.66	0.6	0.79	0.98	0.48	0.19
	μ_1	0.2688	0.2238	0.2135	0.4190	0.8753	0.6283
	μ_2	19.1271	21.2736	16.2567	13.9733	0.6458	0.4102
	p_{01}	0.2388	0.2890	0.2042	0.0633	0.0562	0.0470
	p_{11}	0.6176	0.6337	0.4971	0.4697	0.3404	0.4000

Table 4.3.1 (Continued)

Station	Parameters	Month					
		Jan	Feb	Mar	Apr	May	Jun
King's Park	p	0.99	0.97	0.95	0.91	0.73	0.55
	μ_1	1.0034	0.7552	0.5346	0.3326	0.1973	0.3179
	μ_2	2.6916	6.5405	7.2210	10.6572	22.8450	31.4759
	p_{01}	0.0524	0.0710	0.1026	0.1455	0.2115	0.2960
	p_{11}	0.3478	0.3684	0.3373	0.3604	0.5665	0.6619
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.64	0.6	0.74	0.96	0.98	0.3
	μ_1	0.2101	0.1405	0.1928	0.2686	0.6817	0.7475
	μ_2	20.7178	24.4325	22.3246	22.5766	7.3868	0.5052
	p_{01}	0.2493	0.2722	0.2222	0.0751	0.0664	0.0471
	p_{11}	0.6280	0.6762	0.5412	0.4384	0.3684	0.4130
Station	Parameters	Month					
		Jan	Feb	Mar	Apr	May	Jun
Lau Fau Shan	p	0.99	0.97	0.96	0.92	0.78	0.64
	μ_1	0.9968	0.8119	0.5431	0.3773	0.2267	0.2192
	μ_2	7.8546	6.0461	8.0241	8.8224	18.7183	22.3341
	p_{01}	0.0540	0.0878	0.0929	0.1485	0.2059	0.2633
	p_{11}	0.3111	0.3016	0.3827	0.3684	0.4915	0.6157
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.73	0.65	0.82	0.98	0.98	0.33
	μ_1	0.2686	0.2713	0.2655	0.5256	0.9478	0.7776
	μ_2	14.8103	18.9964	12.3489	11.8950	6.0921	0.5390
	p_{01}	0.2367	0.2831	0.1968	0.0655	0.0587	0.0451
	p_{11}	0.5220	0.5602	0.4522	0.3148	0.4074	0.3810
Station	Parameters	Month					
		Jan	Feb	Mar	Apr	May	Jun
Pak Tam Chung (Tsak Yue Wu)	p	0.98	0.97	0.93	0.87	0.66	0.55
	μ_1	0.9126	0.7892	0.4799	0.3285	0.2343	0.3226
	μ_2	6.3483	6.0013	8.3988	11.2195	26.1691	32.0578
	p_{01}	0.0596	0.0842	0.1267	0.1670	0.2294	0.2987
	p_{11}	0.3061	0.3538	0.3367	0.4091	0.6147	0.6655
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.65	0.64	0.81	0.96	0.98	0.39
	μ_1	0.2141	0.2343	0.1607	0.3103	0.6457	0.8191
	μ_2	21.9139	24.9579	15.7394	19.6007	9.0697	0.5759
	p_{01}	0.2567	0.2633	0.2075	0.0730	0.0649	0.0399
	p_{11}	0.6082	0.5967	0.5029	0.4366	0.4167	0.4651
Station	Parameters	Month					
		Jan	Feb	Mar	Apr	May	Jun
Sha Tin	p	0.99	0.97	0.94	0.88	0.69	0.54
	μ_1	0.9769	0.7593	0.4850	0.3266	0.2206	0.3215
	μ_2	7.2111	6.0890	8.1143	11.5673	25.4969	31.2923
	p_{01}	0.0559	0.0758	0.1120	0.1545	0.2195	0.3143
	p_{11}	0.3191	0.3968	0.3587	0.3917	0.6009	0.6549
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.64	0.61	0.78	0.96	0.97	0.4
	μ_1	0.2271	0.2670	0.1711	0.2606	0.6527	0.8287
	μ_2	23.7213	28.3751	18.1474	16.4981	7.7084	0.5907
	p_{01}	0.2601	0.2611	0.2110	0.0872	0.0714	0.0474
	p_{11}	0.6098	0.6371	0.5220	0.4125	0.4328	0.4490

Table 4.3.1 (Continued)

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Shek Kong	p	0.99	0.98	0.96	0.91	0.74	0.57
	μ_1	0.9769	0.8119	0.5220	0.3293	0.1909	0.2673
	μ_2	7.2111	7.2995	8.3980	9.9549	21.8120	26.1222
	p_{01}	0.0559	0.0668	0.1006	0.1516	0.2368	0.3140
	p_{11}	0.3191	0.3818	0.3415	0.3333	0.5075	0.6236
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.66	0.68	0.77	0.97	0.98	0.38
	μ_1	0.1976	0.2108	0.2397	0.3554	0.9857	0.8144
	μ_2	19.2312	22.8516	20.7622	16.7034	5.6629	0.5741
	p_{01}	0.2599	0.2410	0.1977	0.0650	0.0607	0.0451
	p_{11}	0.5950	0.5895	0.5030	0.4462	0.4000	0.3953

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Ta Kwu Ling	p	0.99	0.97	0.96	0.91	0.74	0.59
	μ_1	1.0146	0.8456	0.5802	0.3323	0.1744	0.2622
	μ_2	7.0203	5.8038	6.9790	10.6024	17.2105	26.4850
	p_{01}	0.0486	0.0710	0.0987	0.1502	0.2190	0.3018
	p_{11}	0.3488	0.3684	0.3537	0.3540	0.5673	0.6130
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.64	0.61	0.79	0.98	0.98	0.33
	μ_1	0.1938	0.2430	0.2107	0.3109	0.8357	0.7760
	μ_2	17.7506	24.7097	15.5780	15.0759	6.6605	0.5353
	p_{01}	0.2500	0.2837	0.2019	0.0707	0.0667	0.0489
	p_{11}	0.6431	0.6160	0.5087	0.4179	0.3898	0.3913

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Tai Mei Tuk	p	0.37	0.99	0.97	0.9	0.78	0.62
	μ_1	0.8056	1.0215	0.6241	0.4024	0.1704	0.2335
	μ_2	0.5626	6.2733	7.6253	8.5814	17.9274	23.7281
	p_{01}	0.0480	0.0615	0.0914	0.1508	0.1963	0.2607
	p_{11}	0.2222	0.2727	0.3056	0.3652	0.5654	0.6400
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.72	0.69	0.85	0.97	0.98	0.1
	μ_1	0.1780	0.1976	0.2222	0.3723	0.9849	0.4790
	μ_2	17.3311	20.5990	11.5402	16.5642	5.5838	0.3057
	p_{01}	0.1990	0.2404	0.1927	0.0651	0.0664	0.0360
	p_{11}	0.6313	0.5877	0.4684	0.4545	0.3684	0.4000

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Tai Mo Shan	p	0.97	0.89	0.82	0.74	0.6	0.54
	μ_1	0.8852	0.7108	0.5806	0.5198	0.5496	0.5557
	μ_2	5.6224	4.7716	9.1291	12.7012	24.7483	28.0905
	p_{01}	0.0679	0.1302	0.1642	0.1934	0.2534	0.2767
	p_{11}	0.3559	0.4175	0.4275	0.5257	0.6349	0.6904
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.58	0.56	0.7	0.94	0.95	0.4
	μ_1	0.3619	0.4179	0.2715	0.3194	0.8332	0.8199
	μ_2	24.9074	28.6152	21.5689	17.4750	4.5119	0.5673
	p_{01}	0.2615	0.3015	0.2177	0.0765	0.0686	0.0476
	p_{11}	0.6642	0.6444	0.5833	0.5060	0.5135	0.4808

Table 4.3.1 (Continued)

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Tate's Cairn	p	0.98	0.93	0.89	0.81	0.62	0.53
	μ_1	0.9042	0.7055	0.5335	0.3634	0.3419	0.3518
	μ_2	6.0927	5.1620	6.8797	12.1766	25.2145	34.2221
	p_{01}	0.0579	0.1102	0.1375	0.1911	0.2547	0.2981
	p_{11}	0.3265	0.3614	0.4103	0.4295	0.6240	0.6794
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.59	0.58	0.74	0.94	0.95	0.47
	μ_1	0.2546	0.2713	0.1887	0.2422	0.5759	0.8663
	μ_2	25.4032	27.2017	21.1964	16.0592	6.7244	0.6272
	p_{01}	0.2686	0.2703	0.2260	0.1096	0.0833	0.0495
	p_{11}	0.6506	0.6618	0.5260	0.4242	0.4819	0.4717
Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Tseung Kwan O	p	0.75	0.97	0.91	0.85	0.69	0.53
	μ_1	0.9716	0.7165	0.4738	0.3335	0.2070	0.3248
	μ_2	0.8382	6.7881	7.9071	10.8046	22.6413	31.0833
	p_{01}	0.0560	0.0777	0.1255	0.1822	0.2312	0.3192
	p_{11}	0.3333	0.3710	0.4128	0.3986	0.5837	0.6678
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.65	0.59	0.77	0.95	0.97	0.24
	μ_1	0.2215	0.2420	0.1799	0.3225	0.5895	0.6898
	μ_2	23.2643	23.5587	20.5047	12.7421	8.5417	0.4590
	p_{01}	0.2560	0.2754	0.2291	0.0882	0.0748	0.0453
	p_{11}	0.6066	0.6533	0.4722	0.4535	0.3750	0.4222
Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Waglan Island	p	0.21	0.99	0.96	0.93	0.82	0.77
	μ_1	0.6531	1.0207	0.7944	0.4201	0.2499	0.2055
	μ_2	0.4280	6.1016	5.2517	8.3887	16.4863	13.4971
	p_{01}	0.0357	0.0685	0.0974	0.1360	0.1739	0.1956
	p_{11}	0.3226	0.3396	0.2933	0.3131	0.5031	0.5842
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.84	0.74	0.89	0.98	0.99	0.11
	μ_1	0.2515	0.2079	0.2538	0.5613	0.9464	0.4950
	μ_2	10.5527	12.9582	14.2510	9.4753	8.0597	0.3121
	p_{01}	0.1776	0.2285	0.1326	0.0513	0.0468	0.0424
	p_{11}	0.5031	0.5613	0.5000	0.4630	0.3953	0.1667
Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Hong Kong Observatory	p	0.99	0.98	0.95	0.9	0.73	0.54
	μ_1	1.0148	0.7464	0.5037	0.3214	0.2012	0.3278
	μ_2	4.9390	7.5290	8.0631	12.1810	23.8319	32.0524
	p_{01}	0.0503	0.0663	0.1049	0.1475	0.2177	0.3123
	p_{11}	0.3095	0.3333	0.3412	0.3514	0.5473	0.6525
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.62	0.56	0.74	0.97	0.97	0.3
	μ_1	0.3243	0.2817	0.1930	0.2568	0.6840	0.7465
	μ_2	23.4123	27.4875	22.3783	16.7074	7.0103	0.5031
	p_{01}	0.2630	0.2768	0.2083	0.0815	0.0670	0.0451
	p_{11}	0.6181	0.6749	0.5602	0.4430	0.4194	0.3953

Table 4.3.1 (Continued)

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Hong Kong International Airport	p	0.99	0.98	0.96	0.9	0.79	0.62
	μ_1	0.9425	0.7124	0.5034	0.3170	0.1713	0.2537
	μ_2	8.5032	8.6510	9.5234	11.2396	18.8354	26.9426
	p_{01}	0.0430	0.0723	0.1020	0.1567	0.1955	0.2648
	p_{11}	0.3421	0.2885	0.3125	0.3333	0.5251	0.6189
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.69	0.64	0.85	0.99	0.98	0.57
	μ_1	0.2515	0.2124	0.2038	0.3990	0.8258	0.9169
	μ_2	18.1876	21.1341	17.4982	17.9420	7.2665	0.7080
	p_{01}	0.2443	0.2730	0.1718	0.0563	0.0530	0.0491
	p_{11}	0.5631	0.5944	0.4690	0.3725	0.4423	0.4286

Table 4.3.2: Mean and Standard deviation of all MCME parameters estimation

Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Mean	p	0.8373	0.9680	0.9387	0.8847	0.7293	0.5867
	μ_1	0.9252	0.8052	0.5573	0.3597	0.2331	0.2994
	μ_2	4.8912	6.4507	7.7161	10.5789	21.6194	27.6147
	p_{01}	0.0517	0.0782	0.1098	0.1580	0.2169	0.2829
	p_{11}	0.3218	0.3540	0.3545	0.3746	0.5567	0.6438
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.6693	0.6287	0.7927	0.9673	0.9413	0.3087
	μ_1	0.2439	0.2405	0.2131	0.3570	0.8040	0.7285
	μ_2	19.8483	22.9979	17.5921	15.7639	6.4280	0.5035
	p_{01}	0.2426	0.2670	0.1984	0.0730	0.0642	0.0448
	p_{11}	0.6014	0.6194	0.5095	0.4298	0.4074	0.4051
Station	Parameters	Jan	Feb	Mar	Apr	May	Jun
Standard deviation	p	0.2675	0.0248	0.0381	0.0492	0.0639	0.0614
	μ_1	0.0981	0.0978	0.0789	0.0518	0.0946	0.0812
	μ_2	2.9033	0.9277	1.0137	1.2835	3.2164	5.0144
	p_{01}	0.0078	0.0183	0.0197	0.0169	0.0212	0.0314
	p_{11}	0.0310	0.0396	0.0382	0.0507	0.0465	0.0302
		Jul	Aug	Sep	Oct	Nov	Dec
	p	0.0631	0.0502	0.0507	0.0148	0.1238	0.1320
	μ_1	0.0474	0.0598	0.0327	0.0926	0.1475	0.1328
	μ_2	3.9417	4.1898	3.5776	3.1187	1.9534	0.1139
	p_{01}	0.0250	0.0197	0.0257	0.0140	0.0087	0.0042
	p_{11}	0.0444	0.0370	0.0352	0.0436	0.0449	0.0705

The conditional transition probability, p_{01}, for dry-day rainfall signifies the likelihood of a conversion from a daily dry day to a wet day event. Conversely, the conditional transition probability, p_{11}, for wet-day rainfall indicates the persistence of daily rainfall events.

It is evident from the data that the values of p_{11} for all stations during the monsoon period (May-September) are consistently in the range of 50% to 65%. This suggests that the probability of continuous rainfall is more than half during this period. In contrast, during the non-monsoon season (January to April), the mean value of p_{01} ranges from approximately 0.05, 0.07, 0.11 to 0.15. For October to December, it is about 0.07 to 0.04. These values imply that the probability of transitioning from dry to wet days is much less than the probability of transitioning from wet to dry days. The analysis highlights the likelihood of successive periods of no rainfall, especially in January, February, October, November, and December when p_{01} is in the range of 0.04 to 0.07.

Furthermore, the conditional transition probability from wet to wet days consistently surpasses the conditional transition probability from dry to wet days in every month at all sites.

The low standard deviations, both below 0.05, indicate that the variability of p_{01} and p_{11} among different stations is very low. Therefore, we can assume that the trends of rainfall and non-rainfall are more or less the same among different stations within the Hong Kong region.

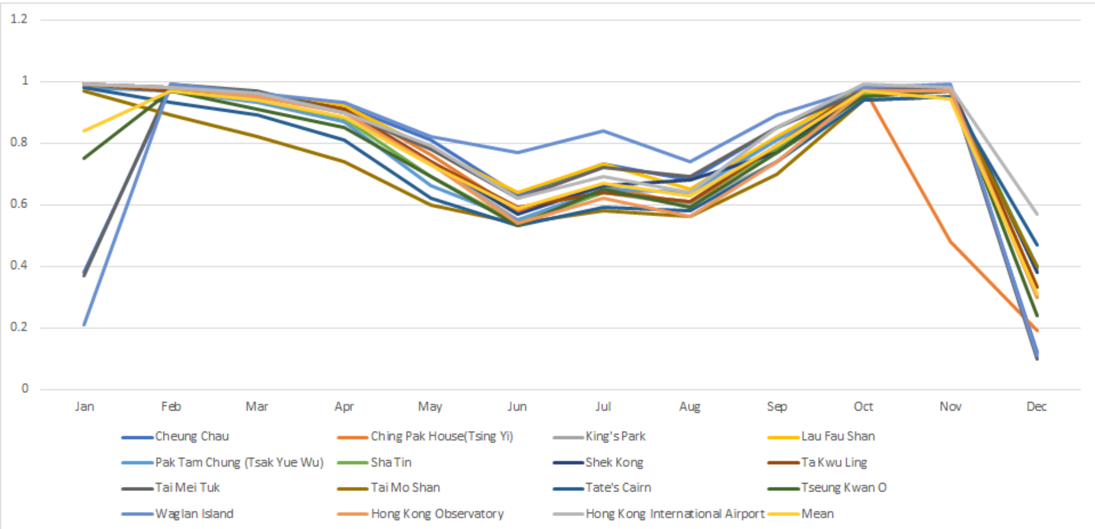


Figure 4.3.1: Line chart of p for different rainfall Stations across Months

The parameter μ_1 represents the average rainfall on dry days (less than 2.5 mm). Upon analyzing the data and referring to Figure 4.3.2, a subtle pattern emerges: the average rainfall on dry days is slightly higher during the non-monsoon season compared to the monsoon season. In the monsoon season, data is notably concentrated around 0.2 at rainfall stations. This empirical observation suggests that daily rainfall on dry days (below 2.5 mm) tends to be minimal during the monsoon season, empirical observations show that daily rainfall below 2.5mm is considered a dry day and vice versa during the monsoon season.

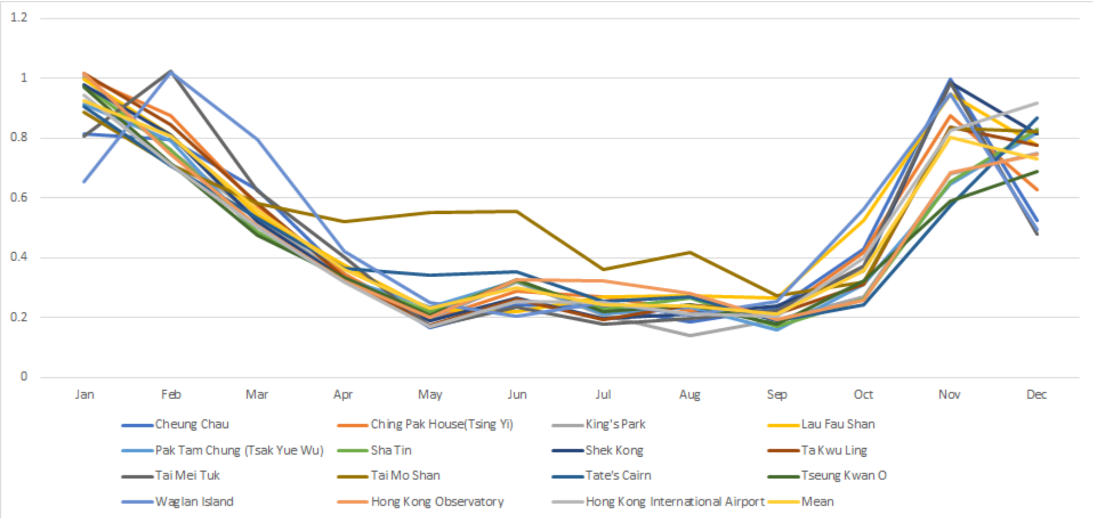


Figure 4.3.2: Line chart of μ_1 for different rainfall stations across months

The parameter μ_2 represents the average rainfall exceeding 2.5 mm. Upon examining the study data and referring to Figure 4.3.3, a distinct trend becomes apparent: rainfall values during the monsoon season are significantly higher than those during the non-monsoon season. This finding aligns with the expected pattern of increased rainfall during the monsoon season. Furthermore, a detailed monthly analysis reveals a positive correlation between higher rainfall months and elevated μ_2 values. This supports the notion that months with increased rainfall exhibit correspondingly higher μ_2 values.

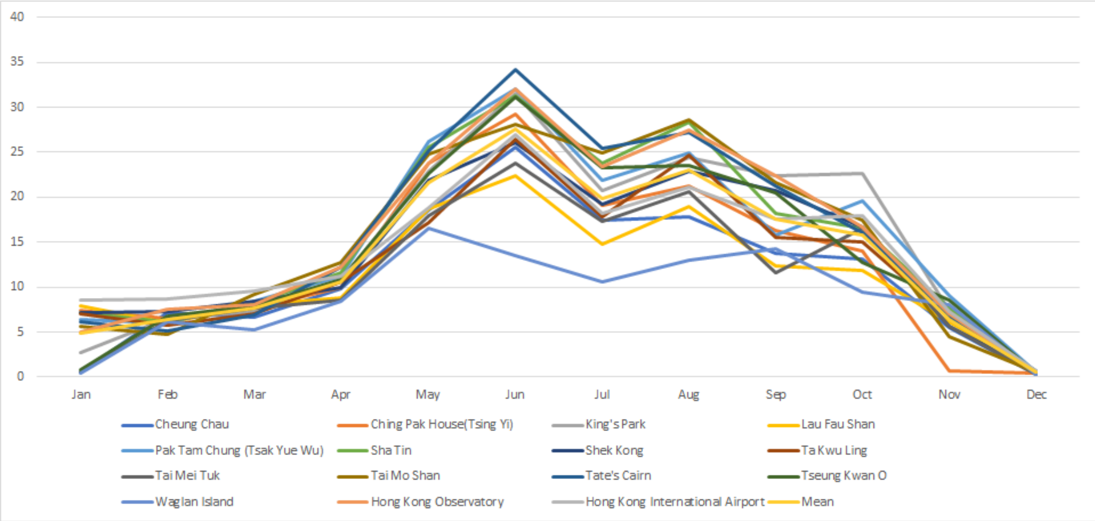


Figure 4.3.3: Line chart of μ_2 for different rainfall stations across months

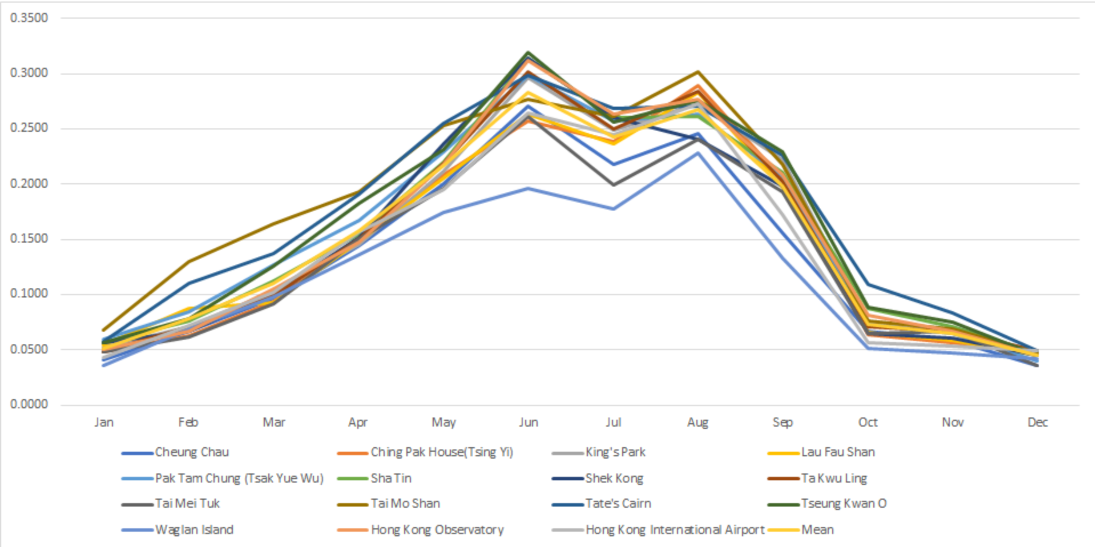


Figure 4.3.4: Line chart of p_{01} for different rainfall stations across months

From the data and Figure 4.3.4, it's evident that during the monsoon season, the probability of transitioning from dry to wet days consistently exceeds 20%, while in the non-monsoon season, these probabilities generally remain below 10%. Notably, the probability of transitioning from dry to wet days in the monsoon season is approximately twice as high as in the non-monsoon season. Moreover, in the non-monsoon season, the probability of continued dryness on the following day is significantly more than 85% if it is also dry on that day. This nuanced analysis sheds light on the seasonal variation in transition probabilities, emphasizing distinct patterns between the monsoon and non-monsoon seasons.

A closer examination of the data and Figure 4.3.5 reveals a noteworthy difference: during the monsoon season, the probability of transitioning from a wet day to a continuously wet day is about 60%, whereas during the non-monsoon season, these transition probabilities typically range between 30% and 50%. This complexity of observations suggests that the probability of continuous wet days is higher in the monsoon season. On the contrary, in the non-monsoon season, the probability of transitioning from wet to dry days exceeds the probability of transitioning to consecutive wet days. This nuanced analysis enriches our understanding of the seasonal dynamics of transition probabilities and provides valuable data support for a deeper understanding of the interactions between wet and dry days during monsoon and non-monsoon periods.

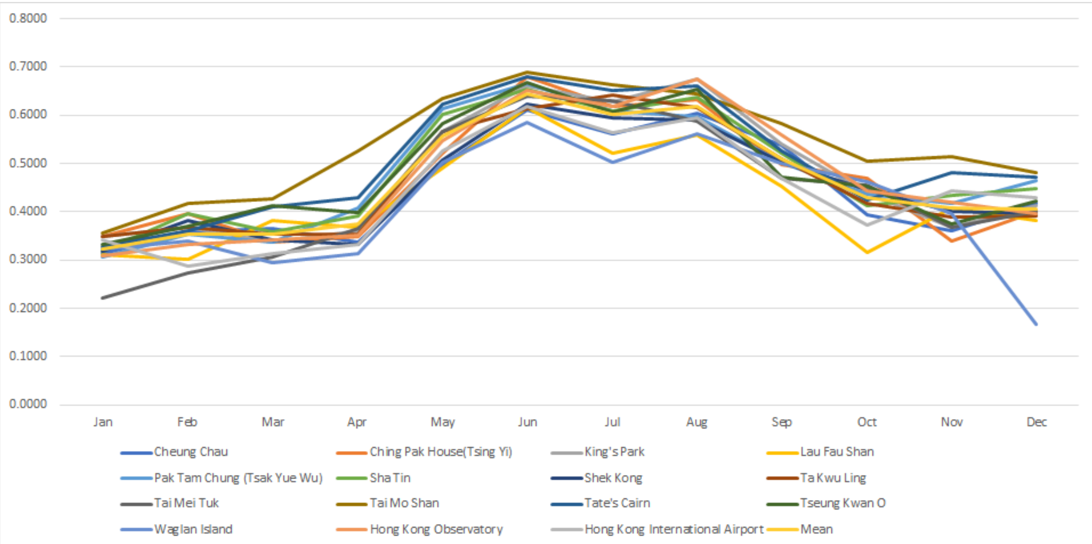


Figure 4.3.5: Line chart of p_{11} for different rainfall stations across months

4.4 Assessment of the MCME Model

Monthly total rainfall data for 20 years were generated and compared with the observed data. The observed and simulated monthly total rainfall intensities are presented in Table 4.4.1. The Relative Error (RE) values for different months across all our stations were averaged, and the results are detailed in Table 4.4.2. We found that the observed monthly total rainfall data differed from the simulated monthly total rainfall data by approximately 20%.

Table 4.4.1: Observed and simulated monthly rainfall intensities

	Cheung Chau			Ching Pak House(Tsing Yi)			King's Park			Lau Fau Shan			Pak Tam Chung (Tsak Yue Wu)		
Month	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE
Jan	656.5	521.9759	20.49%	656.5	1006.967	53.38%	628.4	758.7328	20.74%	664.5	1043.569	57.05%	744.5	896.0851	20.36%
Feb	706.5	845.8555	19.72%	706.5	853.1951	20.76%	836.6	812.3398	2.90%	777.5	809.9622	4.18%	796.2	805.0418	1.11%
Mar	1226.5	921.0721	24.90%	1226.5	921.5708	24.86%	1325.7	911.5145	31.24%	1295	958.6022	25.98%	1566	1144.319	26.93%
Apr	2074	1252.15	39.63%	2074	1288.359	37.88%	2217.2	1361.945	38.57%	1886.5	1187.253	37.07%	2464	1609.692	34.67%
May	5978	3941.448	34.07%	5978	4908.19	17.90%	6117.7	4834.63	20.97%	4686	4010.06	14.42%	7263	7108.856	2.12%
Jun	8597	6364.662	25.97%	8597	8087.939	5.92%	9536.6	8908.302	6.59%	6576	5866.059	10.80%	9678.5	9301.029	3.90%
Jul	5554.5	4085.645	26.44%	5554.5	5102.855	8.13%	6512.9	5518.685	15.27%	4164	3353.968	19.45%	6637.5	5555.045	16.31%
Aug	6938	4770.052	31.25%	6938	5597.852	19.32%	8707.9	6906.233	20.69%	5607.5	4607.446	17.83%	7264	6360.046	12.44%
Sep	4269	2244.018	47.43%	4269	3190.491	25.26%	5785.1	4854.202	16.09%	3186.5	2151.817	32.47%	4819.5	2750.882	42.92%
Oct	1623.5	1137.412	29.94%	1623.5	1284.768	20.86%	2775.7	1956.584	29.51%	1303.5	1066.051	18.22%	2404.5	1705.985	29.05%
Nov	530.5	875.1282	64.96%	530.5	548.0056	3.30%	951.2	809.1474	14.93%	693.5	896.5019	29.27%	1019.5	927.1404	9.06%
Dec	393.5	334.9831	14.87%	393.5	399.1303	1.43%	468.1	475.7885	1.64%	487	496.9458	2.04%	513	522.2719	1.81%
		Mean(RE)	31.64%		Mean(RE)	19.92%		Mean(RE)	18.26%		Mean(RE)	22.40%		Mean(RE)	16.72%

	Sha Tin			Shek Kong			Ta Kwu Ling			Tai Mei Tuk			Tai Mo Shan		
Month	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE
Jan	673	1011.345	50.27%	673	1011.345	50.27%	648.5	987.9536	52.34%	504.5	517.3195	2.54%	784	893.0717	13.91%
Feb	825	800.4839	2.97%	764.5	872.2773	14.10%	746.5	811.8214	8.75%	583	849.0112	45.63%	1004	813.5104	18.97%
Mar	1503.5	1036.323	31.07%	1348.5	1003.134	25.61%	1199	922.6868	23.05%	1105.5	888.101	19.67%	1894.5	1610.476	14.99%
Apr	2449.5	1588.932	35.13%	2195.5	1262.193	42.51%	2215.5	1355.094	38.84%	1857	1173.468	36.81%	2835.5	2405.764	15.16%
May	6839.5	6727.649	1.64%	5919	4745.471	19.83%	5405.5	4055.747	24.97%	5283	3745.108	29.11%	6814.5	6436.885	5.54%
Jun	9644.5	8362.111	13.30%	8019	7409.496	7.60%	7865.5	7480.085	4.90%	7003.5	6283.655	10.28%	8336	8353.866	0.21%
Jul	7041.5	6083.11	13.61%	6125	4924.741	19.60%	6009	4764.706	20.71%	5517	4255.716	22.86%	7479.5	6636.652	11.27%
Aug	8277.5	7430.292	10.24%	6534	5877.79	10.04%	7531.5	6401.136	15.01%	6124.5	5301.433	13.44%	8637	8330.177	3.55%
Sep	5131.5	3244.541	36.77%	4793	3561.941	25.68%	4213	2730.54	35.19%	3333.5	2062.086	38.14%	5429.5	4988.136	8.13%
Oct	2693	1577.215	41.43%	2003	1443.502	27.93%	2141.5	1352.23	36.86%	1923.5	1443.024	24.98%	2475	1730.898	30.06%
Nov	1030.5	879.5194	14.65%	664.5	901.6137	35.68%	783.5	860.5535	9.83%	664	885.6989	33.39%	819.5	843.4779	2.93%
Dec	519	527.9477	1.72%	510	520.9261	2.14%	486	494.9193	1.84%	300	307.4101	2.47%	513.5	520.6212	1.39%
		Mean(RE)	21.07%		Mean(RE)	23.42%			22.69%			23.28%			10.51%

	Tate's Cairn			Tseung Kwan O			Waglan Island			Hong Kong Observatory			Hong Kong International Airport		
Month	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE	Observed	Simulated	RE
Jan	747.5	901.4437	20.59%	608.5	628.6293	3.31%	409	420.4006	2.79%	635.5	867.6951	36.54%	704	987.5117	40.27%
Feb	949	779.6822	17.84%	880	825.3028	6.22%	582.5	889.4372	52.69%	826.9	826.6933	0.02%	875	860.803	1.62%
Mar	1503.5	1187.57	21.01%	1632	1167.871	28.44%	879.5	853.7462	2.93%	1419.4	965.1194	32.01%	1418.6	1055.146	25.62%
Apr	2725.5	2108.916	22.62%	2501.5	1641.335	34.39%	1680.5	1046.408	37.73%	2410.8	1524.982	36.74%	2377.7	1449.595	39.03%
May	7065	6675.11	5.52%	6418	5361.388	16.46%	3874	2933.239	24.28%	6238.5	5547.06	11.08%	5311.2	3820.778	28.06%
Jun	10554.5	10159.04	3.75%	9744.5	9410.475	3.43%	4110.5	2667.936	35.09%	9834.8	8531.65	13.25%	7609.5	7182.585	5.61%
Jul	7894	6344.789	19.63%	6867.5	5895.02	14.16%	3118	2002.095	35.79%	6701.4	6237.432	6.92%	5196.9	4410.194	15.14%
Aug	8409	6829.57	18.78%	7500.5	5919.026	21.08%	4296	3086.885	28.15%	8731.3	8310.809	4.82%	6585.1	5251.409	20.25%
Sep	5660	4581.784	19.05%	5397.5	4135.584	23.38%	3045	1964.832	35.47%	5791.1	4151.769	28.31%	4076.8	2691.957	33.97%
Oct	3003.5	1807.007	39.84%	2221.5	1356.993	38.92%	1200	962.3309	19.81%	2653.1	1493.975	43.69%	1649.4	1343.839	18.53%
Nov	1191.5	818.8389	31.28%	1141	883.2085	22.59%	676	930.3567	37.63%	977.1	828.7836	15.18%	799.2	909.7204	13.83%
Dec	542.5	551.9689	1.75%	432	438.4174	1.49%	310	317.0541	2.28%	467.5	476.2922	1.88%	574.2	587.5697	2.33%
			18.47%		Mean(RE)	17.82%		Mean(RE)	26.22%		Mean(RE)	19.20%		Mean(RE)	20.36%

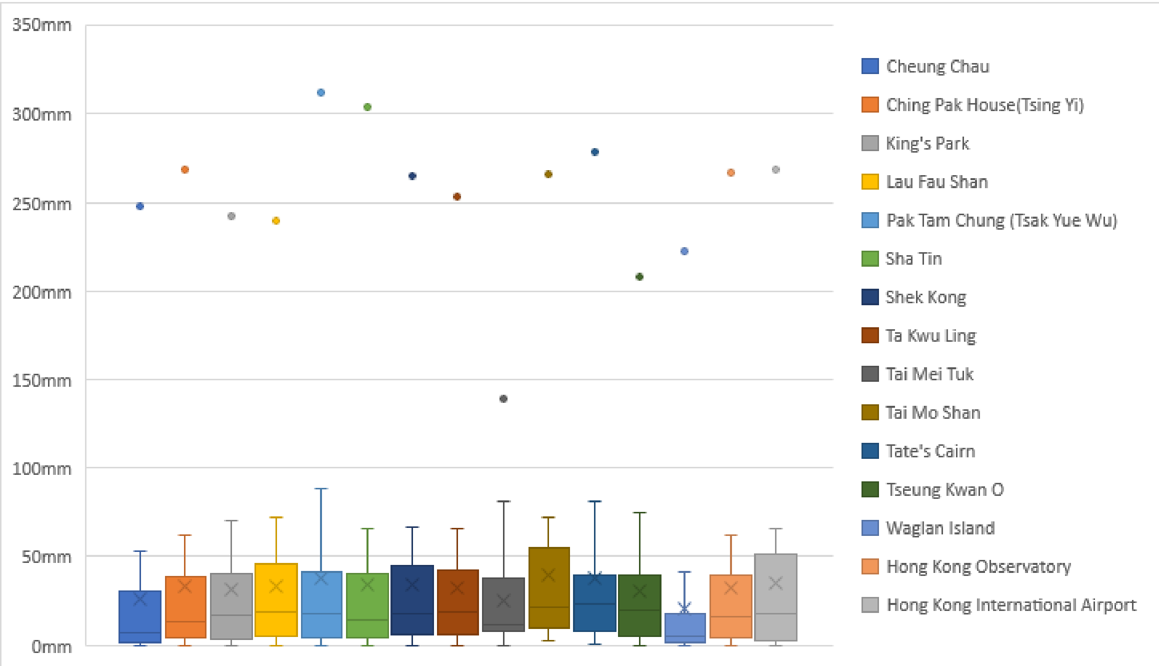
Table 4.4.2: Monthly RE average

Month	Jan	Feb	Mar	Apr	May	Jun
RE(mean)	29.66%	14.50%	23.89%	35.12%	17.07%	10.04%
Month	Jul	Aug	Sep	Oct	Nov	Dec
RE(mean)	17.69%	16.46%	29.89%	29.97%	22.57%	2.74%

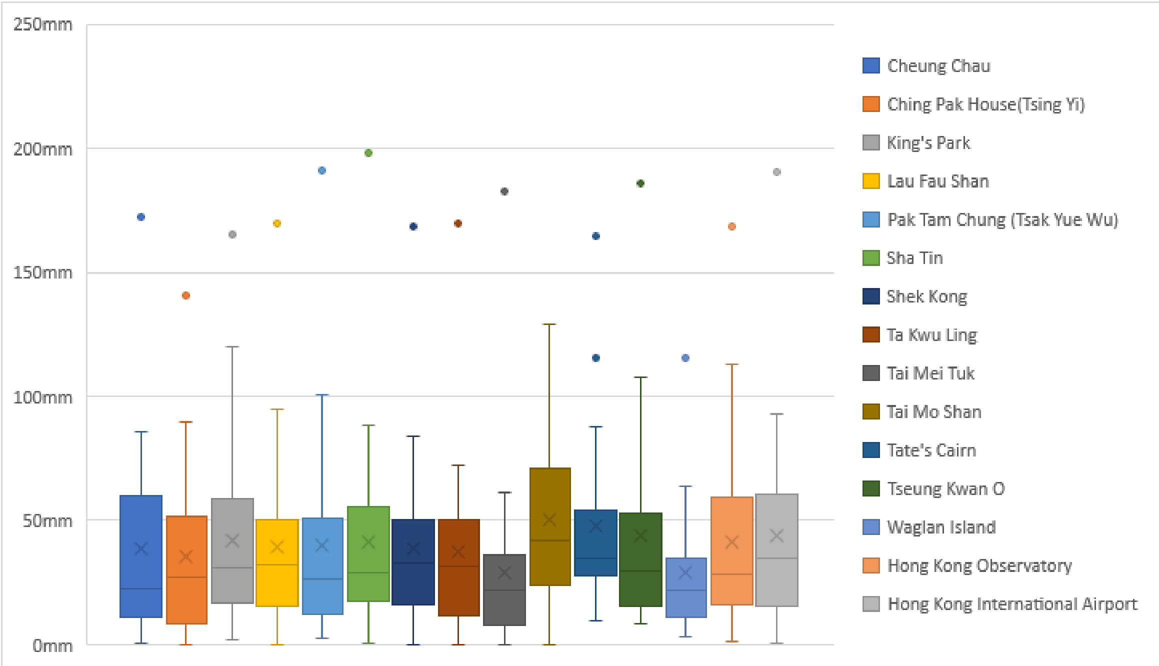
By integrating the relative error (RE) with the boxplot, we gain valuable insights into how measurements deviate from true values. A lower relative error signifies a more precise measurement, indicating higher measurement accuracy. When the corresponding boxplot shows a larger interquartile range (IQR), it suggests that the data are broadly spread, with outliers positioned notably far from the median.

Conversely, a higher relative error implies a greater measurement bias. A smaller interquartile range in the corresponding boxplot suggests that outliers are closer to the median, indicating a more concentrated dataset.

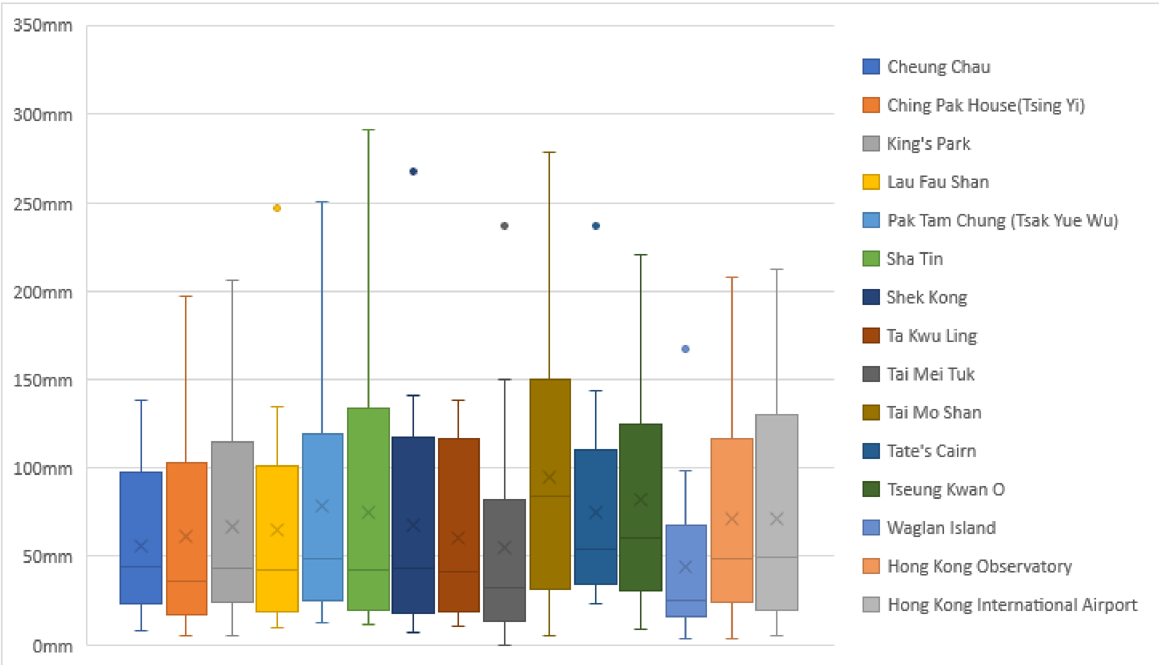
The joint assessment of IQR and relative error in the box-and-whisker plot enables us to evaluate data concentration and accuracy. A smaller IQR and shorter box whiskers in the boxplots indicate greater data concentration, reflecting a more dependable estimation of the observed data. Conversely, larger IQR and longer box whiskers indicate a substantial disparity between the estimated and observed data. This comprehensive analysis provides insights into model performance and strongly supports the credibility of data estimates.



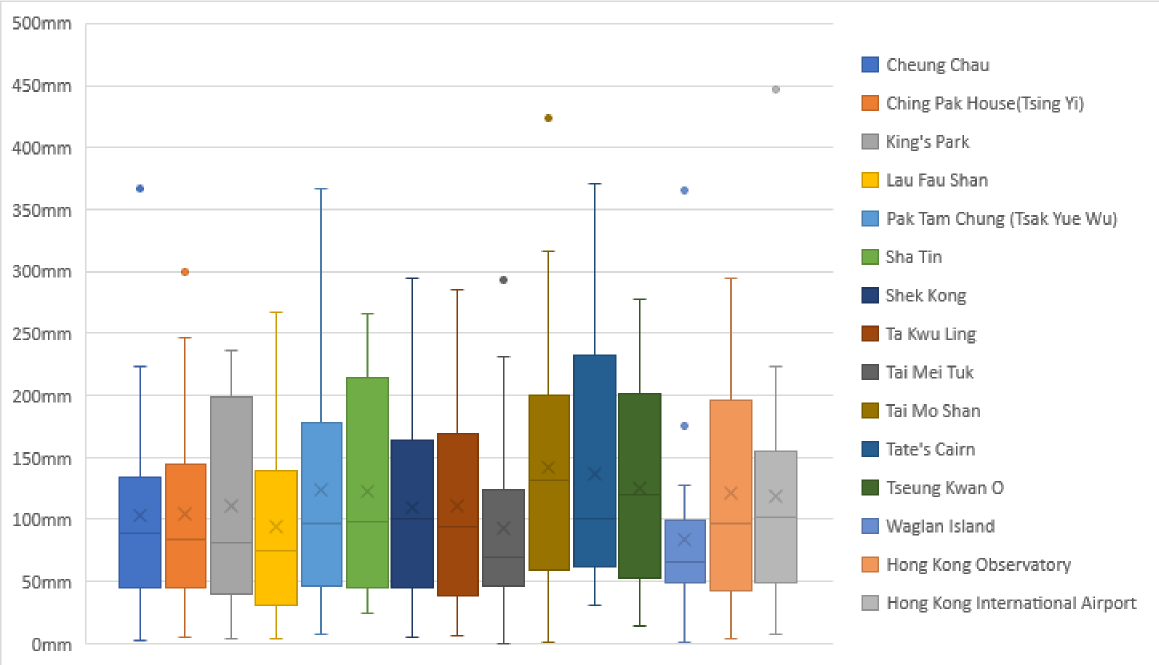
(a) January



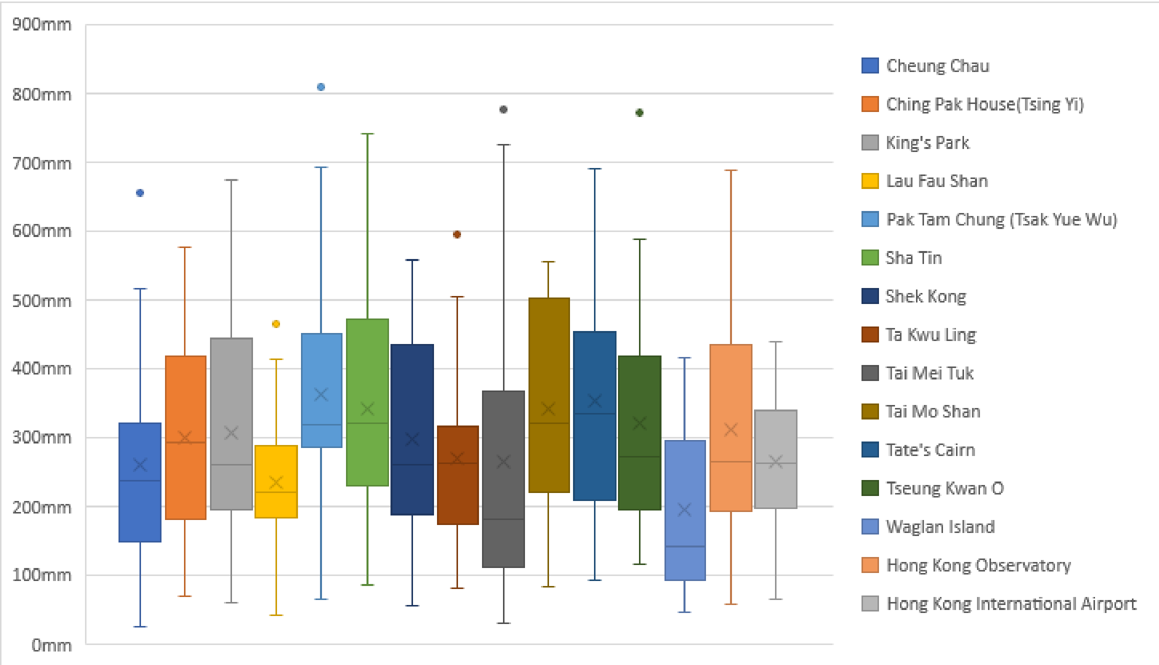
(b) February



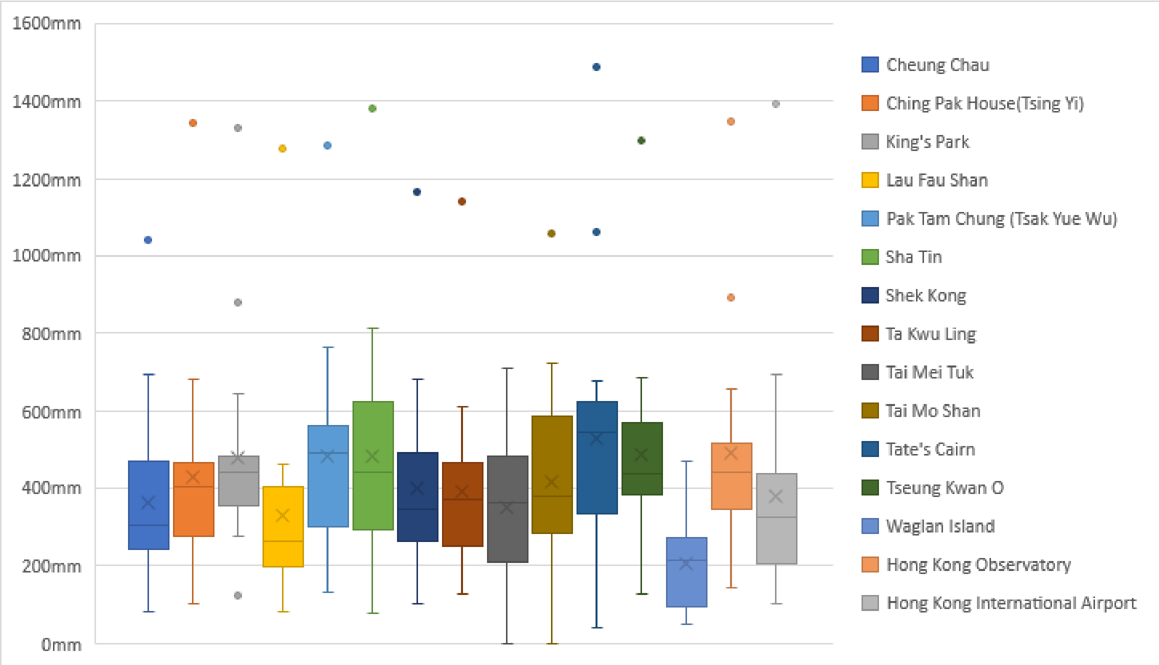
(c) March


500mm
450mm
400mm
350mm
300mm
250mm
200mm
150mm
100mm
50mm
0mm
Cheung Chau
Ching Pak House(Tsing Yi)
King's Park
Lau Fau Shan
Pak Tam Chung (Tsak Yue Wu)
Sha Tin
Shek Kong
Ta Kwu Ling
Tai Mei Tuk
Tai Mo Shan
Tate's Cairn
Tseung Kwan O
Waglan Island
Hong Kong Observatory
Hong Kong International Airport

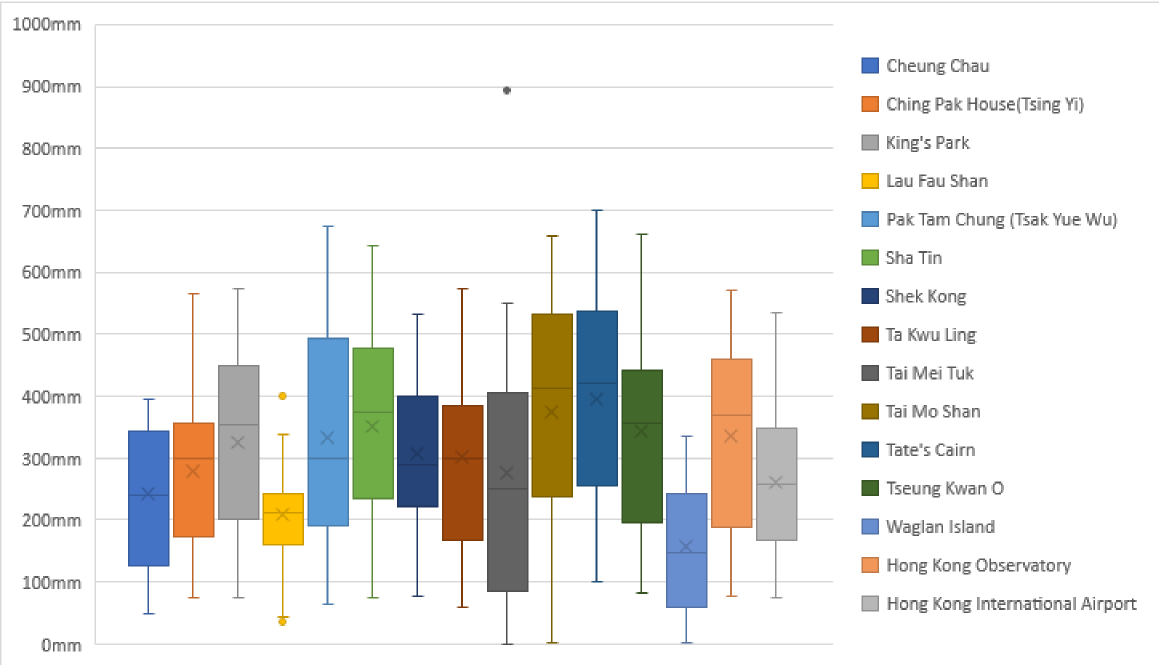
(d) April


900mm
800mm
700mm
600mm
500mm
400mm
300mm
200mm
100mm
0mm
Cheung Chau
Ching Pak House(Tsing Yi)
King's Park
Lau Fau Shan
Pak Tam Chung (Tsak Yue Wu)
Sha Tin
Shek Kong
Ta Kwu Ling
Tai Mei Tuk
Tai Mo Shan
Tate's Cairn
Tseung Kwan O
Waglan Island
Hong Kong Observatory
Hong Kong International Airport

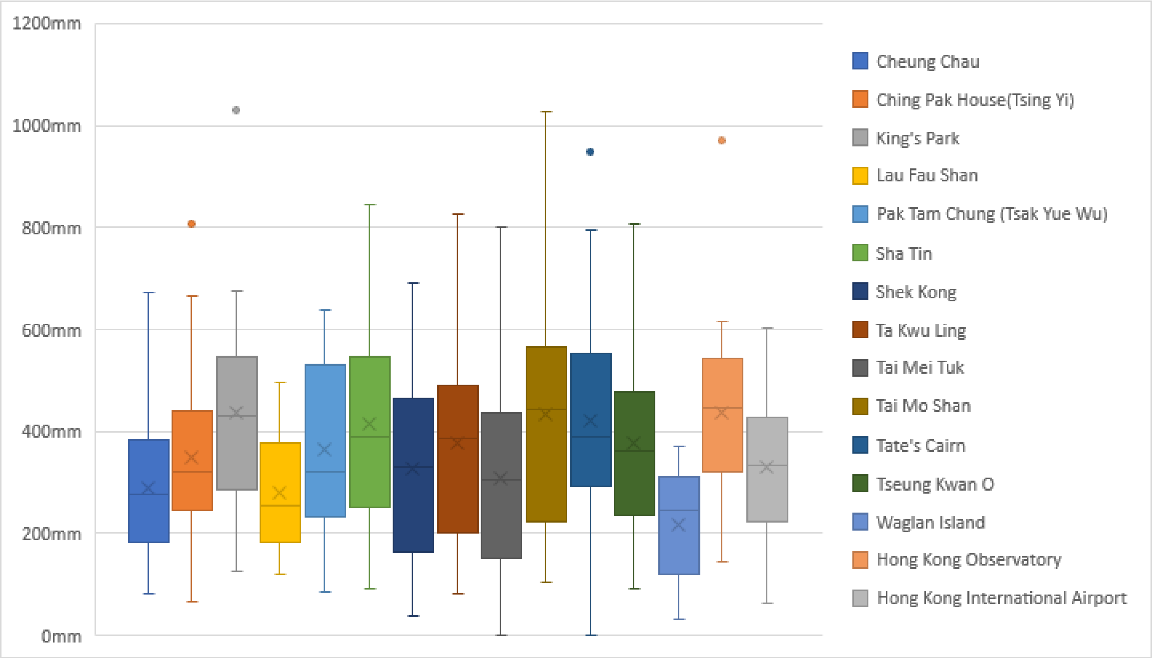
(e) May


1600mm
1400mm
1200mm
1000mm
800mm
600mm
400mm
200mm
0mm
Cheung Chau
Ching Pak House(Tsing Yi)
King's Park
Lau Fau Shan
Pak Tam Chung (Tsak Yue Wu)
Sha Tin
Shek Kong
Ta Kwu Ling
Tai Mei Tuk
Tai Mo Shan
Tate's Cairn
Tseung Kwan O
Waglan Island
Hong Kong Observatory
Hong Kong International Airport

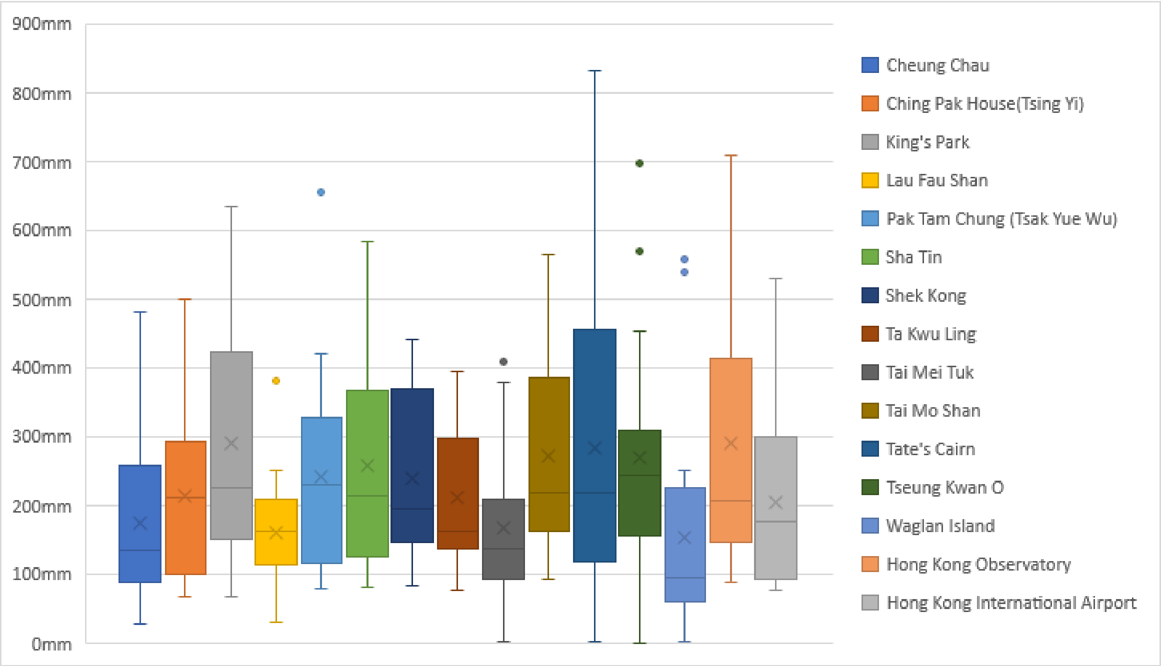
(f) June


1000mm
900mm
800mm
700mm
600mm
500mm
400mm
300mm
200mm
100mm
0mm
Cheung Chau
Ching Pak House(Tsing Yi)
King's Park
Lau Fau Shan
Pak Tam Chung (Tsak Yue Wu)
Sha Tin
Shek Kong
Ta Kwu Ling
Tai Mei Tuk
Tai Mo Shan
Tate's Cairn
Tseung Kwan O
Waglan Island
Hong Kong Observatory
Hong Kong International Airport

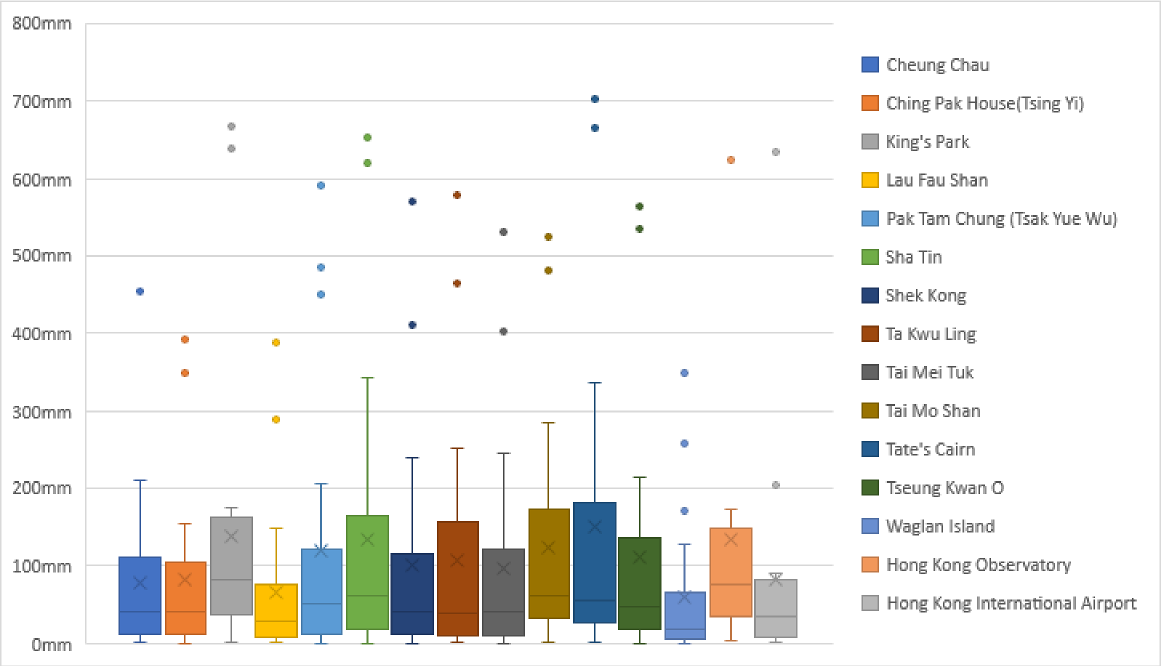
(g) July



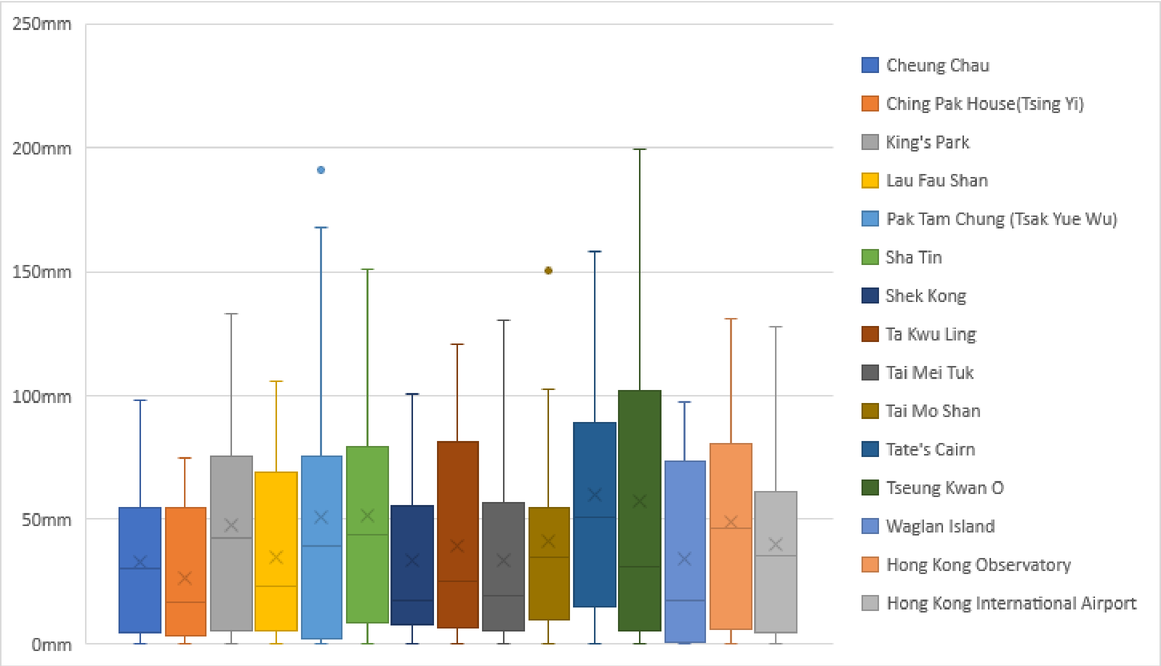
(h) August



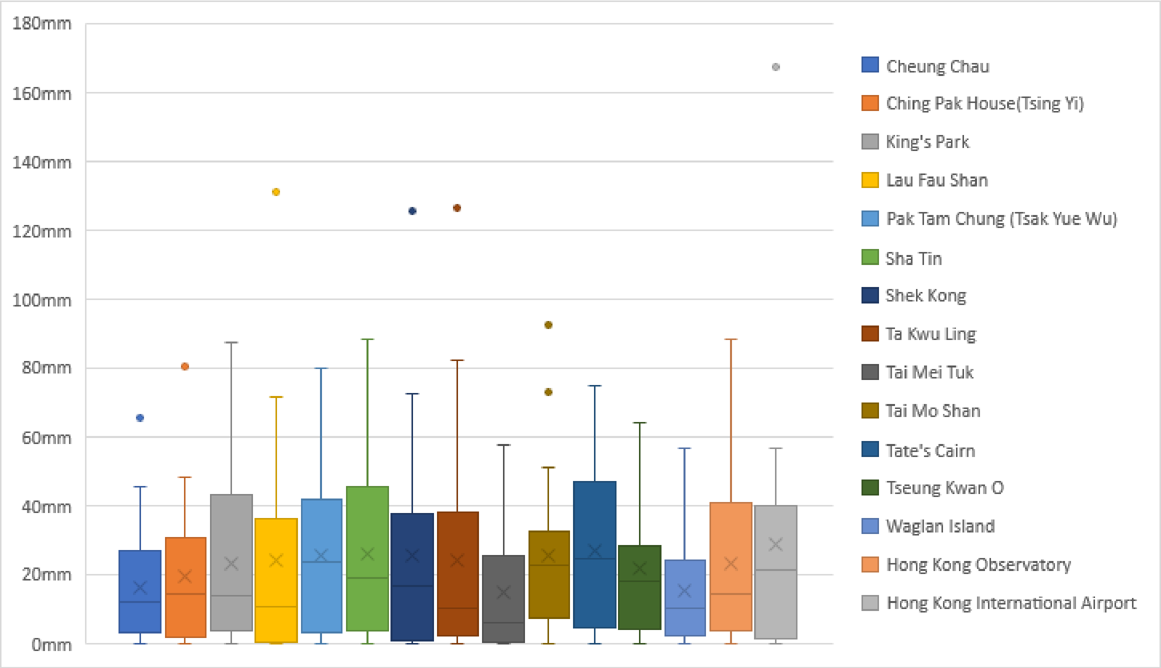
(i) September



(j) October



(k) November



(l) December

Figure 4.4.1: Box plot illustrating monthly rainfall at each rainfall stations

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In this study, we employed a Markov Chain Mixed Exponential Model (MCME model) to characterize rainfall patterns in Hong Kong. This model, combining a first-order two-state Markov chain and a mixed exponential distribution, effectively captures the spatial and temporal variability of rainfall in the region. Parameter estimates, obtained through maximum likelihood estimation for different months, reveal notable variations in daily rainfall. The analysis demonstrates that the MCME model proficiently reflects the diverse characteristics of rainfall in Hong Kong.

The study's findings underscore the significant advantages of the MCME model in describing complex rainfall processes. Notably, it effectively simulates the probability of event occurrences while capturing the diversity of intensities. The model has proven valuable in the absence of historical data and enales the synthesis of rainfall events over different time periods. In conclusion, the MCME model emerges as a robust stochastic simulation tool, offering a more comprehensive and accurate depiction of rainfall processes. Its application provides a powerful means for gaining a deeper understanding and modeling of rainfall dynamics.

5.2 Recommendations for future work

Several recommendations for future work can enhance the MCME model. Consider exploring a 2nd or 3rd order Markov chain, especially during the non-monsoon season where the probability of a dry day transitioning to another dry day is notably high. This adjustment might offer a better fit to the observed reality. Experimenting with various statistical distributions to model the distribution of monthly rainfall could provide insights into finding the most suitable fit. Given the substantial variability in same-month rainfall across different years in Hong Kong, introducing an annual rainfall indicator could be beneficial for correcting rainfall parameters for different years. Additionally, it's advisable to estimate and compare rainfall distribution parameters using different methods to identify the best fit.

REFERENCES

Arumugam, P., and Karthik, S. M., 2018. Stochastic modelling in yearly rainfall at Tirunelveli district, Tamil Nadu, India. *Materials Today: Proceedings*, *5*(1), 1852-1858.

Berhane, T., Shibabaw, N., Awgichew, G., and Kebede, T., 2020. Option pricing of weather derivatives based on a stochastic daily rainfall model with Analogue Year component. *Heliyon*, *6*(1), e03212.

Chin, E. H., 1977. Modeling daily precipitation occurrence process with Markov chain. *Water resources research*, *13*(6), 949-956.

da Silva Jale, J., Júnior, S. F. A. X., Xavier, É. F. M., Stošić, T., Stošić, B., and Ferreira, T. A. E. 2019. Application of Markov chain on daily rainfall data in Paraíba- Brazil from 1995-2015. *Acta Scientiarum. Technology*, *41*, e37186.

El Outayek, S. 2021. *Stochastic modeling of daily precipitation process in the context of climate change*. PhD thesis, McGill University, Montreal.

Kannan, S. K., and Farook, J. A., 2015. Stochastic simulation of precipitation using markov chain-mixed exponential model. *Applied Mathematical Sciences*, *65*(9), 3205-3212.

Hussain, A., 2008. *Stochastic modeling of rainfall processes: a markov chain-mixed exponential model for rainfalls in different climatic conditions*. Master thesis. McGill University, Montreal.

Nguyen, V. T. V., and Mayabi, A. (1991). Probabilistic analysis of summer daily rainfall for the Montreal region. *Canadian Water Resources Journal*, *16*(1), 65-80.

Rolda'n, J., and Woolhiser, D. A., 1982. Stochastic daily precipitation models: 1. A comparison of occurrence processes. *Water resources research*, *18*(5), 1451-1459.

Shibabaw, N., Berhane, T., Kebede, T., and Walelign, A., 2022. Spatio-temporal rainfall distribution and Markov chain analogue year stochastic daily rainfall model in Ethiopia. *Journal of Resources and Ecology*, *13*(2), 210-219.

Yusof, F., Yung, L. M., and Yusop, Z., 2015. Tropical daily rainfall amount modelling using Markov Chain-Mixed Exponential (MCME). *Jurnal Teknologi*, *74*(11).

Zhang, R., and Xia, L., 2012. The application of grey Markov chain model in rainfall projection. *Journal of Chongqing University of Technology (Natural Science)*, 26(10).

APPENDICES

APPENDIX A: Computer Programme Core Code

```
# Loop over B_column from 1 to 12
for (b_value in 1:12) {
  selected_datarain <- D_column[B_column == b_value] # Select the corresponding month data

  # Add each month's results to the vector
  all_total_rainfall <- c(all_total_rainfall, total_rainfall)
  all_total_obrainfall <- c(all_total_obrainfall, total_obrainfall)

  # Log-likelihood function for the mixed exponential distribution
  mcme_likelihood <- function(params, data) {
    p <- params[1]
    beta1 <- params[2]
    beta2 <- params[3]

    if (p <= 0 || beta1 <= 0 || beta2 <= 0) {
      return(-Inf) # Returns a very negative number to avoid calculation errors
    }

    log_likelihood1 <- sum(log(p / beta1 * exp(-data[data < 2.5] / beta1)))
    log_likelihood2 <- sum(log((1 - p) / beta2 * exp(-data[data >= 2.5] / beta2)))

    return(log_likelihood1 + log_likelihood2)
  }

  # parametric estimating function
  estimate_parameters <- function(data, p_range, beta1_range, x_bar) {
    best_likelihood <- -Inf
```

```
  best_params <- c(p = 0, beta1 = 0, beta2 = 0)

  for (p_est in p_range) {
    for (beta1_est in beta1_range) {
      beta2_est <- (x_bar - (p_est / beta1_est)) / (1 - p_est)
      params <- c(p = p_est, beta1 = beta1_est, beta2 = beta2_est)

      likelihood <- mcme_likelihood(params, data)

      if (likelihood > best_likelihood) {
        best_params <- params
        best_likelihood <- likelihood
      }
    }
  }

  return(best_params)
}

#  Initialize parameter ranges and intervals
initial_p_range <- seq(0.01, 0.99, by = 0.01)
initial_beta1_range <- seq(0.01 * x_bar, 0.99 * x_bar, by = 0.01 * x_bar)

#  Estimation of optimal parameters
best_params <- estimate_parameters(selected_datarain, initial_p_range,
initial_beta1_range, x_bar)

# Discretize historical data
rain_discrete <- ifelse(selected_datarain < 2.5, 0, 1)

# Estimating the transfer probability matrix
transition_matrix <- matrix(0, nrow = 2, ncol = 2, dimnames = list(c(0, 1), c(0, 1)))
```

```
  for (i in 1:(length(rain_discrete) - 1)) {
    from_state <- rain_discrete[i]
    to_state <- rain_discrete[i + 1]
    transition_matrix[from_state + 1, to_state + 1] <- transition_matrix[from_state + 1,
to_state + 1] + 1
  }

  # Converting counts to probabilities
  transition_matrix <- transition_matrix / rowSums(transition_matrix)

  # Generate Markov chain time series
  set.seed(123)  # Setting random seeds to ensure reproducibility
  num_steps <- length(selected_datarain)  # Use the length of selected_datarain as the
time step
  rainfall_sequence <- numeric(num_steps)

  current_state <- 0  # Initial state is non-rainfall

  for (i in 1:num_steps) {
    current_state <- sample(c(0, 1), 1, prob = transition_matrix[current_state + 1, ])
    rainfall_sequence[i] <- current_state
  }

  lambda1 <- best_params[2]
  lambda2 <- best_params[3]

  p <- best_params[1]

  # Define a function to generate rainfall data
  generate_rainfall <- function(rainfall_sequence, lambda1, lambda2, p) {
    simulated_rainfall <- numeric(length(rainfall_sequence))

    for (i in seq_along(rainfall_sequence)) {
```

```
    if (rainfall_sequence[i] == 1) {
      simulated_rainfall[i] <- rexp(1, rate = 1/lambda2)
    } else {
      simulated_rainfall[i] <- rexp(1, rate = 1/lambda1)
    }
  }

  return(simulated_rainfall)
}

# Call function to generate simulated rainfall data
rainfall_sequence_data <- generate_rainfall(rainfall_sequence, lambda1, lambda2, p)

# Summation of rainfall_sequence_data
total_rainfall <- sum(rainfall_sequence_data)

# Summation of selected_datarain
total_obrainfall <- sum(selected_datarain)

  sum_rainfall_sequence_data[i] <-
sum(rainfall_sequence_data[start_index:end_index])
  sum_selected_datarain[i] <- sum(selected_datarain[start_index:end_index])
}

# Calculating Relative Error
re_op <- abs(op - p) / op
re_mu1 <- abs(average_below_2.5 - lambda1) / average_below_2.5
re_mu2 <- abs(mean(selected_datarain[selected_datarain >= 2.5]) - lambda2) /
mean(selected_datarain[selected_datarain >= 2.5])

# Storing parameters and sums in lists
monthly_data[[b_value]] <- list(
  month = b_value,
```

```
    x_bar = x_bar,
    best_params = best_params,
    total_rainfall = total_rainfall,
    total_obrainfall = total_obrainfall,
    transition_matrix = transition_matrix,
    rmse = rmse,
    mae = mae,
    op = op,
    mu1 = average_below_2.5,
    mu2 = mean(selected_datarain[selected_datarain >= 2.5]),
    re_op = re_op,
    re_mu1 = re_mu1,
    re_mu2 = re_mu2
  )

# Output parameters, sums, transfer matrices, and RMSE for each month
for (month_data in monthly_data) {
  cat(paste("Month", month_data$month, ":\n"))
  cat("Best Parameters:", month_data$best_params, "\n")
  cat("Total Rainfall:", month_data$total_rainfall, "\n")
  cat("Total_obrainfall:", month_data$total_obrainfall, "\n")
  cat("Transition Matrix:\n")
  print(month_data$transition_matrix)
  cat("op:", month_data$op, "\n")
  cat("mu1:", month_data$mu1, "\n")
  cat("mu2:", month_data$mu2, "\n")
  cat("\n")
}
```