DEVELOPMENT OF DESIGN RAINFALL MODEL USING AMS AND PDS IN PENINSULAR MALAYSIA

CHANG KIAN BOON

DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF ENGINEERING SCIENCE

INSTITUTE OF GRADUATE STUDIES
UNIVERSITY OF MALAYA
KUALA LUMPUR

2014

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Chang Kian Boon (I.C. No: 880923-52-5365)

Registration/Matrix No: KGA110021

Name of Degree: Master of Engineering Science

Title of Dissertation (“this Work”): Development of Design Rainfall Model using AMS and PDS in Peninsular Malaysia

Field of Study: Water Resources Engineering (Civil Engineering)

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship has been acknowledged in this Work;
(4) I do not have any knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature Date 4/1/2015

Subscribed and solemnly declared before,

Witness’s Signature Date

Name:
Designation:

ABSTRACT

In recent years, flood issues have become more frequent in Peninsular Malaysia. This study is focused on the determination of a better approach for deriving rainfall intensity-duration-frequency (IDF) relationship in Peninsular Malaysia, based on 60 selected rainfall stations in Peninsular Malaysia by using two data series: annual maxima series (AMS) and partial duration series (PDS), with their corresponding statistical distribution: generalized extreme value (GEV) distribution and generalized Pareto (GPA) distribution. Besides, the minimum inter-event time (MIT) for separation of rainfall events in Peninsular Malaysia need to be identified for extracting PDS. After some preliminary studies, it is found that to achieve these goals, 3 software packages must be developed as the amount of work required for extracting rainfall data and performing analysis are enormous: RainEMT (for extraction of rainfall data), RainIDF (for derivation of IDF relationship) and RainMap (to display design rainfall effectively). These softwares have been developed and the results of this study show that an MIT of 6 hours is suitable for separating rainfall events for extraction of PDS, and the model based on fitting PDS to the GPA distribution is found to be more suitable than the model based on fitting AMS with the GEV distribution for deriving rainfall IDF relationship in Peninsular Malaysia.

ABSTRAK

Dalam tahun-tahun kebelakangan ini, isu-isu banjir telah menjadi lebih kerap di Semenanjung Malaysia. Kajian ini bertumpu kepada penentuan kaedah yang paling sesuai untuk mendapat lengkongan IDF (intensity-duration-frequency) hujan di Semenanjung Malaysia, berdasarkan 60 stesen yang dipilih di Semenanjung Malaysia dengan menggunakan dua jenis siri data: annual maxima series (AMS) dan partial duration series (PDS), dengan kaedah taburan yang sepadan: taburan generalized extreme value (GEV) dan taburan generalized Pareto (GPA). Selain itu, masa minima (MIT) yang diperlukan untuk memisah data hujan kepada hujan individu perlu ditentukan bagi tujuan mengekstrak PDS. Bagi mencapai matlamat-matlamat ini, beberapa perisian (software) perlu dibina kerana jumlah kerja yang perlu dilakukan untuk mengekstrak data hujan adalah terlalu besar: RainEMT (untuk ekstrak data hujan), RainIDF (untuk menentukan lengkongan IDF) dan RainMap (untuk paparan hujan reka bentuk yang berkesan). Perisian-perisian ini telah dibina dan hasil kajian ini telah menunjukkan bahawa MIT yang berjumlah 6 jam adalah sesuai untuk digunakan untuk pemisahan hujan individu bagi tujuan mengekstrak PDS, serta model yang berdasarkan PDS dengan taburan GPA adalah lebih sesuai apabila dibandingkan dengan model yang berdasarkan AMS dengan taburan GEV, bagi tujuan penerbitan lengkongan IDF di Semenanjung Malaysia.

ACKNOWLEDGEMENTS

This work is a part of a research project (ER029- 2011A), funded by the Ministry of Higher Education (MOHE). All rainfall data used in this work is supplied by the Department of Irrigation and Drainage (DID) Malaysia.

I would like to thank my supervisors Assoc. Prof. Dr. Lai Sai Hin and Assoc. Prof. Dr. Faridah Binti Othman for giving me an opportunity to take part in this project. Their endless support has guided me through the last 2 to 3 years, which have been very trying at times. Dr. Lai has been a fantastic supervisor who provided me with opportunities to push myself harder and further than I normally would. He has been an inspiration and I thank him for that and his emotional support.

Last but not least, I would like to thank my family for giving me endless support and always believe in me.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

τ_2	*L*-CV / Coefficient of *L*-variation
τ_3	*L*-Skewness
τ_4	*L*-Kurtosis
X_T	Quantile value
T	Return period

ABBREVIATIONS

IDF	Intensity-duration-frequency
DDF	Depth-duration-frequency
IDAF	Intensity-duration-area-frequency
EV1	Gumbel / Extreme value type 1
EV2	Frechet / Extreme value type 2
EV3	Weibull / Extreme value type 3
GPA	Generalized Pareto
GEV	Generalized extreme value
PWM	Probability weighted moment
MOM	Method of moment
ML	Maximum likelihood
CDF	Cumulative distribution function
PDF	Probability density function
CV	Coefficient of variation
MIT	Minimum inter-event time
IETD	Inter-event time definition

AMS	Annual maxima series
PDS	Partial duration series
POT	Peak over threshold
ARI	Annual recurrence interval
AD	Anderson-Darling
KS	Kolmogorov-Smirnov
VBA	Visual Basic for Applications
DID	Department of Irrigation and Desalination
NIWA	National Institute of Water and Atmospheric Research
ISSE	Institute for the Study of Society and Environment

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

The rainfall intensity-duration-frequency (IDF) relationship is important for the determination of design rainfall to apply in water resources structural design, urban stormwater management and flood modeling. There are many different procedures available around the globe to develop the rainfall IDF relationship, however, several quantitative measures have to be used to determine the most suitable method for a certain region, such as Peninsular Malaysia. Due to the impact of climate change, with the increase of temperature around the globe, the extreme intensities of rainfall are now higher than ever (Trenberth, 2011). Rapid urbanization of certain areas such as Klang Valley in Peninsular Malaysia is also causing the increase of extreme rainfall intensities of these areas.

In Malaysia, flash flood events have occurred frequently in urban areas such as the Klang Valley. Damages and losses caused by flash floods have been mounting. This shows the importance of choosing the appropriate procedures to develop IDF relationship under the changing climate, and an update to the IDF curves in Peninsular Malaysia is essential.

1.2 Problem Statement

In the publication "Manual Saliran Mesra Alam – Design Rainfall" (DID, 2000), there are 26 and 10 IDF curves for urban areas in Peninsular Malaysia and East Malaysia, respectively. These curves need to be revised as they have not been revised since 1991 and the period of data where the curves derived from are as low as 7 years. A new set of IDF curve need to be generated as they were last developed about 20 years ago and due to the climate change impact in the late year.

There are different methods and approaches to generate rainfall IDF relationship or IDF curves, but the most important step is the fitting of probabilistic distribution into the rainfall data series. The current IDF curves available in Malaysia have been constructed based on Gumbel distribution and Annual Maxima Series (AMS). Researchers have been discussing about the type of distribution that best fit the rainfall data of certain area, with two type of data series (Annual Maxima Series (AMS) and Partial Duration Series (PDS)) (e.g. Ben-Zvi, 2009; Millington et al., 2011). The suitability of the probability distributions and data series could vary on different study regions due to different climate conditions. Although PDS (also known as peak over threshold approach) is commonly used in the application of flood modeling and analysis, it is relatively new in Malaysia for the purpose of deriving rainfall IDF relationships. Therefore, the best distribution with the type of rainfall data series that best fit the rainfall data in Peninsular Malaysia will be determined in this study. The distributions that are included in this study are generalized extreme values (GEV) distribution for AMS, and generalized Pareto (GPA) distribution for PDS, as these were the popular distribution discussed by other researchers (e.g. Koutsoyiannis and Baloutsos (2000) and Ben-Zvi (2009)). Goodness-of-fit test and *L*-moment ratio diagram are used to analyze and determine the best fitting distribution.

1.3 Objectives

The main objectives of this study are listed as below:

- To compare the different data series or approaches (i.e. Annual Maxima Series (AMS) and Partial Duration Series (PDS)) with their corresponding probability distributions (i.e. GEV and GPA distributions) in order to determine to best approach for derivation of rainfall IDF relationship in Peninsular Malaysia.
- To provide an update of IDF curves that applied the most suitable data series and distribution found in this study with the latest rainfall data from the department of irrigation and drainage (DID) Malaysia.
- To produce IDF curves for 60 rainfall stations in Peninsular Malaysia.

1.4 Scope of Study

The study area of this research covers the whole Peninsular Malaysia. Where the rainfall stations with at least 15 years of rainfall record and are still in operations are used for analysis. Moreover, the statistical or probability distributions that are covered in this study are GEV and GPA distributions. Two types of data techniques are determined, which is the annual maxima series (AMS) and partial duration series (PDS). The threshold selection for the PDS data is fixed at a certain level and the method to estimate the parameters of the distribution in order for comparison purpose is *L*-moment method (there are other parameter estimation methods such as: maximum likelihood (ML), method of moments (MOM) and probability weighted moments (PWM)). The L-moment ratio diagram is used as goodness-of-fit test as it provides an overall comparison of the methods used in this study. In order to develop the IDF relationship with the best fitting distribution and data series determined in this study, the one-step least squares method proposed by Koutsoyiannis et al. (1998) is used, although there are plenty of other methods available for generation of IDF relationship.

1.5 Significance of Study

In order to achieve the main objectives in this study, 3 different software packages have to be developed. These 3 reusable hydrological software packages will also benefit students, engineers and researchers around the world. This study will provide more comprehensive and reliable IDF curves in Malaysia for engineers, researchers, planners, etc. The outcomes of this study will also provide useful information to relevant government agency/department such as Department of Irrigation and Drainage (DID) for the formulation of new regulations for water infrastructure management as well as changes in design practices.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter reviews some of the most popular statistical or probability distributions for fitting of rainfall data, with different type of data series such as Annual Maxima Series (AMS) and Partial Duration Series (PDS), which is also known as peak over threshold (POT) approach. Besides, there are different type of quantitative measures that can be used to determine the most appropriate distribution and data series for Peninsular Malaysia, which are known as goodness-of-fit tests (such as the *L*-moment ratio diagram).

2.2 Statistical Methods

Several numbers of extreme event distributions are used in the field of hydrology. The methods that are included in this study are generalized extreme value (GEV) distribution and generalized Pareto (GPA) distribution.

The current method used by DID (Department of Irrigation and Drainage) Malaysia which is also one of the most popular method is Gumbel or EV1 distribution. In some studies where long period of data (e.g. more than 100 years) is not available, the EV1 distributions have found to be as fit as the GEV distribution to the rainfall data and EV1 is preferred in that case since it only has two parameters while GEV has three parameters (e.g. Mohymont et al., 2004).

In recent years, more studies (e.g. Koutsoyiannis and Baloutsos, 2000; Ben-Zvi, 2009; Millington et al., 2011) have shown that GEV distribution is more appropriate than EV1 distribution. These studies have expressed skepticism for the appropriateness of EV1 distribution for rainfall extremes, which show that the EV1 distribution tends to underestimate the largest extreme rainfall amounts. A study by Koutsoyiannis and Baloutsos (2000) shows that with a long record of annual rainfall (i.e. 136 years), the underestimation of EV1 distribution is quite substantial (e.g. 1:2). This fact must be considered as a warning against the widespread use of the EV1 distribution for rainfall extremes, therefore the importance of this study to analyze the best fitting distribution for Malaysia can be seen, especially when the current intensity-duration-frequency curves of DID Malaysia are based on the EV1 distribution.

Besides the concern of inappropriateness of EV1 distribution, recent studies have also demonstrated the increasing popularity of rainfall analysis and extreme hydrological events modeling based on generalized Pareto (GPA) distribution with partial duration series (PDS) or peak over threshold (POT) approach over the conventional method (e.g. EV1 and GEV distributions with annual maximum series). This can be seen by some studies (e.g. Koutsoyiannis and Baloutsos, 2000; Ben-Zvi, 2009), which show the superiority of GPA/PDS approach over the GEV/AMS approach. Therefore, this study is focuses on comparing the GPA/PDS model with the GEV/AMS model to see which model best represent the rainfall data of the Peninsular Malaysia region.

2.2.1 GEV distribution

The generalized extreme value (GEV) distribution is a family or combination of Gumbel (EV1), Frechet (EV2) and Weibull (EV3) distributions. It is worth noting that GEV distribution makes use of 3 parameters: location (ξ), scale (α) and shape (k).

In recent years, more studies (e.g. Koutsoyiannis and Baloutsos, 2000; Ben-Zvi, 2009; Millington et al., 2011) have shown that GEV distribution is more appropriate than other distributions that are commonly used for fitting annual maxima series (e.g. Gumbel and Log-Pearson Type III distributions). These studies have expressed skepticism for the appropriateness of Gumbel distribution (a family member of the GEV distribution) for rainfall extremes, which show that the Gumbel distribution tends to underestimate the largest extreme rainfall amounts. A study by Koutsoyiannis and Baloutsos (2000) shows that with a long record of annual rainfall (i.e. 136 years), the underestimation of Gumbel distribution is quite substantial (e.g. 1:2). Zalina et al. (2002) found that GEV distribution are the best fitted distribution for annual maxima series in Peninsular Malaysia.

The CDF (cumulative distribution function) and PDF (probability density function) of GEV (Hosking, 1997) are defined as:

(2.1)

$$F(x) = \exp\ \{-\left(1 - \frac{k(x-\xi)}{\alpha}\right)^{\frac{1}{k}}\}$$

(2.2)

$$f(x) = \alpha^{-1} \exp\ [-(1-k)y - \exp(-y)]$$

Where,

(2.3)

$$y = -k^{-1} \log\left[1 - \frac{k(x-\xi)}{\alpha}\right], \text{when } k \neq 0$$

x is the random variable of interest, ξ is the location parameter, α is the scale parameter and k is the shape parameter.

The location parameter, ξ represents the shift of a distribution in a given direction on the horizontal axis. The scale parameter, α shows how spread out the distribution is, and locates where the bulk of the distribution lies. The shape parameter, k shows the shape of the distribution and governs the tail of each distribution. It is the shape parameter that specifies one of the three asymptotic extreme-value distributions: EV1 ($k = 0$), EV2 ($k < 0$) or EV3 ($k > 0$).

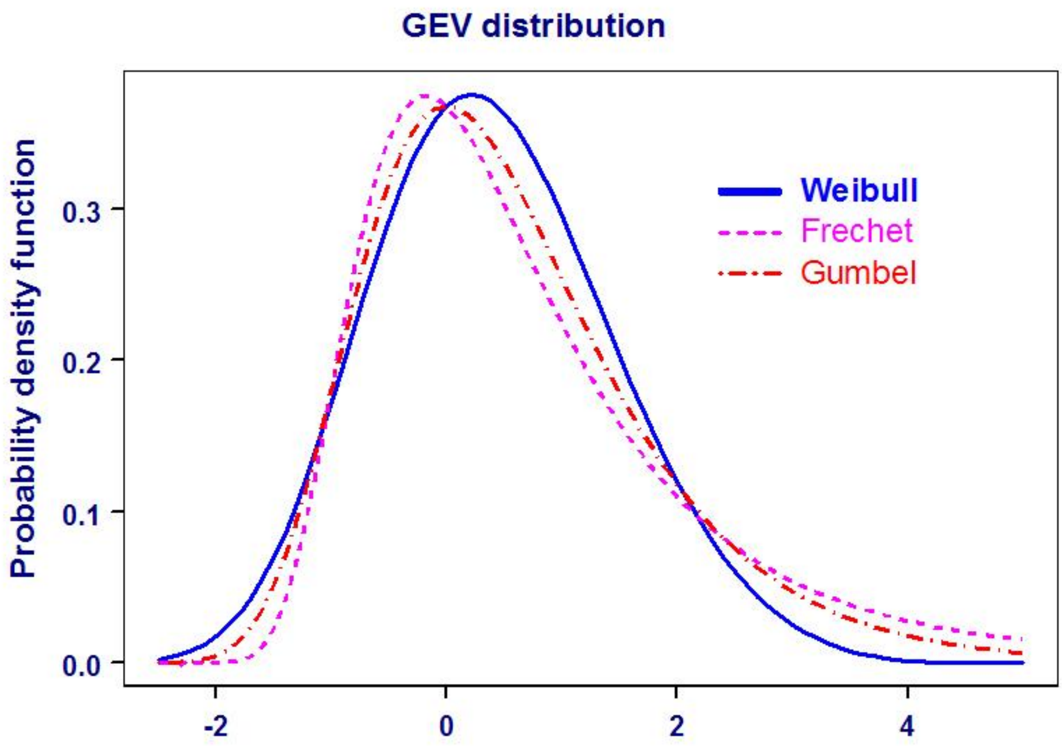


Figure 2.1: Shapes and Tails of GEV Distribution in the Form of Gumbel (EV1), Frechet (EV2) and Weibull (EV3). Adapted from ISSE (2011).

As shown in Figure 2.1, Gumbel is a distribution with a light upper tail and positively skewed. Frechet has a heavy upper tail and infinite higher order moments, and Weibull is a distribution with a bounded upper tail. EV1 is effective for small sample sizes, however when the sample size is greater than 50, GEV shows a better overall performance (Cunnane, 1989).

The extreme quantile, X_T of the corresponding return period, T and duration from the annual maxima series can be computed by using the inverse CDF of the GEV distribution:

(2.4)

$$X_T = \xi + \frac{\alpha\left\{1-\left(-\ln\left(1-\frac{1}{T}\right)\right)^{\kappa}\right\}}{\kappa}, \text{ when } \kappa \neq 0$$

(2.5)

$$X_T = \xi - \alpha \ln\left(-\ln\frac{1}{T}\right), \text{ when } \kappa = 0$$

It is worth noting that the GEV distribution turns into Gumbel distribution when the shape parameter, κ is equal to zero. Gumbel distribution is often chosen for its ease of use, since it only consists of 2 parameters (without the shape parameter). However, by implementing the automated distribution fitting function in RainIDF, all 3 parameters of GEV distribution can be estimated easily and thus, eliminates the 2-parameter advantage of the Gumbel distribution

2.2.2 Generalized Pareto distribution

The GPA distribution is one of the most popular distributions used for partial duration or POT analysis (e.g. Beguería, 2005; Ben-Zvi, 2009; Palynchuk and Guo, 2008). The CDF and PDF of GPA distribution as defined by Hosking and Wallis (1997) is:

(2.6)

$F(x) = 1 - e^{-y}$

(2.7)

$f(x) = \alpha^{-1} e^{-(1-\kappa)y}$

Where,

(2.8)

$y = -\kappa^{-1} \log\left\{1 - \frac{\kappa(x-\xi)}{\alpha}\right\}$, when $\kappa \neq 0$

(2.9)

$y = \frac{(x-\xi)}{\alpha}$, when $\kappa = 0$

ξ is the location parameter, α is the scale parameter and κ is the shape parameter. Special cases: $\kappa = 0$ is the exponential distribution with 2 parameters; $\kappa = 1$ is the uniform distribution on the interval $\xi \leq x \leq \xi + \alpha$.

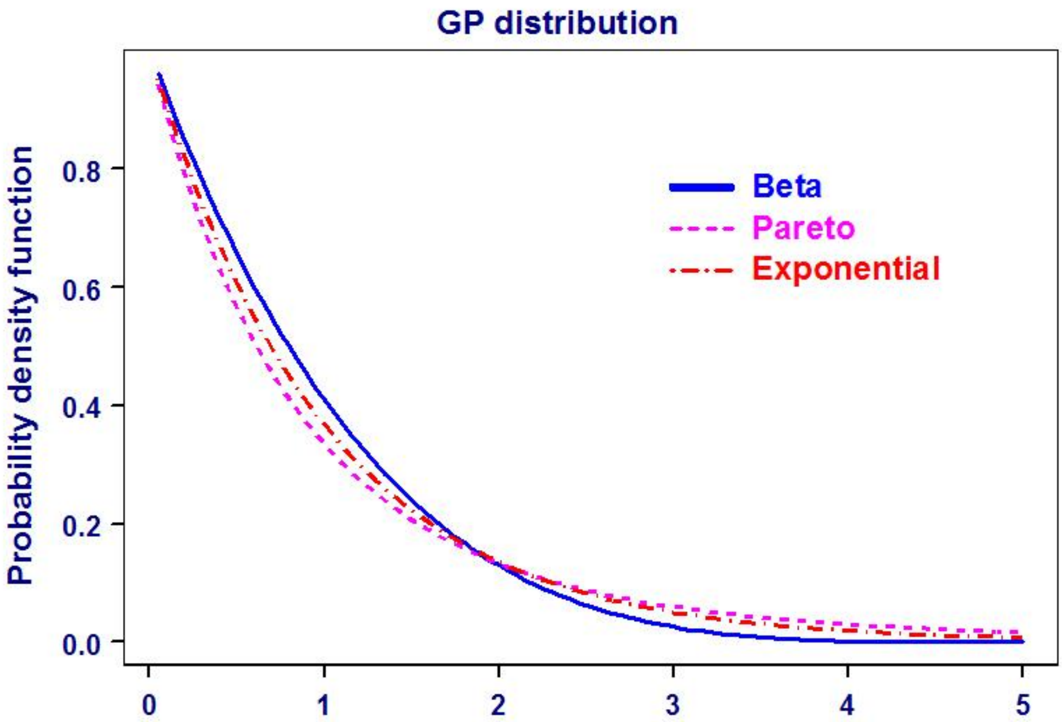


Figure 2.2: Shape of GPA distribution in the Form of Beta, Pareto and Exponential. Adapted from ISSE (2011).

As shown in Figure 2.2, Exponential is a light-tailed distribution with a "memoryless" property. Where as Pareto is a heavy-tailed distribution which sometimes called the power law and Beta is a bounded distribution. When $k > 0$ the generalized Pareto distribution is equivalent to the Pareto distribution and $k < 0$ gives the Beta distribution.

The location parameter, ξ is actually the threshold of the data series. The threshold value is usually known when fitting partial duration series to the GPA distribution. In this case, the 2 parameters (2-P) GPA distribution is used for fitting partial duration series, where only the scale and shape parameters are estimated with *L*-moments. Given that the average number of events per year, λ is known with the corresponding threshold x_0, the quantile of a specific duration with T-year return period can be calculated by:

(2.10)

$$X_T = x_0 + \frac{\alpha}{\kappa}\left[1 - \left(\frac{1}{\lambda T}\right)^{\kappa}\right], \text{ when } \kappa \neq 0$$

(2.11)

$$X_T = x_0 + \alpha \ln\left(\frac{1}{\lambda T}\right), \text{ when } \kappa = 0$$

The 2-P GPA distribution has a different formula for parameter estimation with *L*-moments compared to the 3 parameters (3-P) GPA distribution. Although the 2-P GPA distribution is preferred for fitting partial duration series, 2-P and 3-P are both included in RainIDF. The 2-P GPA distribution requires the user to specify the threshold values for each data series, while the 3-P GPA distribution estimates the location parameter from the data series.

2.3 Parameter Estimation Techniques

There are a few methods for fitting distributions to data, for example: MOM (method of moments), ML (maximum likelihood) and PWM (probability-weighted moments). They are used to estimate the parameters of the distributions. Environment Canada uses, and recommends the MOM technique for estimation of EV1 parameters (Millington et al., 2011). MOM is also known as one of the oldest, simplest and most popular method of estimating parameters. MOM is originally proposed by Gumbel (1941) to fit the EV1 or Gumbel distribution. Unfortunately, it is not so suitable to be used in the field of hydrology as most of the hydrologic variables are more or less skewed, therefore MOM represents a small or large loss of efficiency in estimation (Shin, 2009). According to Madsen et al. (1997b), in general, PDS-MOM should be used for negative shapes of the distribution fitted (heavy tailed); while one should use AMS-MOM for moderately positive shapes; and ML for large positive shapes (light tailed). Heavy tail and light tail reflects the rate of increase of the physical variable when its exceedance probability declines. Heavy tailed distributions increase faster than the exponential rate, while light tailed distributions are slower. However, Ben-Zvi (2009) has concluded that all good fits of the GPA distribution and most of the good fits of the Gumbel distribution by use of the PWM are found better than those by use of the MOM. PWM has been described as a simple and efficient method for fitting distributions to data (e.g., Koutsoyiannis et al., 1998).

Another method of parameter estimation is known as the *L*-moments, which has been used by Millington et al. (2011). *L*-moments are based on PWM, however *L*-moments provide a higher degree of accuracy and ease of use. As mentioned by Hosking and Wallis (1997), PWM uses weights of the CDF but it is difficult to interpret the moments as scale and shape parameters for probability distributions. The method of *L*-moments, rather than a completely new method, is actually a modification of PWM. PWM is used by the *L*-moments method to calculate parameters that are easy to interpret and also can be used to calculate parameters for statistical distribution. Millington et al. (2011) found that the method of *L*-moments is easy to work with and more reliable as they are less sensitive to outliers, thus provide an advantage. Rowinski (2001) has discovered that the MOM techniques are only able to apply to a limited range of parameters, whereas *L*-moments can be more widely used. Therefore, the method of *L*-moments has been chosen to estimate the parameters of the statistical distributions in this study.

2.3.1 Probability Weighted Moments

Probability weighted moments are required to calculate *L*-moments. Firstly, data is arranged in ascending order before it is applied to the following equations (Cunnane, 1989):

(2.12)

$$M_{100} = sample\ mean = \frac{1}{N}\sum_{i=1}^{N} Q_i$$

(2.13)

$$M_{110} = \frac{1}{N}\sum_{i=1}^{N} \frac{(i-1)}{(N-1)} Q_i$$

(2.14)

$$M_{120} = \frac{1}{N}\sum_{i=1}^{N} \frac{(i-1)(i-2)}{(N-1)(N-2)} Q_i$$

(2.15)

$$M_{130} = \frac{1}{N}\sum_{i=1}^{N} \frac{(i-1)(i-2)(i-3)}{(N-1)(N-2)(N-3)} Q_i$$

where, N is the sample size, Q is the data value and i is the rank of the value in ascending order.

2.3.2 *L*-moments

Equations for *L*-moments are listed as below (Cunnane, 1989):

(2.16)

$$\lambda_1 = L1 = M_{100}$$

(2.17)

$$\lambda_2 = L2 = 2M_{110} - M_{100}$$

(2.18)

$\lambda_3 = L3 = 6M_{120} - 6M_{110} + M_{100}$

(2.19)

$\lambda_4 = L4 = 20M_{130} - 30M_{120} + 12M_{110} - M_{100}$

The 4 *L*-moments are derived from PWM. As mentioned by Hosking and Wallis (1997), the *L*-moments (λ_1 and λ_2), the *L*-moment ratios (*L*-CV (τ_2), *L*-Skewness (τ_3) and *L*-Kurtosis (τ_4)) are the most useful quantities for summarizing probability distributions. Note that λ_1 and λ_2 are also known as the *L*-location or mean of the distribution and *L*-scale, respectively. The quantity of *L*-CV is analogous to the ordinary coefficient of variation. Instead of an abbreviation of "*L*-coefficient of variation", it would be more appropriate to describe *L*-CV as "coefficient of *L*-variation" in words (Hosking and Wallis, 1997). The equations of these *L*-moment ratios are given as (Hosking and Wallis, 1997):

(2.20)

$\tau_2 = L2/L1$

(2.21)

$\tau_3 = L3/L2$

(2.22)

$\tau_4 = L4/L2$

2.4 Sampling Techniques

Two kinds of samples are utilized in flood or rainfall frequency analysis: one that includes the peaks for every year in the observational period which is known as annual maximum series (AMS); the second one that includes all the peaks for events that exceed a given threshold which is known as partial duration series (PDS). PDS are usually derived from event maximum series (EMS). In the field of hydrology, the use of AMS has been very popular such as their fitting in GEV, Gumbel, Lognormal and Log-Pearson Type 3 distribution whilst PDS is not. However, some recent studies have evenly considered and included two kinds of sample (e.g., Katz et al., 2002; De Michele and Salvadori, 2005; Ben-Zvi, 2009).

In the early studies of statistic analysis, Gumbel (1954) considered AMS as more suitable than PDS. However, Kisiel et al. (1971) found that that an event series is more informative than a monthly or an annual series. Rasmussen et al. (1994) discovered PDS provides a more complete description of flood processes than AMS, which support this finding. According to Todorovic (1978), the construction of a stochastic model for AMS is hampered by many difficulties, whereas for PDS has a solid theoretical base. Pikand (1975) showed that the GPA distribution is a limiting form for the distribution of independent exceedances over high thresolds such as PDS. Whereas Smith (1984) found that for a large number of events in a year, the GEV distribution is a limiting form for the distribution of AMS.

A recent study by Ben-Zvi (2009) demonstrates the feasibility of using large partial duration series (PDS) derived from event maximum series (EMS) by fitting the GPA distribution to them, and its superiority over the conventional practice. The conventional practice here refers to GEV, Gumbel and Lognormal distribution, which are fitted to annual maximum series (AMS). Ben-Zvi (2009) found that the best fitted GPA distribution to PDS are superior to the other alternatives tested. Koutsoyiannis and Baloutsos (2000) also found that GPA/PDS approach is more appropriate than the other approach with AMS. We will be comparing two approaches in this study: GPA/PDS and GEV/AMS.

2.4.1 Rainfall Events Separation

Separation of rainfall data to their individual isolated events has to be performed to obtain EMS and PDS. A typical criterion used to separate individual rainfall event from continuous rainfall event is the period without rainfall between the rainfall events. This has been known as the minimum inter-event time (MIT). If the period or inter-event time is shorter than the MIT, the events will be identified as a single continuous event. On the other hand, when the inter-event time is longer or equal to the MIT, the rainfall events will be separated and isolated as different events. Ben-Zvi (2009) has chosen 24 hours inter-event time to separate rainfall events, while Adams et al. (1986) proposed MIT values between 1 and 6 hours for urban applications. Ahmad (2008) calculated a MIT value of 3 hours for rainfall events in Peninsular Malaysia. A similar approach to identify individual events based on an inter-event time is called inter-event time definition (IETD). IETD has the same function as MIT, with a different abbreviation. Palynchuk and Guo (2008) have selected an IETD of 6 hours for identification of individual rainfall events at Toronto, Canada. However, rainfall events separated by using inter-event time alone is not enough for this study as the separated rainfall events are considered as EMS, where PDS is needed in this study. A certain threshold has to be set to obtain PDS within the EMS.

2.4.2 Threshold Selection

A PDS's associated threshold may affect its properties. The lower the value of the threshold, the larger the PDS size and vice versa. A large sized PDS are more serially correlated and might be less suitable for probabilistic analysis; a small sized PDS will result in larger series, which could be less sensitive to sampling variations (Ben-Zvi, 2009). More details about the selection of the threshold level for this study are available in the methodology section of this report.

2.5 Goodness-of-Fit Tests

Goodness-of-fit tests are usually performed to find the distribution that best fit to a given data. However, these tests cannot be used to pick the best distribution, rather to reject possible distributions. They calculate test-statistics to analyze how well the data fits the given distribution. They are usually used to describe the differences between the observed data values and the expected values from the distribution being tested. Some of the goodness-of-fit tests are Anderson-Darling (AD) test, Kolmogorov-Smirnov (KS) test, and Chi-Squared (x^2) test which have been used by Millington et al. (2011) to compare the best fitting distribution.

2.5.1 Anderson-Darling Test

The Anderson-Darling test gives more weight to the tail of the distribution than KS test, which in turn leads to the AD test being stronger, and having more weight than the KS test (Millington et al., 2011). Stephens (1986) has also concluded that the AD is more powerful than other tests commonly used. By that means, the use of AD test alone is enough to determine the goodness of fit for the distributions in this study.

Some modifications to the AD test is proposed by Ahmed et al. (1988), with an emphasis on the upper or lower tail. However, Arshad et al. (2002) did not find this modification improves the power of the test. The test statistic of the Anderson-Darling (Stephens, 1986) is:

(2.23)

$$A^2 = \left(\Sigma\left[(2j-1)\ln\left(F_j\right) + (2n+1-2j)\ln(1-F_j)\right]/n - n\right) * \left(1 + \frac{0.2}{\sqrt{n}}\right)$$

Where A^2 is the statistic, j is the position in an ascending order of magnitudes, n is series size and F_j is the non-exceedance probability of the jth smallest value in the series, computed through the distribution fitted.

The lower the value of A^2, the better the distribution fits to the corresponding data. Significance levels for rejecting the fit of certain distributions has been presented by Stephens (1986), where $A^2 = 0.48$ indicates a level of 25% and $A^2 = 1.04$ indicates a level of 1%. A good fit is considered whenever $A^2 \leq 0.48$ while the fit is rejected when $A^2 \geq 1.04$.

When the fit is not rejected, the non-exceedance probabilities associated with the given recurrence intervals can be obtained through the relationship below:

(2.24)

$$F(x) = 1 - \frac{N}{nT(x)}$$

Where N is the number of years recorded and $T(x)$ is the recurrence interval of x.

2.5.2 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test statistic takes the greatest vertical distance from the empirical and theoretical cumulative distribution functions (CDF) into account. A hypothesis is rejected if the test statistic is greater than the critical value for a selected significance level. For example, at significance level of $\alpha = 0.05$, the corresponding critical value is 0.12555. The test statistic (D) of Kolmogorov-Smirnov test is:

(2.25)

$$D = \max(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i))$$

The KS test checks whether the two data samples come from the same distribution. Although the KS test can be served as a goodness-of-fit test, it is less powerful than AD test and not included in this study.

2.5.3 Chi-Squared Test

According to Cunnane (1989), the Chi-Squared test has not been considered as a high power statistical test and is not very useful. The test statistic (x^2) resembles a normalized sum of squared deviations between observed and theoretical frequencies. It is based on binned data where the number of bins (k) is given by:

(2.26)

$$k = 1 + log_2 N$$

where N = sample size

The test statistic (x^2) for Chi-Squared test is determined by:

(2.27)

$$x^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

where,

O_i is the observed frequency

E_i is the expected frequency given by,

$E_i = F(x_2) - F(x_1)$, x_1 and x_2 are the limits of the ith bin

At significance level of $\alpha = 0.05$, the critical value is 12.592. In the case where the test statistic (x^2) is greater than this critical value, the hypothesis is rejected.

2.6 *L*-moment Ratio Diagram

An *L*-moment ratio diagram consists of *L*-Skewness (τ_3) and *L*-Kurtosis (τ_4) of the sample data set, which is plotted against constant lines and points of known statistical distributions of interest. It is a common method used in regional frequency analysis where the fitting of the observed data is determined by comparing the values against the fitted regional data.

Several researchers (e.g. Ben-Zvi and Azmon, 1996 and Millington et al., 2011) have used *L*-moment diagrams in conjuction with goodness-of-fit tests. Ben-Zvi and Azmon (1996) have first applied the *L*-moment diagram in order to screen out the inappropriate candidate distributions, and then the Anderson-Darling test was used to examine the descriptive performance of the screened distributions. They have concluded such two-stage procedure, which applies quantitative measures in both stages, would reduce the subjectivity involved with the selection of a probability distribution, thus improve the credibility of the predicted high discharges. According to Ben-Zvi and Azmon (1997), a selection of distribution that only been screened through with *L*-moment diagram still applies certain subjective considerations, and is advisable to strengthen the share of objective by a joint use of another quantitative measure, such as the goodness-of-fit test (e.g. Anderson-Darling test).

In this study, we choose to use *L*-moment diagram to compare the results of all rainfall stations. If the result is hard to differentiate the goodness-of-fit of two sets of data, then only Anderson-Darling test will be used. The purpose of this approach is to obtain a bigger image about the quantitative measures of all the included distributions in this study. Three parameter distributions (i.e. GEV, GPA) are plotted as a line that corresponds to the varying shape parameters. The expressions for τ_4 are given as functions of τ_3, and are approximated as (Hosking and Wallis 1997):

For GEV distribution:

(2.28)

$$\tau_4 = 0.10701 + 0.1109\tau_3 + 0.84838\tau_3^2 - 0.06669\tau_3^3 + 0.00567\tau_3^4 - 0.04208\tau_3^5 + 0.03673\tau_3^6$$

For GPA distribution:

(2.29)

$$\tau_4 = 0.20196\tau_3 + 0.95924\tau_3^2 - 0.20096\tau_3^3 + 0.04061\tau_3^4$$

2.7 Rainfall Intensity-Duration-Frequency Relationship

The rainfall Intensity-Duration-Frequency (IDF) relationship is known as one of the most commonly used tools in the field of hydrology and water resources engineering. The establishment of such relationship was done as old as in 1932 (Bernard, 1932) and can be represented in the form of empirical IDF formulas and IDF curves, which are commonly required for design purposes of water resources projects. The IDF relationship is actually a mathematical relationship between the rainfall intensity, the duration, and the return period (Koutsoyiannis et al., 1998).

One of the most challenging problems faced when constructing a reliable IDF curve is the absence of long record rainfall data. Therefore, to reduce the error and uncertainties, only rainfall stations with more than 15 years of recorded data is included in this study, where most of the chosen stations have around 30 years of rainfall data record. Besides, the most suitable statistical distribution and data series for Peninsular Malaysia will be determined before they are applied in the development of IDF curves in this study.

2.7.1 The Empirical IDF Formula

The empirical IDF formula used by Bernard (1932) is known as:

(2.30)

$$i = \frac{\lambda T^{\kappa}}{(d+\theta)^{\eta}}$$

Where i is the rainfall intensity (mm/hour) of the corresponding d-duration (hour) and T-year return period. All parameters must be positive values and $0 < \eta < 1$.

CHAPTER 3

METHODOLOGY

3.1 Introduction

In order to develop a new set of IDF curves for Peninsular Malaysia, the best fitting distribution and its corresponding data series must first be determined (either GEV/AMS or GPA/PDS). Then, the chosen distribution and data series, along with the selected rainfall stations in Peninsular Malaysia, return periods and durations of intensity will be used to develop rainfall IDF relationship with the proposed one-step least squares method (Koutsoyiannis et al., 1998). To achieve the steps mentioned, there are 3 software packages to be developed to achieve their purposes. Therefore the methodology is split into 3 parts based on each software and the method or steps involved to achieve our targeted results.

3.2 General Research Design and Procedure

Since our study area includes the whole Peninsular Malaysia, the first step is to collect hydrological data from Department of Irrigation and Drainage (DID) Malaysia. After the data is obtained, it has to be filtered (from empty records and outliers) and 60 proper sites are selected based on certain criteria (such as completeness of data; years of record; location of the station). 16 out of these 60 stations will be used to perform analysis on identifying the most suitable MIT (minimum inter-event time) for separation of rainfall events in Peninsular Malaysia, as it is needed to extract PDS data. After that, two different series of data (AMS and PDS within EMS) has to be extracted from these 60 stations for fitting into probabilistic distributions. The fitting of these distributions is tested with *L*-moment ratio diagram and goodness-of-fit tests (if necessary). The GEV and GPA distribution with their relative data series are used to construct updated IDF curves for the 60 rainfall stations in Peninsular Malaysia for comparison. The general research procedure of this study is shown in Figure 3.1.

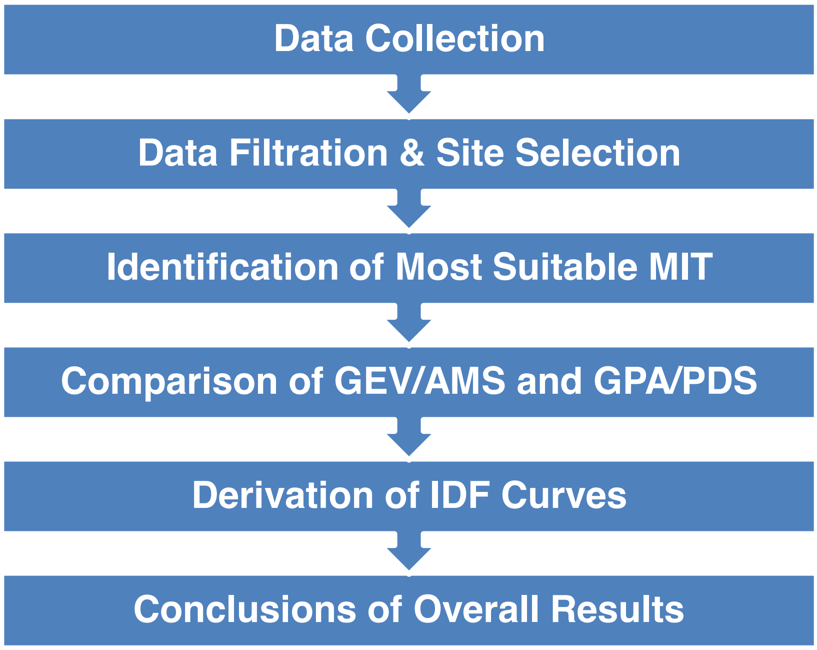


Figure 3.1: GEV/AMS and GPA/PDS Development

3.3 Study Area

The study area covers the whole Peninsular Malaysia, which is within the tropical climate. The climate of Peninsular Malaysia is governed by the northeast and southwest monsoons. The northeast monsoon commences from early November until March, while the southwest monsoon usually starts from early June until September. They are separated by two short inter-monsoon seasons which heavy rainfall is recorded (Ahmad, 2008). The 60 selected stations that are used in this study are shown in APPENDIX A.

3.4 Development of Software Packages

In this study, all the methodology, calculation and algorithm are built into software packages so that the community can reuse them in future. Therefore, we will discuss each of these software and the steps used to achieve our results. These software packages are:

- RainEMT – A Microsoft Access add-in used to extract and process the required rainfall data such as AMS and PDS.
- RainIDF – A Microsoft Excel add-in that is used to generate IDF relationship and plot IDF curves based on GEV/AMS and GPA/PDS approach.
- RainMap – A standalone software that showcases the locations of rainfall stations with their corresponding design rainfall and coefficients of empirical IDF formula.

3.5 RainEMT

RainEMT (Rainfall Event Mining Tool) is a database tool developed in Microsoft Access using Visual Basic for Applications. It allows easy extraction of large-scale and meaningful rainfall event data from a large time series rainfall dataset (i.e. 5, 10 or 15 minutes interval). The rainfall events are separated with a user-defined minimum inter-event time (MIT), which is also known as inter-event time definition (IETD). The output data includes the following: total yearly or monthly rainfall events with minimum storm duration categories; annual maximum rainfall or annual maxima series; event maximum rainfall with threshold selection (partial duration series / peak over threshold). These extracted data are used for various applications such as rainfall pattern analysis, climate change analysis, extreme value statistics, development of rainfall intensity-duration-frequency relationship, etc.

3.5.1 Software description

In recent years, the use of event maximum rainfall and partial duration series (PDS) or peak over threshold (POT) approach for development of rainfall intensity-duration-frequency, intensity-duration-area frequency and depth-duration-frequency relationship has gained an increasing popularity (e.g. Ben-Zvi, 2009; De Michele et al., 2011; Palynchuk and Guo, 2008). Rainfall PDS or POT data are also used in climate change analysis and simulation (Kyselý et al., 2010). However, technical difficulties faced in obtaining such data have caused drawbacks of large-scale application of event maxima and PDS/POT approach. Moreover, more analysis on rainfall events and patterns should be performed to obtain more accurate and convincing parameters especially on selection of threshold level (Beguería, 2005). Such analysis also requires extensive separation and calculation of rainfall events, which exceed the capabilities of current hydrological database software available in the industry.

A new database software tool, RainEMT is introduced to overcome these issues where such data can be obtained easily with just one click. To obtain or calculate rainfall events from rainfall data, a minimum inter-event time (MIT) or inter-event time definition (IETD) is used. With RainEMT, more detailed analysis on rainfall events can be performed and large-scale extraction of multiple interval partial duration series / peak over threshold rainfall data is made possible. Moreover, the interval of input rainfall data can be as short as 5 minutes, which gives a very accurate and detailed output data.

RainEMT is developed in Microsoft Access by using Visual Basic for Applications. The reason why Microsoft Access is used instead of Microsoft Excel is that current version of Microsoft Excel (Excel 2010) has a limit of around 1 million rows of data, and the older versions are even fewer than this. A set of 5 minutes interval rainfall data with 40 years of record length has more than 4 millions rows of data. This led us to choose Microsoft Access to serve our purpose, where there is no limitation on maximum rows of data. However, a database created in current version of Microsoft Access (Access 2010) has a file size limit of 2 GB. Although a dataset with 40 years of 5 minutes rainfall is only around 100 MB in size, one should take note of this limitation when importing several datasets into the same database. RainEMT is built as a form application in Microsoft Access (Figure 3.2). User can choose the table contains the imported rainfall data, type of data to be extracted and key in or choose the desired parameters.

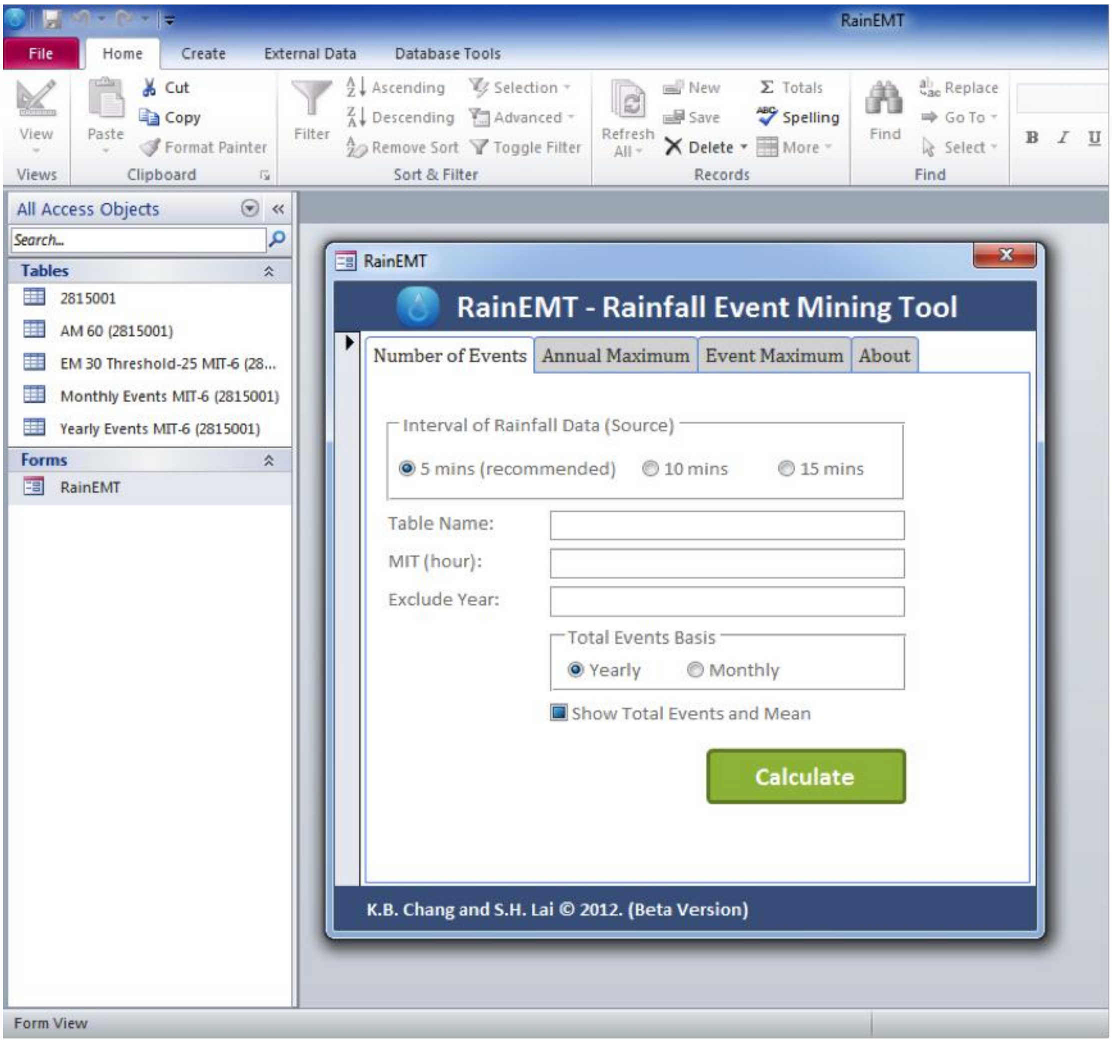


Figure 3.2: Screen interface of RainEMT in Microsoft Access.

3.5.2 Importing rainfall data

The time series data that are used as input data for RainEMT is exported from hydrological database software such as NIWA Tideda. In this paper, the sample dataset contains time series data of 5 minutes interval rainfall (ranging from 6/29/1970 to 12/9/2011) in a comma separated sheet (.csv) format. It is exported from its raw data in NIWA Tideda for Station 2815001 in Selangor, Malaysia. Note that the dataset file can be any format, as long as Microsoft Access supports it. This dataset is then imported into RainEMT through Microsoft Access in table '2815001'. Although RainEMT supports 5, 10 and 15 minutes rainfall data, the use of 5 minutes interval data is highly recommended in order to produce more reliable result.

The input data is categorized into three fields: 'Date', 'Time' and 'Rain mm'. These are the default field names for time series rainfall data exported from NIWA Tideda. One should rename the field names for data exported from other hydrological database software, if they are different from the field names mentioned earlier. The 'Date' field must be in 'MM/DD/YYYY' format (e.g. 6/29/1970), where 'Rain mm' shows that the depth of rainfall is in millimeter measurement. Although the 'Time' field is usually included in the dataset exported from hydrological database software, it is not used in the algorithmic calculation of RainEMT. Therefore, user may choose to exclude the 'Time' field when importing data into RainEMT to reduce disk usage. Addition of primary key when importing dataset to RainEMT is recommended as it helps Microsoft Access to sort multiple rows of data in the database (Fig. 3.3).

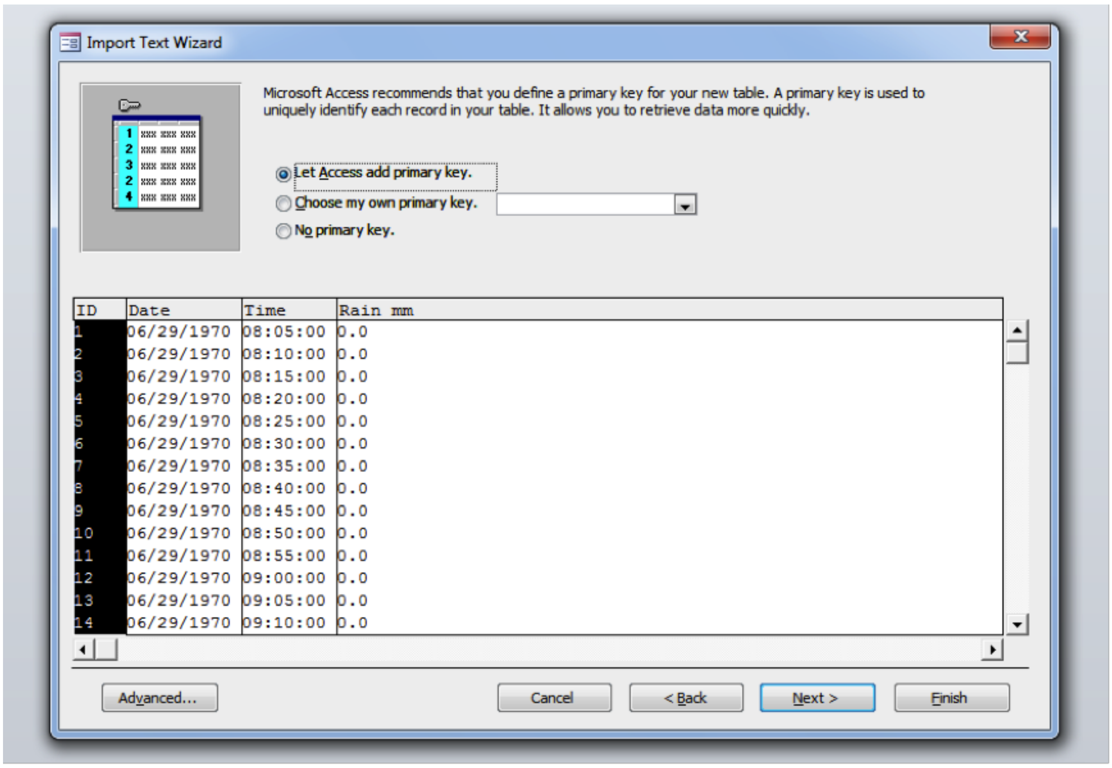


Figure 3.3: Addition of primary key when importing dataset into RainEMT using Import Text Wizard in Microsoft Access.

3.5.3 Separation of rainfall events with minimum inter-event time

Rainfall data exported from hydrological database software contains zero value (dry period) and non-zero value (wet period). To separate rainfall events, a guideline of minimum dry period length between two wet periods has to be set. Such guideline is known as minimum inter-event time (MIT) (e.g. Dunkerley, 2008; Haile et al., 2010; Haile et al., 2011) or inter-event time definition (IETD) (e.g. Balistrocchi and Bacchi, 2011; Palynchuk and Guo, 2008). If the dry period between two wet periods is equal or longer than the minimum inter-event time, they are considered as two separated rainfall events; if the dry period is shorter than the minimum inter-event time, they are considered as a single rainfall event.

The algorithm of RainEMT loops the input data from the first record until the end of record, year by year. In each year, to be considered as a rainfall event (including the first and last events), each event must begin and end with a dry period that satisfied the minimum inter-event time (in hour/hours) specified by the user. In some cases, users may choose to exclude unwanted years within the dataset (e.g. incomplete years). At the end of the loop, RainEMT creates a new table and insert the output data into the table. The table name contains the abbreviation and value of chosen parameters, and the table name of input data. For example, 'Yearly MIT-6 (2815001)' shows that the output yearly events are based on an MIT of 6 hours from the input data of table 2815001. The output data in the table can be executed directly from Microsoft Access to Microsoft Excel for further analysis if needed.

3.5.4 Application of RainEMT: extracting yearly and monthly rainfall events

RainEMT is capable of calculating total rainfall events in a yearly or monthly basis based on a minimum inter-event time (in hour/hours) specified by user. Generally, small minimum inter-event time will give a higher number of separated rainfall events and vice versa. Beside total events, number of events categorized into their minimum storm duration (i.e. 0.5, 1, 2, 3, 4, 5, 6, 12 and 24 hours) is also calculated. Such minimum storm duration categories allow a very detailed study and analysis of rainfall events with their corresponding minimum storm durations, with a variation on minimum inter-event time.

In the case where the user is interested of the mean and total yearly or monthly events for the entire period of record, RainEMT can calculate and include them in the last row of the output data. On the other hand, user also has the ability to define the year/years to be excluded from calculation and output data (Figure 3.4). The output yearly and monthly data are listed in new tables created in RainEMT. For yearly data, each row of data will start with the year of the data, followed by total events and events with minimum storm duration categories. Monthly data is listed in the same way as yearly data did, except that it contains a month field between the year and total events fields. They are organized in the way that is very convenient for plotting of rainfall trend or pattern charts when exported to Microsoft Excel.

To demonstrate the usage of RainEMT, yearly (Figure 3.5) and monthly (Figure 3.6) rainfall events of Station 2815001 are extracted into table ‘Yearly Events MIT-6 (2815001)’ and table ‘Monthly Events MIT-6 (2815001)’ with minimum inter-event time (MIT) of 6 hours. The parameters for extraction of these data is shown in Fig. 3.4, where year 1970, 1986, 1989 and 1990 are excluded due to a large amount of missing data in these years. These tables are exported to Microsoft Excel where charts for yearly rainfall events and monthly rainfall events of year 2010 are plotted (Figure 3.7 and Figure 3.5).

By observing Figure 3.7, we can see that number of yearly events seems to be around 150 events per year, where year 1982 and 2005 are unusual cases. Such cases could also indicate that there might be a significant amount of missing data, which should be examined from the hydrological database software that contains the raw data. The monthly rainfall patterns of year 2010 (Figure 3.8) shows that February and October are the dry months with the least number of rainfall events. These results are based on a MIT of 6 hours and will vary accordingly with a different value of MIT.

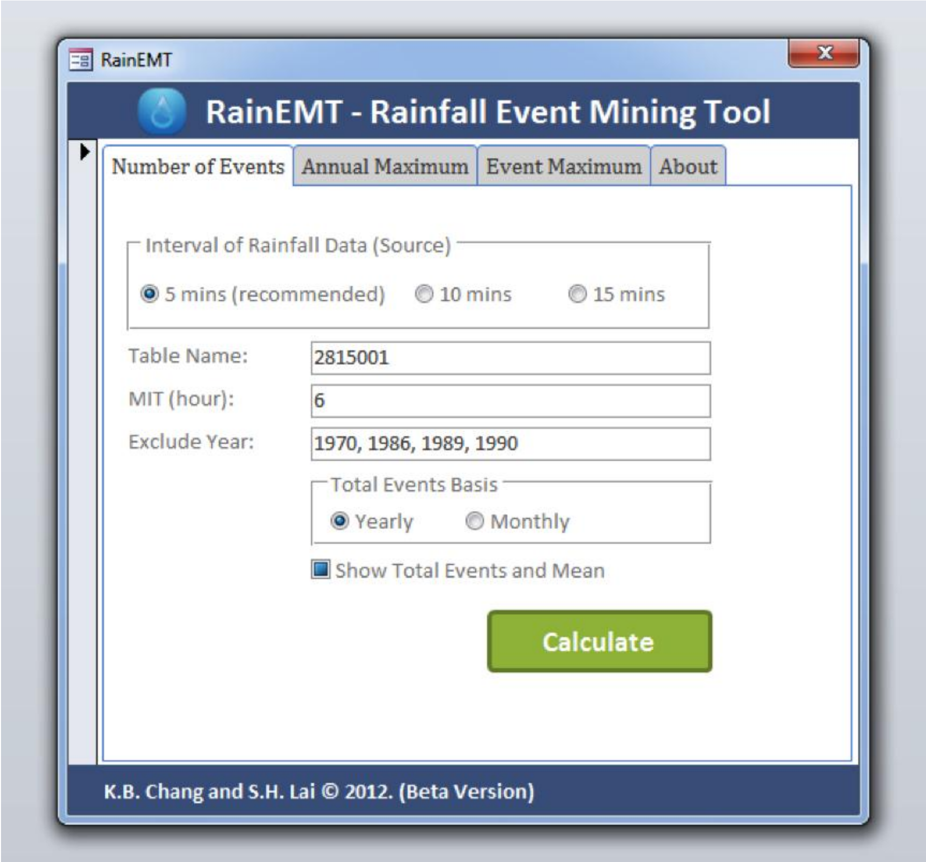
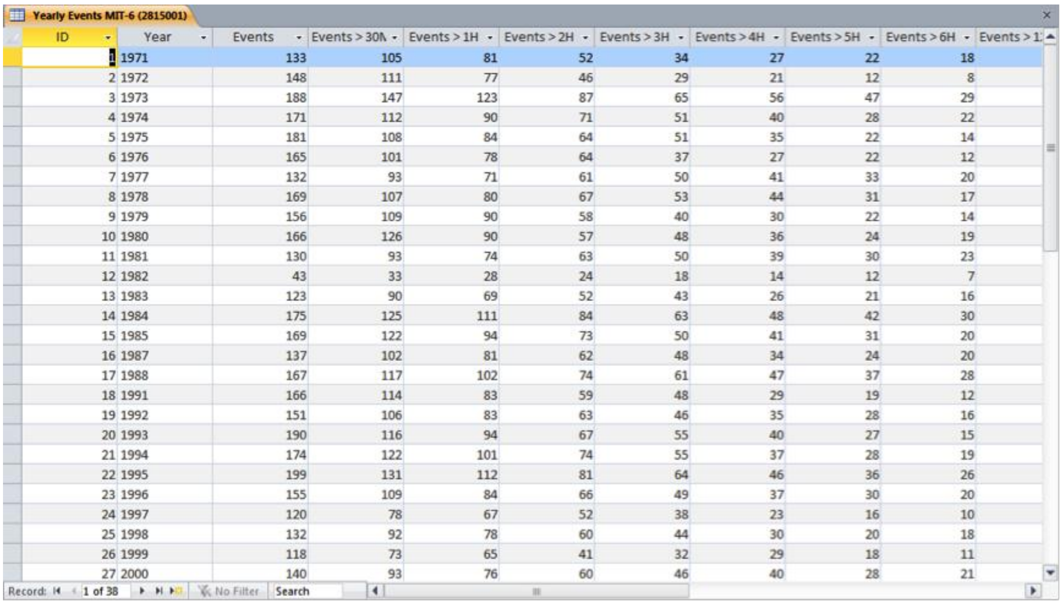


Figure 3.4: User interface for extraction of number of yearly and monthly events.

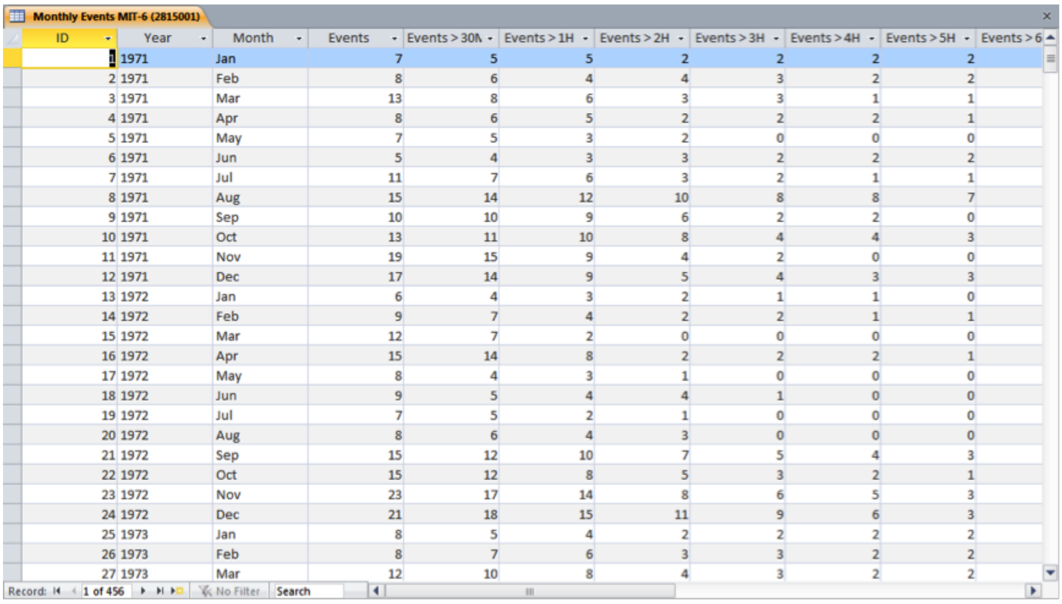


Yearly Events MIT-6 (2815001)

ID	Year	Events	Events > 30M	Events > 1H	Events > 2H	Events > 3H	Events > 4H	Events > 5H	Events > 6H	Events > 1:
1	1971	133	105	81	52	34	27	22	18	
2	1972	148	111	77	46	29	21	12	8	
3	1973	188	147	123	87	65	56	47	29	
4	1974	171	112	90	71	51	40	28	22	
5	1975	181	108	84	64	51	35	22	14	
6	1976	165	101	78	64	37	27	22	12	
7	1977	132	93	71	61	50	41	33	20	
8	1978	169	107	80	67	53	44	31	17	
9	1979	156	109	90	58	40	30	22	14	
10	1980	166	126	90	57	48	36	24	19	
11	1981	130	93	74	63	50	39	30	23	
12	1982	43	33	28	24	18	14	12	7	
13	1983	123	90	69	52	43	26	21	16	
14	1984	175	125	111	84	63	48	42	30	
15	1985	169	122	94	73	50	41	31	20	
16	1987	137	102	81	62	48	34	24	20	
17	1988	167	117	102	74	61	47	37	28	
18	1991	166	114	83	59	48	29	19	12	
19	1992	151	106	83	63	46	35	28	16	
20	1993	190	116	94	67	55	40	27	15	
21	1994	174	122	101	74	55	37	28	19	
22	1995	199	131	112	81	64	46	36	26	
23	1996	155	109	84	66	49	37	30	20	
24	1997	120	78	67	52	38	23	16	10	
25	1998	132	92	78	60	44	30	20	18	
26	1999	118	73	65	41	32	29	18	11	
27	2000	140	93	76	60	46	40	28	21	

Record: 1 of 38 No Filter Search

Figure 3.5: Table contains output data of yearly events extracted from Station 2815001 with MIT of 6 hours.



Monthly Events MIT-6 (2815001)

ID	Year	Month	Events	Events > 30M	Events > 1H	Events > 2H	Events > 3H	Events > 4H	Events > 5H	Events > 6
1	1971	Jan	7	5	5	2	2	2	2	
2	1971	Feb	8	6	4	4	3	2	2	
3	1971	Mar	13	8	6	3	3	1	1	
4	1971	Apr	8	6	5	2	2	2	1	
5	1971	May	7	5	3	2	0	0	0	
6	1971	Jun	5	4	3	3	2	2	2	
7	1971	Jul	11	7	6	3	2	1	1	
8	1971	Aug	15	14	12	10	8	8	7	
9	1971	Sep	10	10	9	6	2	2	0	
10	1971	Oct	13	11	10	8	4	4	3	
11	1971	Nov	19	15	9	4	2	0	0	
12	1971	Dec	17	14	9	5	4	3	3	
13	1972	Jan	6	4	3	2	1	1	0	
14	1972	Feb	9	7	4	2	2	1	1	
15	1972	Mar	12	7	2	0	0	0	0	
16	1972	Apr	15	14	8	2	2	2	1	
17	1972	May	8	4	3	1	0	0	0	
18	1972	Jun	9	5	4	4	1	0	0	
19	1972	Jul	7	5	2	1	0	0	0	
20	1972	Aug	8	6	4	3	0	0	0	
21	1972	Sep	15	12	10	7	5	4	3	
22	1972	Oct	15	12	8	5	3	2	1	
23	1972	Nov	23	17	14	8	6	5	3	
24	1972	Dec	21	18	15	11	9	6	3	
25	1973	Jan	8	5	4	2	2	2	2	
26	1973	Feb	8	7	6	3	3	2	2	
27	1973	Mar	12	10	8	4	3	2	2	

Record: 1 of 456 No Filter Search

Figure 3.6: Table contains output data of monthly events extracted from Station 2815001 with MIT of 6 hours.

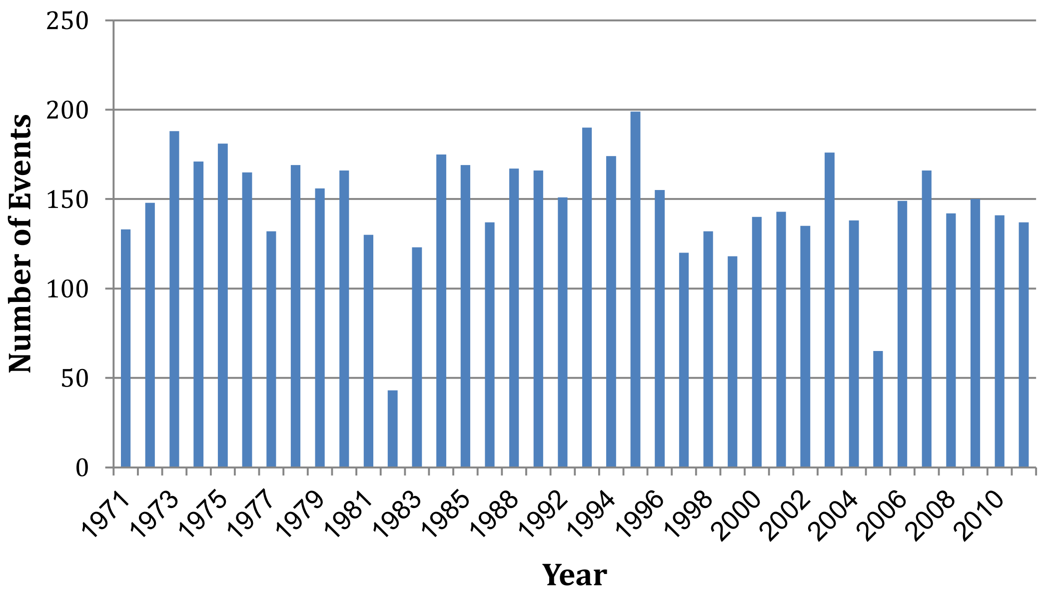


Figure 3.7: Histogram shows number of yearly events with MIT of 6 hours for Station 2815001.

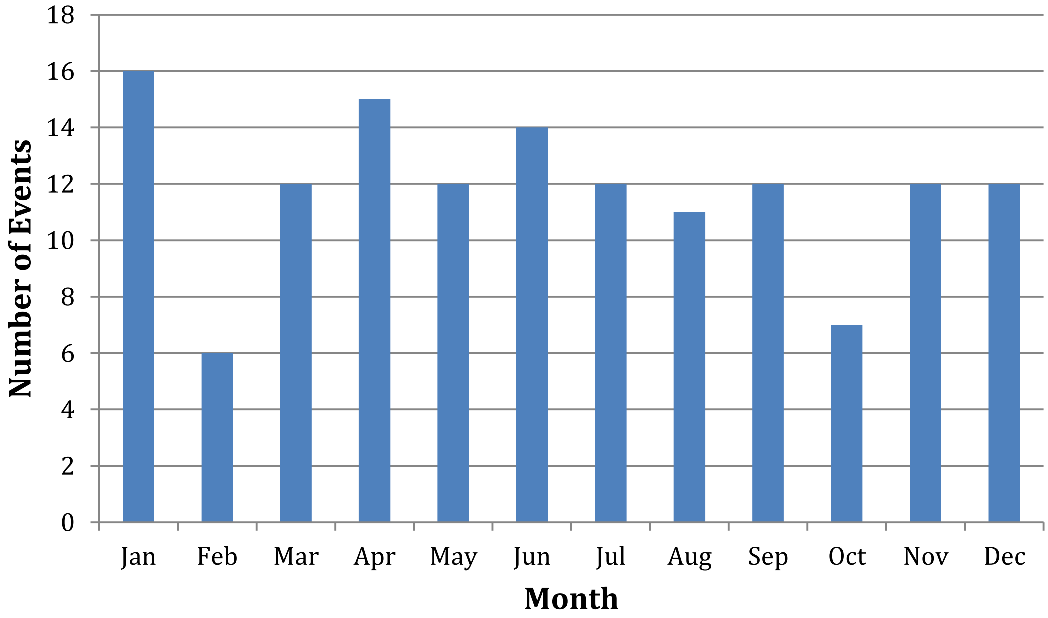


Figure 3.8: Histogram shows number of monthly events at year 2010 with MIT of 6 hours for Station 2815001.

3.5.5 Application of RainEMT: extracting annual and event maximum rainfall

3.5.5.1 Annual maximum rainfall

The annual maximum or extreme rainfall can be extracted from most hydrological database software (e.g. NIWA Tideda). Although the extraction of annual maximum series does not require the use of minimum inter-event time, this function is also included in RainEMT for the ease of extracting multiple types of extreme rainfall data (with event maximum rainfall or partial duration series / peak over threshold). The annual maximum rainfall for a specific interval is extracted together with its occurrence date into a new table created in RainEMT.

The common usages of annual maximum rainfall are extreme value analysis, trend and statistics (e.g. Adamowski and Bougadis, 2003; Katz et al., 2002; Kuo et al., 2011;

Villarini et al., 2011). Extreme value statistics have played an important role in engineering practice for water resources design and management (Katz et al., 2002). For example, the development of rainfall intensity-duration-frequency or depth-duration-frequency curves from annual maximum rainfall allows the prediction of maximum rainfall intensity for a specific return period (e.g. Ben-Zvi, 2009; Koutsoyiannis and Baloutsos, 2000; Madsen et al., 2009; Overeem et al., 2008; Van de Vyver and Demarée, 2010).

By entering the table name of the source rainfall data and the desired retrieval interval (Figure 3.9), annual maximum rainfall of 60 minutes interval is retrieved from Station 2815001. The output table 'AM 60 (2815001)' contains the retrieved annual extreme rainfall and the date where such extreme event occurred (Figure 3.10). Aside from application for rainfall intensity-duration-frequency curve and extreme pattern analysis, dates of such occurrence of extreme rainfall are also useful for integration in flood analysis.

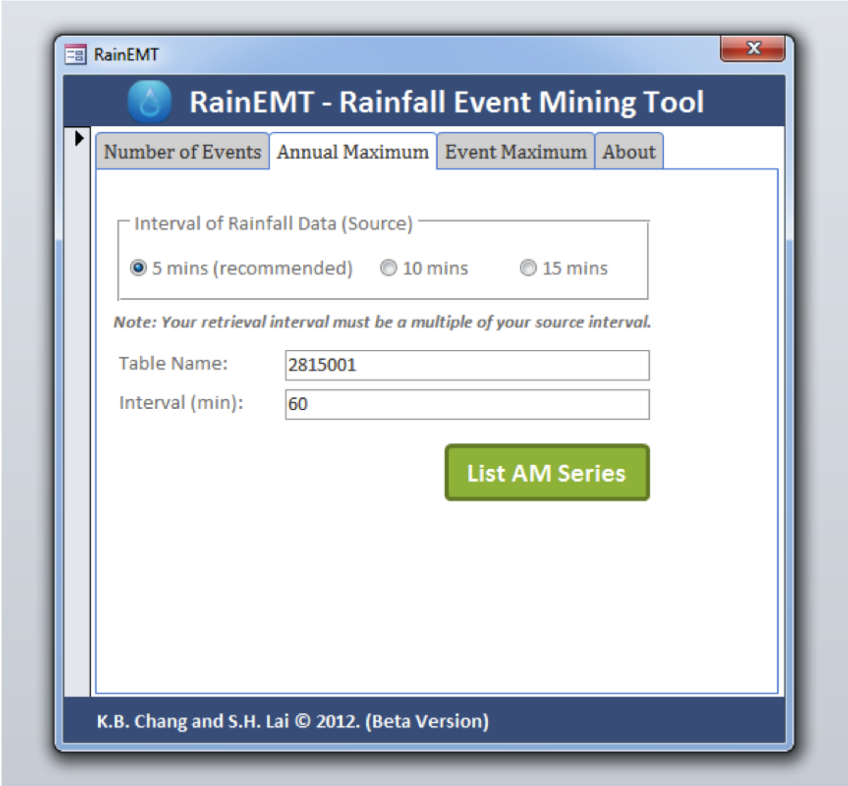
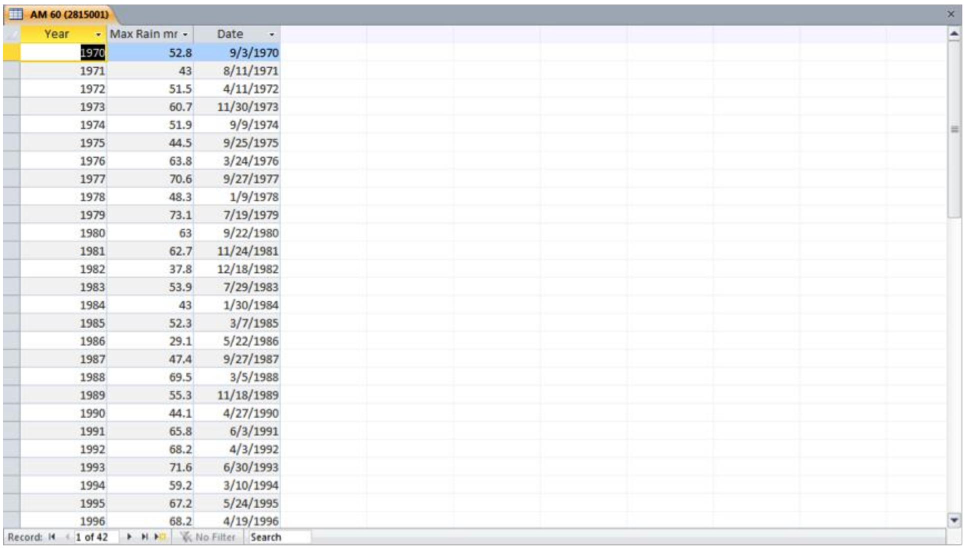


Figure 3.9: User interface for extraction of annual maximum rainfall.


AM 60 (2815001)

Year	Max Rain mr	Date
1970	52.8	9/3/1970
1971	43	8/11/1971
1972	51.5	4/11/1972
1973	60.7	11/30/1973
1974	51.9	9/9/1974
1975	44.5	9/25/1975
1976	63.8	3/24/1976
1977	70.6	9/27/1977
1978	48.3	1/9/1978
1979	73.1	7/19/1979
1980	63	9/22/1980
1981	62.7	11/24/1981
1982	37.8	12/18/1982
1983	53.9	7/29/1983
1984	43	1/30/1984
1985	52.3	3/7/1985
1986	29.1	5/22/1986
1987	47.4	9/27/1987
1988	69.5	3/5/1988
1989	55.3	11/18/1989
1990	44.1	4/27/1990
1991	65.8	6/3/1991
1992	68.2	4/3/1992
1993	71.6	6/30/1993
1994	59.2	3/10/1994
1995	67.2	5/24/1995
1996	68.2	4/19/1996

Record: 1 of 42 No Filter Search

Figure 3.10: Annual maximum rainfall of 60 minutes interval extracted from Station 2815001.

3.5.5.2 Event maximum rainfall and partial duration series

The maximum depth of a specific interval (e.g. 30 minutes) of a rainfall event is known as event maximum rainfall. In the case where the storm duration of a rainfall event is shorter than the specified interval, the event maximum rainfall of that interval is the total depth of the entire rainfall event. The extraction of event maximum rainfall is the most challenging and time consuming part of any analysis that requires such data. This is the main reason that leads to the development of RainEMT, and is also one of the most powerful and useful features of RainEMT.

By specifying the required parameters, an event maximum series can be extracted with just one click in RainEMT (Figure 3.11). When a zero threshold value is used, the extracted data includes event maximum series for all rainfall events separated by the specified minimum inter-event time. Meanwhile, a non-zero threshold value will include event maximum series that exceed or equal to the specified threshold value only (which is known as partial duration series or peak over threshold).

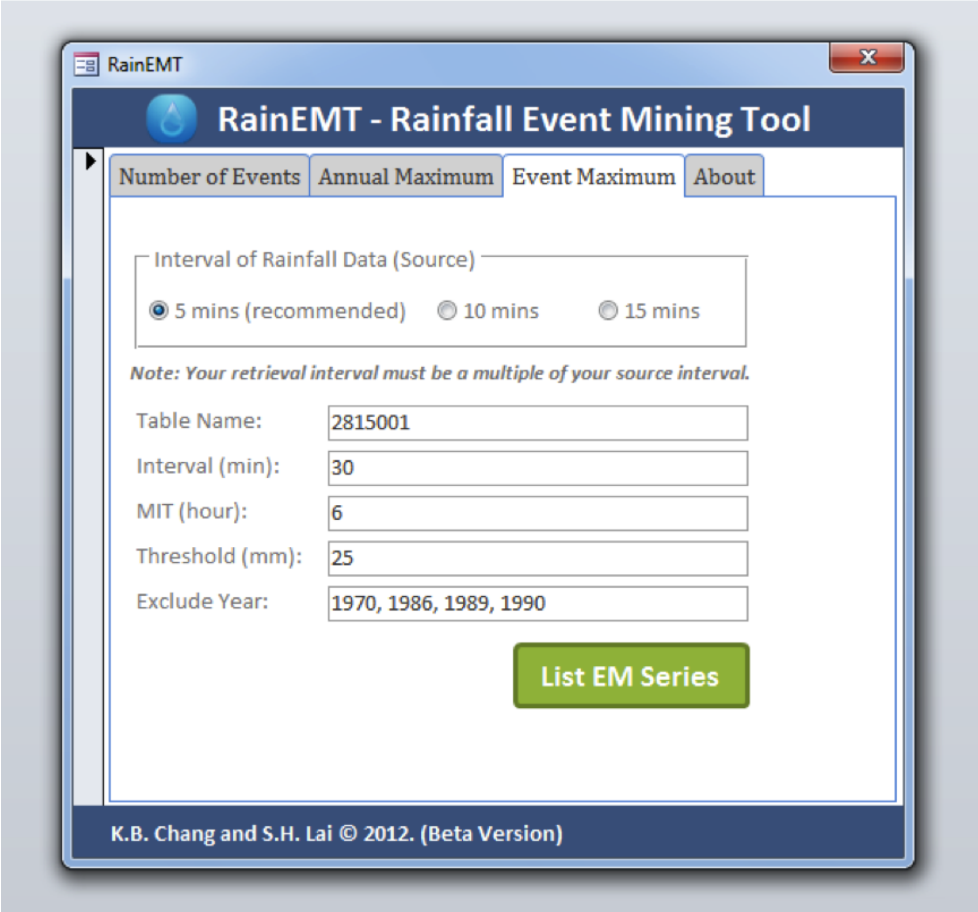


Figure 3.11: User interface for extraction of event maximum rainfall.

RainEMT does not only extract the depth of the event maximum rainfall. It also includes the occurrence date of the event maximum rainfall, storm duration (multiple of the input data's interval, in minutes), total depth (mm) and average intensity (mm/hr) of the rainfall event. This set of rainfall event data can be used in a wide range of analysis and study. Most importantly, such useful data can be extracted easily and therefore, large-scale extraction can be done in a short period of time.

To obtain the partial duration series or peak over threshold data of Station 2815001 with 30 minutes interval, parameters such as MIT of 6 hours, threshold value of 25 mm and years to exclude is entered (Figure 3.11). The table 'EM 30 Threshold-25 MIT-6 (2815001)' contains the output data of the partial duration series, which includes the storm duration, total depth and average intensity of the rainfall event (Figure 3.12). This table can be exported to Microsoft Excel where further threshold requirements such as minimum average intensity or total depth of rainfall event can be applied, if necessary.

EM 30 Threshold-25 MIT-6 (2815001)

ID	Year	Max Rain mr	Date	Duration mi	Depth mm	Avg Intensit
1	1971	30	7/3/1971	215	50.3	14
2	1971	31.7	8/12/1971	365	46	7.6
3	1971	28.1	12/12/1971	60	31.4	31.4
4	1972	26.4	1/9/1972	140	34.9	15
5	1972	25.8	2/10/1972	190	40	12.6
6	1972	37.3	4/11/1972	75	52.4	41.9
7	1972	25.8	4/28/1972	60	30.3	30.3
8	1972	35.7	7/7/1972	80	40.3	30.2
9	1972	34.1	12/29/1972	40	37.4	56.1
10	1973	31.2	1/8/1973	360	48.5	8.1
11	1973	31.6	2/16/1973	80	52.2	39.1
12	1973	30.6	2/25/1973	685	97.5	8.5
13	1973	47.3	2/28/1973	185	51.8	16.8
14	1973	36.6	3/9/1973	70	52.8	45.3
15	1973	27	4/29/1973	200	56.2	16.9
16	1973	49	5/31/1973	460	67	8.7
17	1973	25.9	10/26/1973	80	42.2	31.6
18	1973	25.6	11/26/1973	70	36.4	31.2
19	1973	50.6	11/30/1973	175	63.9	21.9
20	1974	30.5	4/9/1974	105	39.4	22.5
21	1974	33.4	9/9/1974	725	70.8	5.9
22	1974	34.7	9/28/1974	880	49.2	3.4
23	1975	25.3	1/2/1975	185	54.5	17.7
24	1975	27.4	3/9/1975	70	30.9	26.5
25	1975	34	9/25/1975	100	45.9	27.5
26	1975	29.2	10/14/1975	255	33.7	7.9
27	1975	36.6	10/30/1975	155	44.9	17.4

Record: 1 of 279 No Filter Search

Figure 3.12: Partial duration series or peak over threshold output data with a threshold value of 25 mm and MIT of 6 hours extracted from Station 2815001.

3.5.6 Finding the most suitable rainfall minimum inter-event time (MIT)

Two rainfall events are normally separated by a dry period in between. This rainless period has to satisfy a chosen guideline, which is known as minimum inter-event time (MIT) or inter-event time definition (IETD). The objective of this study is to illustrate a few analyses to identify the optimum MIT for separation of rainfall events with two criteria: as many events as possible and as independent as possible. Rainfall events data contain the annual number of rainfall events and together with the number of events that exceed a certain range of durations (0.5, 1, 2, 3, 4, 5, 6, 12 and 24 hours) from 16 rainfall stations in Peninsular Malaysia have been extracted with RainEMT. The relationship between annual numbers of events, event duration categories, and MIT (ranged from 1 to 24 hours) are investigated. These analyses have given consistent results, with effects of geographical and nongeographic parameters observed. The range of unsuitable MIT and the optimum MIT for Peninsular Malaysia has been identified based on these results.

3.5.6.1 Separation of rainfall events

One of the most common and widely used methods for the separation of rainfall events is the use of minimum inter-event time (MIT) (e.g. Dunkerley, 2008; Haile et al., 2010; Haile et al., 2011; Heneker et al., 2001; Powell et al., 2007), which is also known as inter-event time definition (IETD) (e.g. Balistrocchi and Bacchi, 2011; Branham and Behera, 2010; Guo and Adams, 1998; Guo and Baetz, 2007; Palynchuk and Guo, 2008). The dry period between two wet periods are known as inter-event time, and if it is equal to or longer than the desired MIT or IETD, they are considered as two individual events. Apparently, the chosen length of MIT or IETD directly affects the number of rainfall events separated in a fixed period of record.

A difficult task comes with the use of MIT, which is the selection of the proper MIT criterion. The selection of the MIT length is usually associated with the type of intended study or application (runoff modeling, partial duration series, flood studies, canopy drying time, etc.), and directly affects the characteristics of the separated rainfall events (Shamsudin et al., 2010). In this study, the objective is to identify the optimum MIT that includes as many rainfall events and as independent as possible, which is used in extracting rainfall partial duration series. In previous studies, 6-hour MIT appears to be commonly used in Toronto, Canada (e.g. Guo and Adams, 1998; Palynchuk and Guo, 2008). A method is proposed to select the MIT when the analyzed rainfall data has a CV near 1 (Restrepo-Posada and Eagleson, 1982), but is found to be inappropriate as sometimes CV near 1 at any separation time less than 12 hours is observed (Powell et al., 2007). Guo and Adams (1998) mentioned that a more objective way to select the suitable MIT is by examining the relationship between MIT and the average annual number of rainfall events.

The shorter the MIT, the more the events separated, but the independence of these events is reduced at the same time. On the other hand, the longer the MIT, the independence of the separated events increases, but the number of events separated decreases. Therefore, the optimum or best choice of MIT is the intermediate point between the unsuitable MIT (which is too short or too long). In this study, several analyses are carried out on annual number of rainfall events, event duration categories and MIT ranging from 1 to 24 hours for 16 (out of selected 60) rainfall stations in Peninsular Malaysia. Their relationships are analyzed and these analyses are used to identify the unsuitable range of MIT, by taking account of the rainfall characteristics in Peninsular Malaysia. The goal of this study is to identify this unsuitable range of MIT, which leads to the identification of the optimum MIT (to apply on all 60 selected rainfall stations) for use in extraction of rainfall partial duration series in Peninsular Malaysia.

3.5.6.2 Preparation and extraction of data with RainEMT

In order to determine the most suitable MIT to use on selected 60 stations in Peninsular Malaysia, separation of rainfall events with various MIT (i.e. 1 to 24 hours) are first performed on 16 chosen rainfall stations among the 60 selected stations (Figure 3.13). These 16 stations have been categorized into 4 regions based on monthly modification factor regional division for Peninsular Malaysia (DID, 2010), to study the differences between these regions. These regions are separated based on their differences in annual rainfall distribution. The northern region has the highest rainfall distribution from September to October within its own region; central region from October to November; eastern region from November to December and southern region from December to January. Rainfall data with 5 minutes interval (for 16 out of 60 selected rainfall stations) with the numbers of complete years are listed in Table 3.1. The climate of Peninsular Malaysia is very much governed by the monsoons, where the northeast monsoon occurs from May to August, while the southwest monsoon occurs from November to February (Suhaila et al., 2011). Station 8 and 9 are located at Kuala Lumpur, the federal capital of Malaysia.

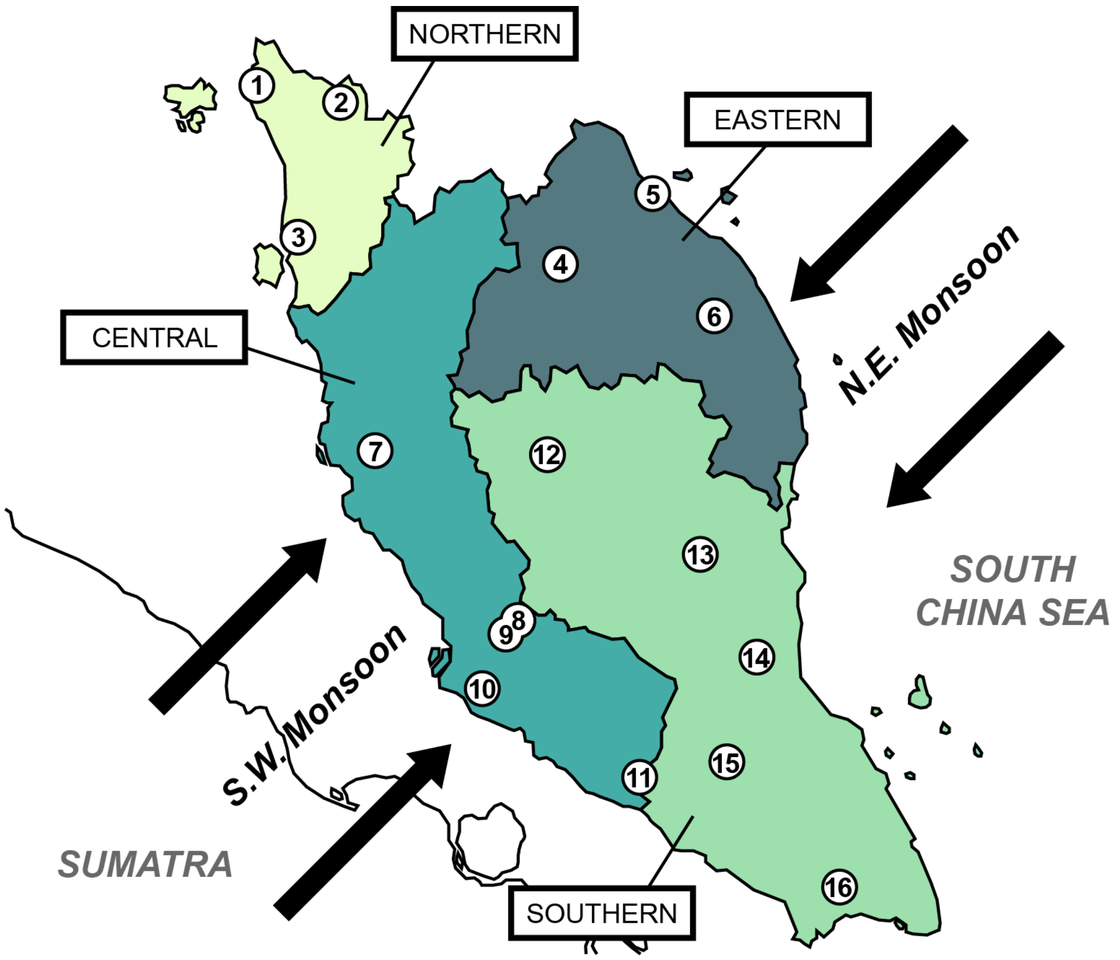


Figure 3.13: Locations of 16 rainfall stations in Peninsular Malaysia.

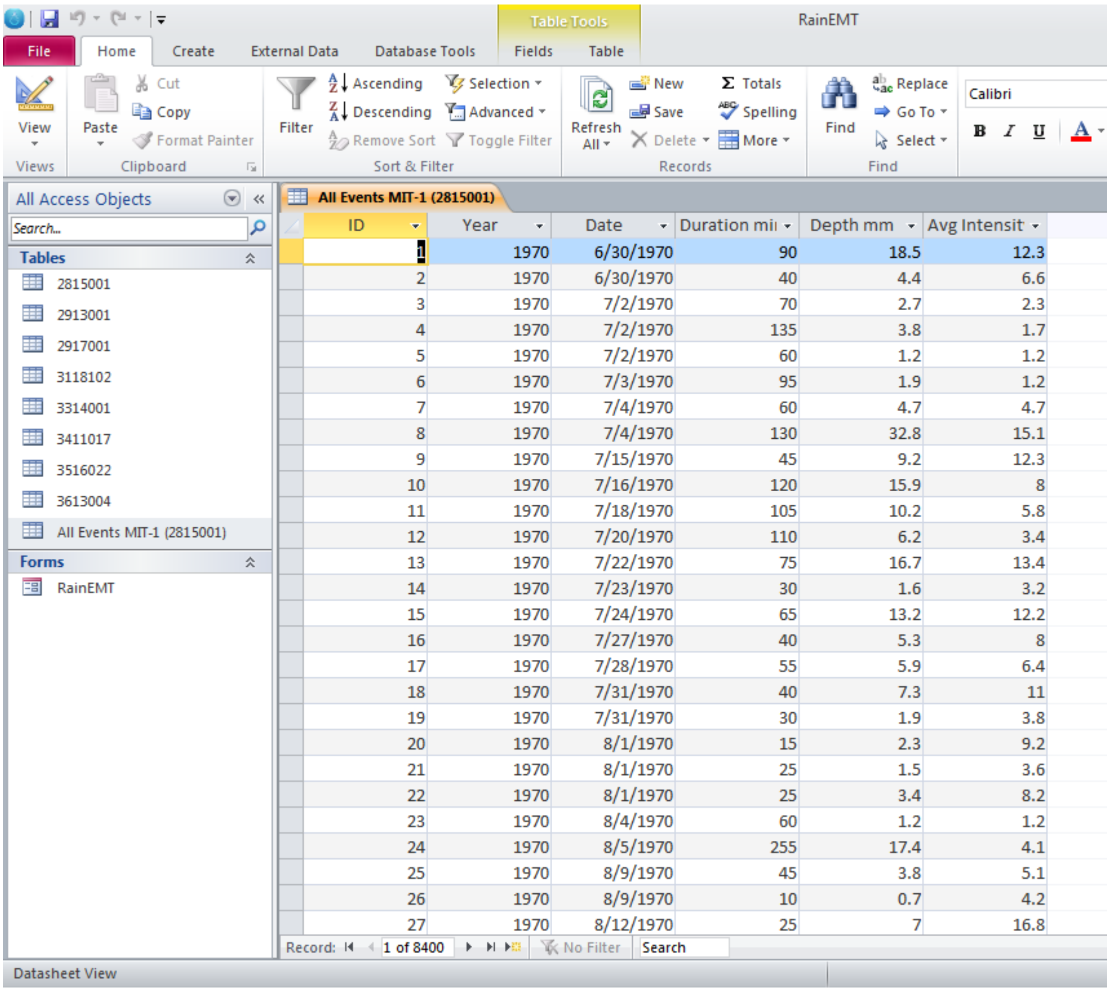
Large-scale extraction of rainfall events with their corresponding storm duration, separated with various MIT (1-24 hour) is a very time consuming task. 5-minute time series rainfall data are first exported from NIWA Tideda (a hydrological database software contains our raw rainfall data), and then they are imported into Microsoft Access for extraction of rainfall events with RainEMT. Large-scale extractions of events data in this study are performed with ease with the use of RainEMT.

RainEMT extracts annual number of rainfall events, with number of events that exceed certain durations (i.e. 0.5, 1, 2, 3, 4, 5, 6, 12 and 24 hours). In each year, to be considered as a rainfall event, a wet period must be surrounded by two dry periods (before and after), which equals to or exceeds the selected MIT. Incomplete years are specified in RainEMT to exclude them from data extraction. The extraction of these rainfall events data (e.g. Figure 3.14) are repeated across MIT ranging from 1 to 24 hours, with 1-hour increment for all stations included in this study.

The extracted data are used to perform three analyses to study the effect of MIT on rainfall events: the relationship of annual number of events under different MIT; distribution of rainfall event with various duration categories of yearly events based on various MIT; and the difference of yearly rainfall events between MIT. The yearly events used here are the average of complete years on the target rainfall station. These analyses performed not only helps to study the characteristics of rainfall events in Peninsular Malaysia, but also used to identify the unsuitable range of MIT.

Table 3.1: Stations information.

Station No.	Region	Observational Period	Years
1	Northern	Nov 1974 – Dec 2011	33
2	Northern	July 1970 – Nov 2011	35
3	Northern	July 1970 – Oct 2011	39
4	Eastern	Sept 1971 – Oct 2011	34
5	Eastern	July 1970 – Nov 2011	27
6	Eastern	July 1971 – Nov 2011	31
7	Central	July 1970 – Nov 2011	37
8	Central	Dec 1972 – Oct 2011	37
9	Central	Dec 1992 – Oct 2011	17
10	Central	June 1970 – Dec 2011	38
11	Central	July 1970 – Nov 2011	41
12	Southern	July 1974 – Nov 2011	29
13	Southern	July 1975 – Nov 2011	29
14	Southern	Aug 1981 – Nov 2011	21
15	Southern	June 1970 – Sept 2011	33
16	Southern	Sept 1980 – Oct 2011	24



ID	Year	Date	Duration mi	Depth mm	Avg Intensit
1	1970	6/30/1970	90	18.5	12.3
2	1970	6/30/1970	40	4.4	6.6
3	1970	7/2/1970	70	2.7	2.3
4	1970	7/2/1970	135	3.8	1.7
5	1970	7/2/1970	60	1.2	1.2
6	1970	7/3/1970	95	1.9	1.2
7	1970	7/4/1970	60	4.7	4.7
8	1970	7/4/1970	130	32.8	15.1
9	1970	7/15/1970	45	9.2	12.3
10	1970	7/16/1970	120	15.9	8
11	1970	7/18/1970	105	10.2	5.8
12	1970	7/20/1970	110	6.2	3.4
13	1970	7/22/1970	75	16.7	13.4
14	1970	7/23/1970	30	1.6	3.2
15	1970	7/24/1970	65	13.2	12.2
16	1970	7/27/1970	40	5.3	8
17	1970	7/28/1970	55	5.9	6.4
18	1970	7/31/1970	40	7.3	11
19	1970	7/31/1970	30	1.9	3.8
20	1970	8/1/1970	15	2.3	9.2
21	1970	8/1/1970	25	1.5	3.6
22	1970	8/1/1970	25	3.4	8.2
23	1970	8/4/1970	60	1.2	1.2
24	1970	8/5/1970	255	17.4	4.1
25	1970	8/9/1970	45	3.8	5.1
26	1970	8/9/1970	10	0.7	4.2
27	1970	8/12/1970	25	7	16.8

Figure 3.14: Rainfall event data (separated with 1-hour MIT) extracted from Station 10 by using RainEMT in Microsoft Access.

3.6 RainIDF

RainIDF, a software tool for derivation of rainfall intensity-duration-frequency (IDF) relationship is developed as an Excel add-in by using Visual Basic for Applications (VBA). The tool is integrated with two of the most widely used statistical distributions for determination of IDF relationship: the generalized extreme value (GEV) distribution for annual maxima series, and the generalized Pareto (GPA) distribution for partial duration series. It provides automated distribution fitting for rainfall data in the form of annual maxima or partial duration series for multiple intervals, solving and plotting of rainfall intensity-duration-frequency curves. RainIDF uses the Solver add-in function in

Excel to solve the coefficients of the empirical IDF formula in one-step. The methodology built into RainIDF is discussed and rainfall IDF relationships for several stations in Peninsular Malaysia are derived and compared. RainIDF is available for download on GitHub (http://github.com/kbchang/rainidf) as an Excel add-in.

3.6.1 Software description

The rainfall intensity-duration-frequency (IDF) relationship is an important tool for the determination of design rainfall in water resources structural design, urban stormwater management, flood modeling, etc. In some cases, depth-duration-frequency (DDF) and intensity-duration-area-frequency (IDAF) relationships are used, which serve the same purpose as IDF relationship. Generally, there are two types of rainfall data series that are widely applied for derivation of IDF relationship: annual maxima series (AMS) and partial duration series (PDS). There are different methods or approaches to derive rainfall IDF relationship (e.g. Ben-Zvi, 2009; De Michele et al., 2011; Koutsoyiannis and Baloutsos, 2000; Madsen et al., 2009; Overeem et al., 2008; Palynchuk and Guo, 2008; Van de Vyver and Demarée, 2010), where the selection of good fitting probabilistic distributions for the target region is very important.

RainIDF is developed by using Visual Basic for Applications (VBA) and can be installed as an Excel add-in. Basically, input data in the form of annual maxima or partial duration series for multiple intervals are inserted into an Excel worksheet and the RainIDF add-in will fit the data with the corresponding probabilistic distribution (generalized extreme value (GEV) or generalized Pareto (GPA) distribution), list out all statistical parameters and optimization procedures to obtain the empirical IDF formula. Besides, it also takes the advantage of Excel's chart plotting functionality to plot the IDF curves automatically based on the derived empirical IDF formula. The output return periods or annual recurrence intervals (ARI) for the IDF relationship derived are 3-month, 6-month, 9-month, 1-year, 2-year, 5-year, 10-year, 20-year, 50-year and 100-year. For other return periods or ARI, it has to be performed manually based on the parameters of the fitted distribution.

3.6.2 Extraction of annual maxima and partial duration series

Before extracting annual maxima or partial duration series (by using the previous software package, RainEMT), the data must be filtered to exclude years with a reasonable amount of missing data. Besides, inappropriate data values due to certain malfunction errors of the recording rain gauge or hydrological database software have to be carefully identified. A way to identify this type of invalid data is by comparing the depth and the duration of the rainfall event. For the rainfall data used in this study, we have identified some problematic years where all the rainfall events have the same duration, which are then excluded from analysis.

The extraction of annual maxima series is fairly simple and straightforward. Annual maximum rainfall for a particular duration or interval is obtained by selecting the largest value of rainfall depth for that particular duration in each year. Besides using annual maximum rainfall for derivation of IDF relationship, some researchers have also studied the trends of the annual maximum rainfall (e.g. Adamowski and Bougadis, 2003; Kuo et al., 2011). Meanwhile, partial duration series (also known as peak-over-threshold or POT approach) consists of all the rainfall events that exceed a certain threshold value. The use of partial duration series is common in flood analysis, until recent studies show an increasing popularity of using partial duration series in rainfall analysis, especially for derivation of IDF relationship (e.g. Beguería, 2005; Ben-Zvi, 2009; Palynchuk and Guo, 2008).

Before extracting partial duration series, it is recommended to determine individual events from rainfall data, as this will increase the independence of the extracted data. To identify individual rainfall event (as mentioned in section 3.5.7.1), a minimum inter-event time can be used. If the dry period between two wet periods is equal or more than the minimum inter-event time, they are considered as two separated events. A minimum inter-event time of 6 hours is commonly used (e.g. Guo and Adams, 1998; Palynchuk and Guo, 2008). A method to select minimum inter-event time based on a coefficient of variation (CV) near 1 is proposed (Restrepo-Posada and Eagleson, 1982), but is found to be inappropriate (Powell et al., 2007). In this segment of study, a minimum inter-event time of 6 hours is adopted for separation of rainfall events as it is found to be the most appropriate MIT for Peninsular Malaysia (see section 4.2.6).

The most uncertain parameter when extracting partial duration series is the threshold value (Beguería, 2005). Similar to minimum inter-event time, threshold value also directly affects the number of rainfall events extracted. Most of the previous researches regarding partial duration series are applied on flood analysis. As for usage in rainfall analysis, a method of choosing threshold values is based on the result of goodness-of-fit test (Ben-Zvi, 2009). Madsen et al. (2002) have implied common threshold values in all the studied stations, which result in the range of 2.5–3.2 for regional average number of exceedances per year. Palynchuk and Guo (2008) have chosen a threshold value of 25mm for their study area in Toronto, Canada. Beguería (2005) has concluded that a unique optimum threshold value cannot be found. In this segment of study, arbitrary thresholds are first attempted, then they are adjusted to produce the desired average number of events per year (around 3 or 2-4 events per year).

3.6.3 Derivation of IDF relationship with RainIDF

By installing RainIDF add-in on Excel 2007 or 2010 (Windows PC), rainfall IDF relationship can be computed automatically based on input data series (annual maxima or partial duration). It is straightforward to derive IDF relationship from annual maxima series, while for partial duration series there are some required parameters such as number of recorded years and threshold value (if the 2-P generalized Pareto distribution is chosen). Figure 3.15 shows the interface of RainIDF in Excel, where the RainIDF menu buttons are located at the top right corner of the home tab. The first step towards generation of IDF relationship is to import annual maxima or partial duration series into Excel spreadsheet, with the header containing the interval value of the data series in minutes. The input data can be in any format (e.g. .txt (text) files and .csv (comma separated values) files) as long as they can be imported (or copied and pasted) into the Excel spreadsheet. By selecting the header range (see Figure 3.15), RainIDF can identify and locate all the data series below the header range (up to 30 sets of data series), and obtain the interval information of the data series from the header. In Figure 3.15, the headers of the partial duration series are selected for generation of IDF relationship, where the 2-P generalized Pareto distribution is selected.

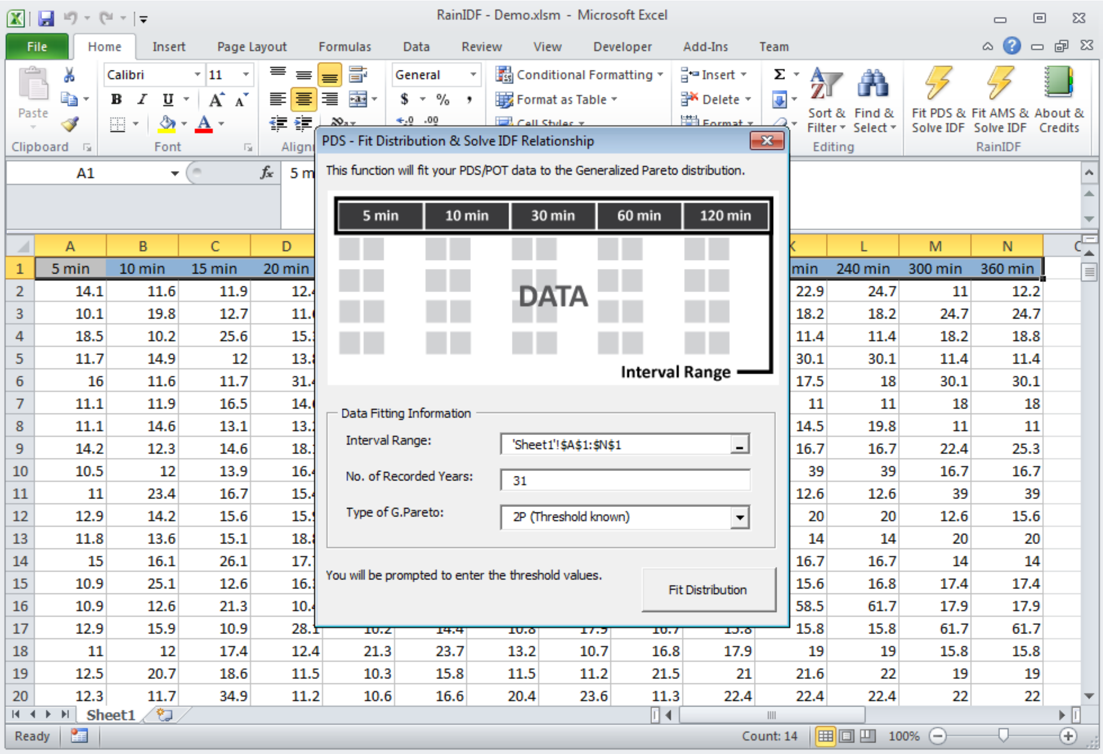


Figure 3.15: Interface of RainIDF add-in with partial duration series input parameters form.

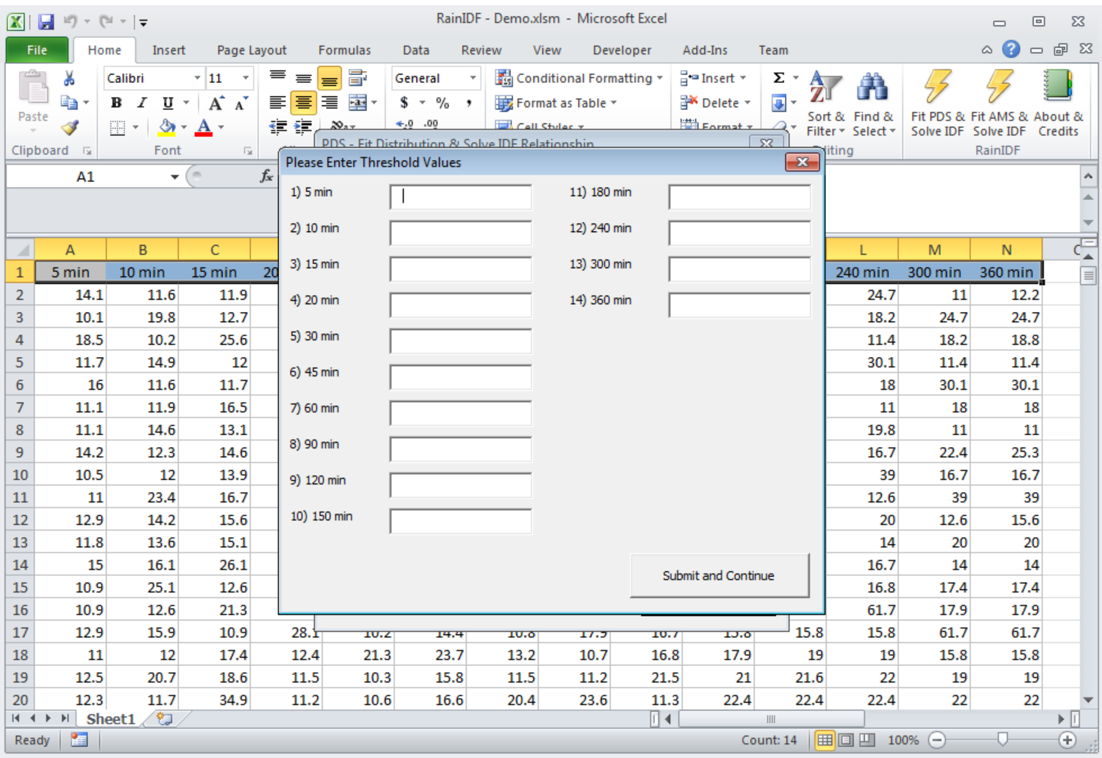
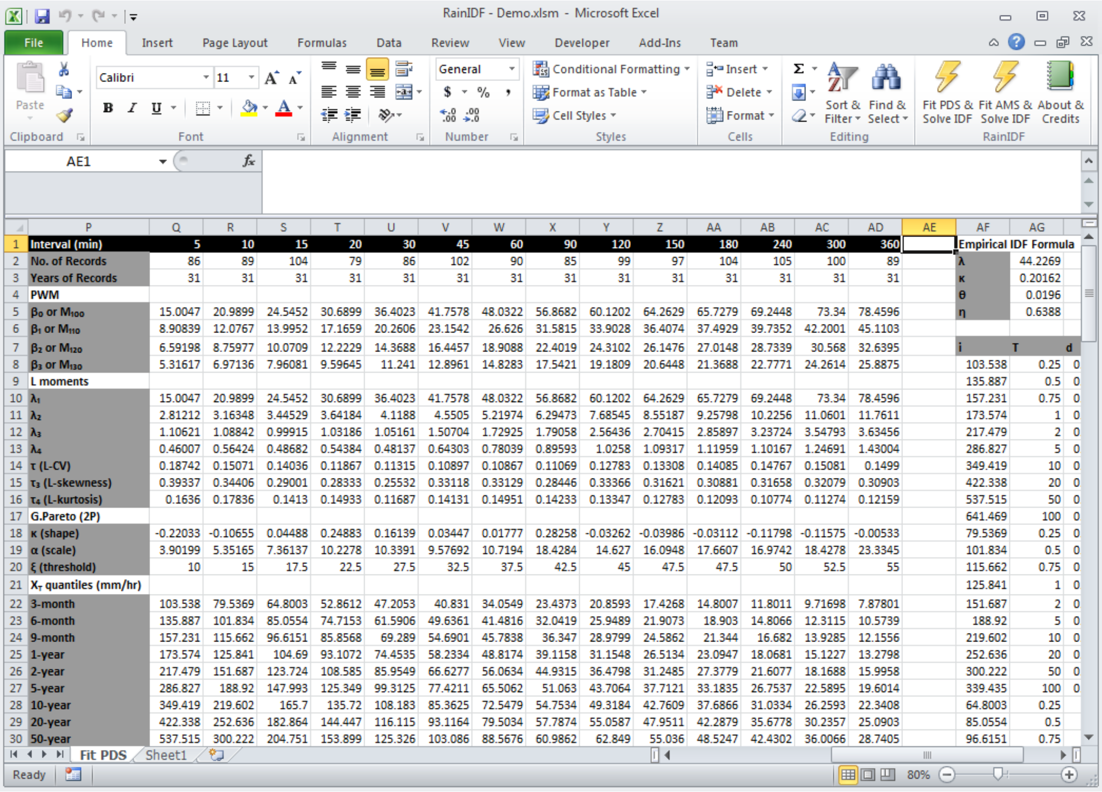


Figure 3.16: Input form for entering threshold values for their corresponding interval.



	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	
1	Interval (min)	5	10	15	20	30	45	60	90	120	150	180	240	300	360		Empirical IDF Formula		
2	No. of Records	86	89	104	79	86	102	90	85	99	97	104	105	100	89		λ	44.2269	
3	Years of Records	31	31	31	31	31	31	31	31	31	31	31	31	31	31		κ	0.20162	
4	PWM																θ	0.0196	
5	β_0 or M_{100}	15.0047	20.9899	24.5452	30.6899	36.4023	41.7578	48.0322	56.8682	60.1202	64.2629	65.7279	69.2448	73.34	78.4596		η	0.6388	
6	β_1 or M_{110}	8.90839	12.0767	13.9952	17.1659	20.2606	23.1542	26.626	31.5815	33.9028	36.4074	37.4929	39.7352	42.2001	45.1103				
7	β_2 or M_{120}	6.59198	8.75977	10.0709	12.2229	14.3688	16.4457	18.9088	22.4019	24.3102	26.1476	27.0148	28.7339	30.568	32.6395		i	T	d
8	β_3 or M_{130}	5.31617	6.97136	7.96081	9.59645	11.241	12.8961	14.8283	17.5421	19.1809	20.6448	21.3688	22.7771	24.2614	25.8875		103.538	0.25	0
9	L moments																135.887	0.5	0
10	λ_1	15.0047	20.9899	24.5452	30.6899	36.4023	41.7578	48.0322	56.8682	60.1202	64.2629	65.7279	69.2448	73.34	78.4596		157.231	0.75	0
11	λ_2	2.81212	3.16348	3.44529	3.64184	4.1188	4.5505	5.21974	6.29473	7.68545	8.55187	9.25798	10.2256	11.0601	11.7611		173.574	1	0
12	λ_3	1.10621	1.08842	0.99915	1.03186	1.05161	1.50704	1.72925	1.79058	2.56436	2.70415	2.85897	3.23724	3.54793	3.63456		217.479	2	0
13	λ_4	0.46007	0.56424	0.48682	0.54384	0.48137	0.64303	0.78039	0.89593	1.0258	1.09317	1.11959	1.10167	1.24691	1.43004		286.827	5	0
14	τ (L-CV)	0.18742	0.15071	0.14036	0.11867	0.11315	0.10897	0.10867	0.11069	0.12783	0.13308	0.14085	0.14767	0.15081	0.1499		349.419	10	0
15	τ_3 (L-skewness)	0.39337	0.34406	0.29001	0.28333	0.25532	0.33118	0.33129	0.28446	0.33366	0.31621	0.30881	0.31658	0.32079	0.30903		422.338	20	0
16	τ_4 (L-kurtosis)	0.1636	0.17836	0.1413	0.14933	0.11687	0.14131	0.14951	0.14233	0.13347	0.12783	0.12093	0.10774	0.11274	0.12159		537.515	50	0
17	G.Pareto (2P)																641.469	100	0
18	κ (shape)	-0.22033	-0.10655	0.04488	0.24883	0.16139	0.03447	0.01777	0.28258	-0.03262	-0.03986	-0.03112	-0.11798	-0.11575	-0.00533		79.5369	0.25	0
19	α (scale)	3.90199	5.35165	7.36137	10.2278	10.3391	9.57692	10.7194	18.4284	14.627	16.0948	17.6607	16.9742	18.4278	23.3345		101.834	0.5	0
20	ξ (threshold)	10	15	17.5	22.5	27.5	32.5	37.5	42.5	45	47.5	47.5	50	52.5	55		115.662	0.75	0
21	X_T quantiles (mm/hr)																125.841	1	0
22	3-month	103.538	79.5369	64.8003	52.8612	47.2053	40.831	34.0549	23.4373	20.8593	17.4268	14.8007	11.8011	9.71698	7.87801		151.687	2	0
23	6-month	135.887	101.834	85.0554	74.7153	61.5906	49.6361	41.4816	32.0419	25.9489	21.9073	18.903	14.8066	12.3115	10.5739		188.92	5	0
24	9-month	157.231	115.662	96.6151	85.8568	69.289	54.6901	45.7838	36.347	28.9799	24.5862	21.344	16.682	13.9285	12.1556		219.602	10	0
25	1-year	173.574	125.841	104.69	93.1072	74.4535	58.2334	48.8174	39.1158	31.1548	26.5134	23.0947	18.0681	15.1227	13.2798		252.636	20	0
26	2-year	217.479	151.687	123.724	108.585	85.9549	66.6277	56.0634	44.9315	36.4798	31.2485	27.3779	21.6077	18.1688	15.9958		300.222	50	0
27	5-year	286.827	188.92	147.993	125.349	99.3125	77.4211	65.5062	51.063	43.7064	37.7121	33.1835	26.7537	22.5895	19.6014		339.435	100	0
28	10-year	349.419	219.602	165.7	135.72	108.183	85.3625	72.5479	54.7534	49.3184	42.7609	37.6866	31.0334	26.2593	22.3408		64.8003	0.25	
29	20-year	422.338	252.636	182.864	144.447	116.115	93.1164	79.5034	57.7874	55.0587	47.9511	42.2879	35.6778	30.2357	25.0903		85.0554	0.5	
30	50-year	537.515	300.222	204.751	153.899	125.326	103.086	88.5676	60.9862	62.849	55.036	48.5247	42.4302	36.0066	28.7405		96.6151	0.75	

Figure 3.17: Spreadsheet contains parameters of the fitted data series and IDF relationship.

Since the 2-P GPA distribution is chosen, user will be prompted to enter the corresponding threshold value based on the selected range of headers (Figure 3.16). If the 3-P GPA distribution is selected, this step is skipped, as the threshold or location parameters will be estimated from the data series by using the L-moments method. User may choose to use a set of common threshold value of all rainfall stations (which produces varied average number of events per year), or arbitrary threshold values that requires adjustment for different stations to produce a desired average number of events per year (e.g. 2-4 events). (After the threshold values are entered, RainIDF will automatically filter data value that is lower (if there is any) than its corresponding threshold value. RainIDF creates a new spreadsheet (Figure 3.17) containing all the important parameters (such as PWM, L-moments, distribution parameters and quantiles), coefficients of empirical IDF formula and IDF curves (Figure 3.18).

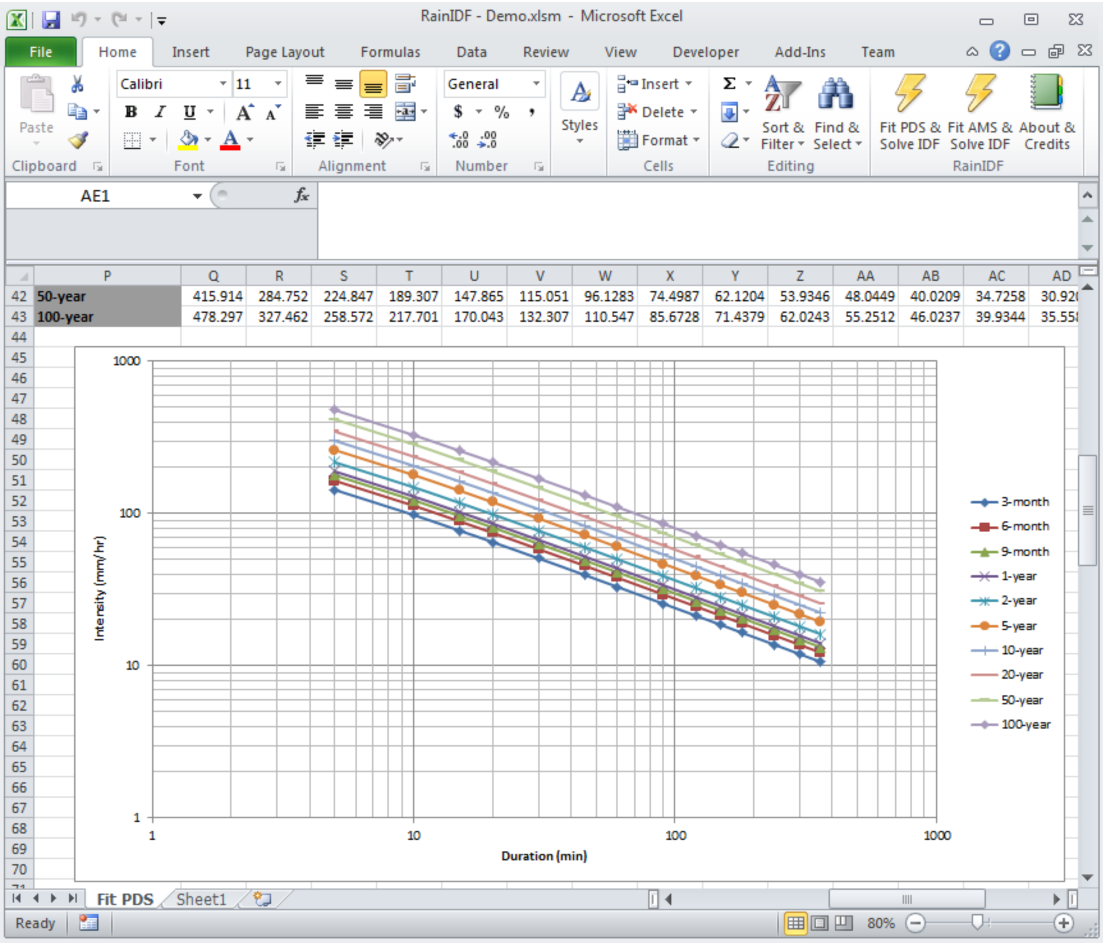


Figure 3.18: Rainfall IDF curves plotted automatically with RainIDF Excel add-in.

RainIDF calls the Solver utility function in Excel, to perform one-step least squares method for solving and optimizing the coefficients of the empirical IDF formula. Other methods and details about one-step least squares method are discussed in Koutsoyiannis et al. (1998). The coefficients of the solved empirical IDF formula are listed in the generated spreadsheet (see top right corner of Figure 3.17). The optimized quantiles from the empirical IDF relationship are calculated and plotted into IDF curves (Figure 3.18). It is worth noting that the Solver add-in embedded in Excel must be enabled in order to derive IDF relationship with RainIDF, else a warning message box will appear as RainIDF fails to call the Solver add-in function.

3.7 RainMap

The purpose and function of RainMap is to showcase all the design rainfall of every single stations of this study in one place. RainMap is coded in Visual Basic and integrated with Bing Maps powered by Microsoft. The dynamic mapview allows users to zoom and locate the rainfall stations accurately, especially when finding the nearest located rainfall stations to their site. When a rainfall station is selected in RainMap, it will display the empirical IDF formula and GPS coordinates of the station, and with an option to view the design rainfall of the station. RainMap is a creative and effective visual tool for displaying design rainfall with the location of the rainfall stations that will ease and change the way in which future water scientists and engineers store and view their design rainfall data.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter will be divided into different sections with their corresponding results and discussions that will contribute to the overall conclusions of this study.

4.2 Rainfall characteristics and minimum inter-event time (MIT) in Malaysia

This section of study will show the results and discussions on identifying the most suitable MIT for Peninsular Malaysia, along with some rainfall characteristics of the regions. The extraction and preparation of data from 16 rainfall stations are discussed in section 3.5.7.2.

4.2.1 Relationship of annual number of rainfall events with MIT

The average annual numbers of rainfall events extracted are plotted with their corresponding MIT (Figure 4.1). These include all yearly rainfall events and rainfall events that exceed a certain duration such as 0.5, 1, 2, 3, 4, 5, 6, 12 and 24 hours. In general, the number of rainfall events decrease with the increase in MIT. By observing the trend of 'All Events' in Figure 4.1A, there is a significant drop on number of yearly events for MIT shorter than 6 hours. After MIT of 5 or 6 hours, the drop seems to be gradual and gives a smooth line (Figure 4.1A). The trend lines of 'All Events' for all observed rainfall stations are quite similar (Figure 4.1A-E). For some stations (such as Figure 4.1B-E), the dropping rate of 'All Events' seems to change slightly at long MIT (e.g. more than 20 hours).

By comparing ‘All Events’ with ‘Events > 30 mins’ in Figure 4.1, the high number of rainfall events and the significant drop of rainfall events at short MIT (e.g. less than 6 hours) are mostly consist of short duration events (30 minutes and below). Rainfall events that exceed 1, 2, 3, 4, 5 and 6 hours are low at short MIT, and increase to a certain point with the increment of MIT, followed by stable trends (small decrease or increase) and get closer to each other (Figure 4.1). These trends show that MIT affects all short and long durations of rainfall events, and this is the reason where the second analysis (a study of rainfall distribution with event duration categories based on various MIT) is performed.

The number of rainfall events exceed 12 hours generally increases with the increment of MIT. At most stations (e.g. Figure 4.1B-E), the number of events that exceeds 12 hours increases at a low rate at short MIT, continued by stable increases, and followed with a low rate of increases at long MIT. Comparisons of this observation and the previous discussed trend of ‘All Events’ indicate that short and long MIT are causing dramatic rates of change on the annual number of events observed, where there seem to have a range with a stable and more consistent rate in between.

An outstanding trend of rainfall events that exceed 24 hours has been observed. Throughout all stations studied, the number of rainfall events that exceed 24 hours is near zero or very small at short MIT. Upon reaching a certain point, the numbers take off and increase significantly with the increment of MIT (especially after 12 hours). In this case, the high number of rainfall events above 24 hours does not represent the characteristics of rainfall events in Peninsular Malaysia well, where most of the rainfall occurred in Peninsular Malaysia is convective.

4.2.2 Distribution of rainfall duration categories under different MIT

The distribution of annual rainfall events is divided into various event duration categories in percentage (Figure 4.2A). The event duration categories starts from 5 minutes as 5 minutes rainfall data are used. They are plotted with MIT ranging from 1 to 24 hours (Figure 4.2) for each station. In this analysis, instead of investigation the number of events associated with different MIT, it studies the formation of rainfall event distribution by the percentage of different event duration categories at different MIT. Again, all the stations show a very similar pattern of rainfall distribution across the range of MIT tested.

By looking at Figure 4.2, rainfall events with duration of 5 to 30 minutes seem to accumulate around 50% of the total rainfall at MIT of 1 hour. This shows that the separation of rainfall events that is less independent than each other, where some of these short duration events should be considered as a single combined event. The percentage of 5 to 30 minutes rainfall events drops significantly between MIT of 1 to 5/6 hours. After that, the percentage of 5 to 30 minutes seems to have a stable and low decreasing rate with the increment of MIT.

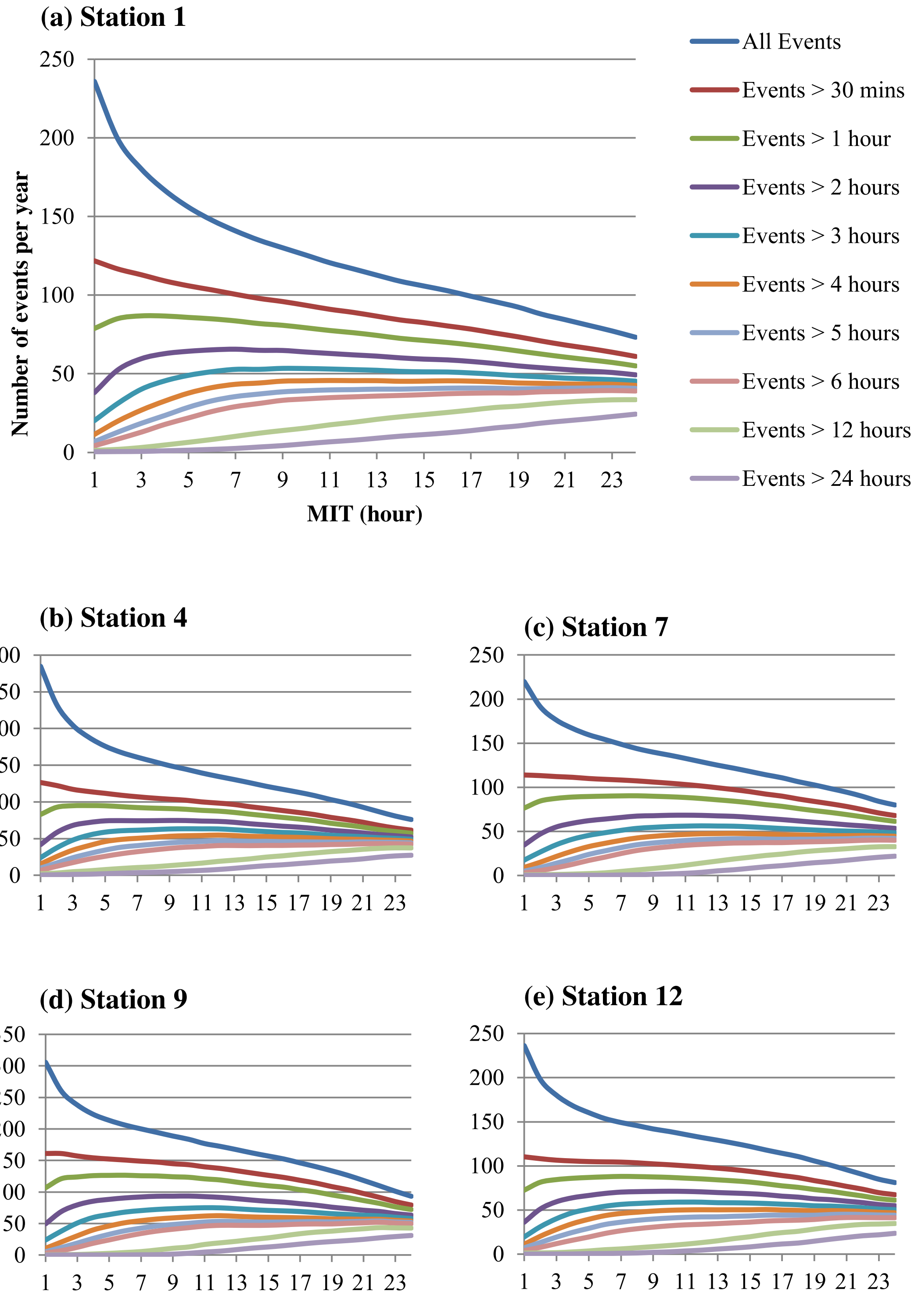


Figure 4.1: Number of rainfall events versus MIT (hour) at different stations.

Figure 4.2 shows that the percentage of rainfall events with duration of 35 to 60, 65 to 120, 125 to 180, 185 to 240, 245 to 300, 305 to 360 minutes seem to be quite similar with no significant changes through the range of 1 to 24 hours MIT. Only slight increases or decreases are observed at short and long MIT. However, the percentage of rainfall events from 365 to 720 minutes occupies a very small portion of the total rainfall events separated at MIT of 1 hour. This percentage increases from MIT of 1 hour until 10 hours, where the increment in percentage is lowered and changed to decrement with further increment of MIT. These observations show that a short MIT (e.g. 1 to 3 hours) is not recommended for application in extracting partial duration series as it tends to separate long duration events (365 to 720 minutes) into short duration events (which produces too many short duration events, especially from 5 to 30 minutes), where the independences of these events are questionable.

The percentage of rainfall events with duration of 725 to 1440 minutes also seems to be very low at short MIT (e.g. 1 to 3 hours). With the increment of MIT, they increase at a high rate until it slows down with little or no changes at long MIT. Note that percentage of rainfall events (from 725 to 1440 minutes in terms of duration) at MIT above 12 hours seems to be very high (around 10% to 18%), when compared with MIT shorter than 12 hours.

When we look at the percentage of rainfall events with duration above 24 hours (or 1440 minutes) of Figure 4.2, it remains at near zero or very small percentage for MIT below 10 hours. When the length of MIT increases, this percentage increases significantly. These increases are undesired as more short duration events are combined into long duration events, which decrease the number of rainfall events. Note that the percentage of rainfall events above 24 hours is more than 20% at long MIT (e.g. 22 hours and above).

4.2.3 Difference in number of rainfall events between different MIT

In order to have a closer look at the rate of change of total events with increment of MIT at Figure 4.1, the 1-hour differences for number of events between the corresponding MIT are plotted for comparison (Figure 4.3). By looking at Figure 4.3, the differences of number of events are very high and with significant drops (from MIT of 1 to 2 until 4 to 5 hours). With the increment of MIT, the differences reach a consistent and stable rate. However, at long MIT (e.g. above 20 hours), the differences seem to fluctuate, but remain lower than the differences observed at short MIT (see Figure 4.3C and D).

A good explanation for the pattern observed at short MIT (e.g. 4 hours and below) is a large number of rainfall events separated at this range of MIT is more dependent of each other, when compared with MIT that is longer than 4 hours. By combining with the observations of previous analyses, most of them appear to be short duration events (30 minutes and below). The rainfall events separated becomes less dependent with the increment of MIT, which then gives a similar difference between MIT. Theoretically, the number of events separated decreases with the increment of MIT (e.g. Dunkerley, 2008; Shamsudin et al., 2010), and this decrease should be quite consistent if these events are independent, which can be seen in Figure 4.3, particularly after the MIT of 6 hours.

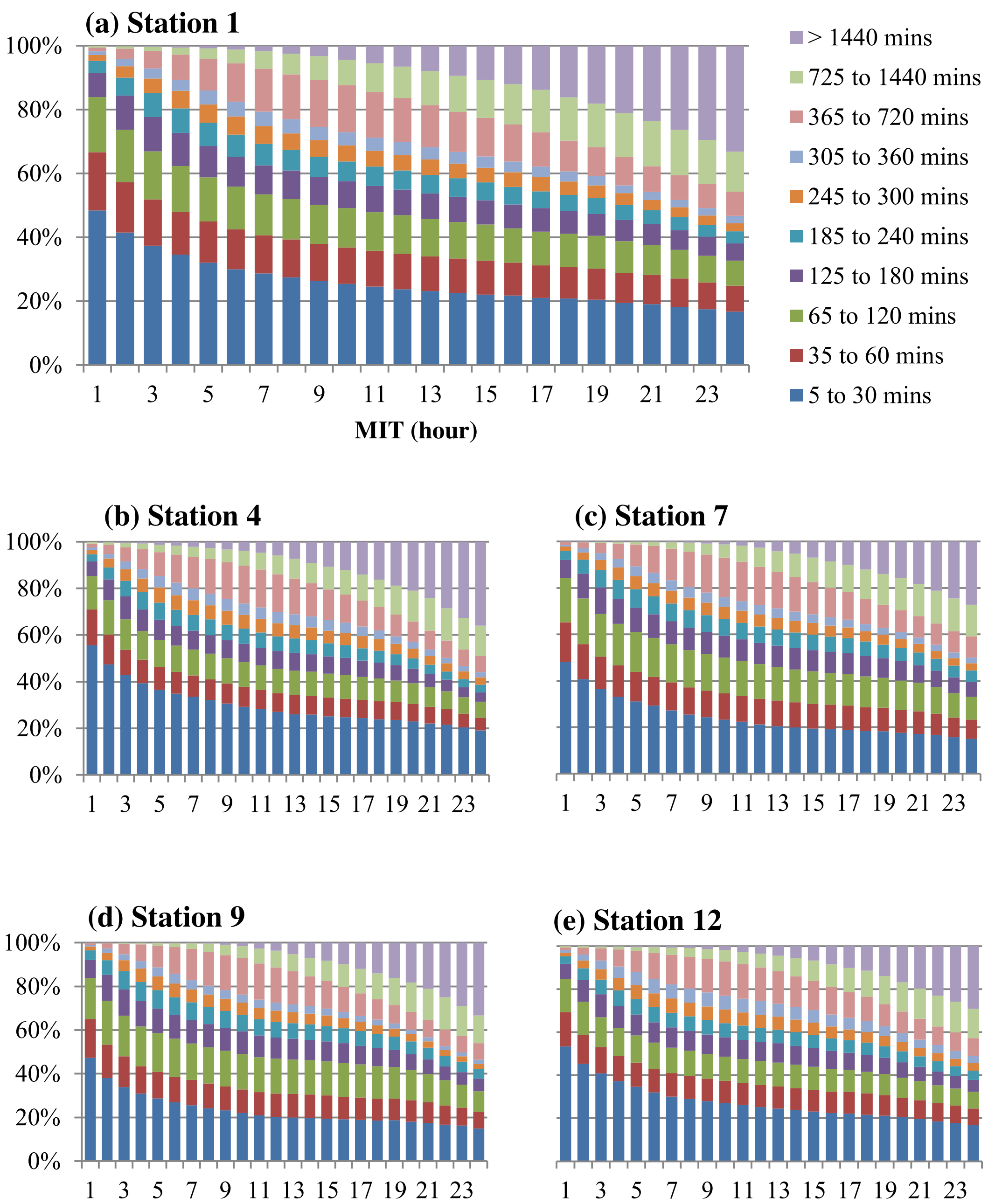


Figure 4.2: Percentage of rainfall events at various duration versus MIT (hour) of different stations.

The fluctuations of the differences as mentioned earlier at long MIT could mean that, at long MIT, some events that are supposed to be independent have been combined as a single event. This could cause some extreme events being excluded and neglected, when they are merged with extreme events where the depths of these events are larger.

For extreme event statistics, this could mean a great loss of important data or information. Hence, it is important to choose the proper range of MIT that includes as many events as possible (in this case, the shorter the MIT, the more the number of events), but at the same time making sure that these events are independent (or less dependent) of each other (especially for MIT that is too short). This shows that there is a need to identify the undesired or unstable range of short MIT, and to use the shortest possible and stable MIT for the separation of rainfall events.

It is very obvious that the short MIT (e.g. less than 4 hours) are undesired or unstable by observing their differences in Figure 4.3. However, it is very hard to tell when is the changing point or the beginning of the stable MIT. A quantitative measure or guideline is needed for this purpose. When the difference of events for a MIT with its following 1-hour MIT exceeds this guideline, that MIT will be categorized as unstable. At first, we have attempted to see if a general threshold value for all the stations in this study can be determined. However, it appears to be inappropriate as the differences between MIT of some stations are quite different from other stations, and we believe that it is due to the influence of geographical parameters such as wind regime.

Therefore, it is found that every station requires its own threshold value, which is computed based on the rainfall characteristics and results of the station itself. A suitable and reasonable way to compute this value is by obtaining the average or mean difference of yearly events between MIT. Observations show that the differences of events after MIT of 6 hours maintained at a low and quite consistent level for all stations in Figure 4.3, by averaging the value of differences including the differences of those short MIT which is significantly higher (especially for MIT below 4 hours), the resulted average or mean difference will be a threshold value that is higher than most of the stable MIT but lower than those MIT that give abnormally high and unstable values.

The mean difference for all stations are calculated and listed in Table 4.1, along with the list of MIT that exceed the mean difference. The results give a consistent rejection for MIT of 4 hours and below, where 5-hour appear to be inappropriate for 3 stations (station 1, 6 and 15). While most of the stations do not have MIT of 5 hours that exceed the mean difference, the differences for 5-hour MIT for these stations are very close to the values of mean differences. Therefore, 5-hour seems to be the changing point of the unstable to stable MIT, which makes it not suitable to be listed as stable MIT. The finalized unstable range of MIT for Peninsular Malaysia is 5 hours and below.

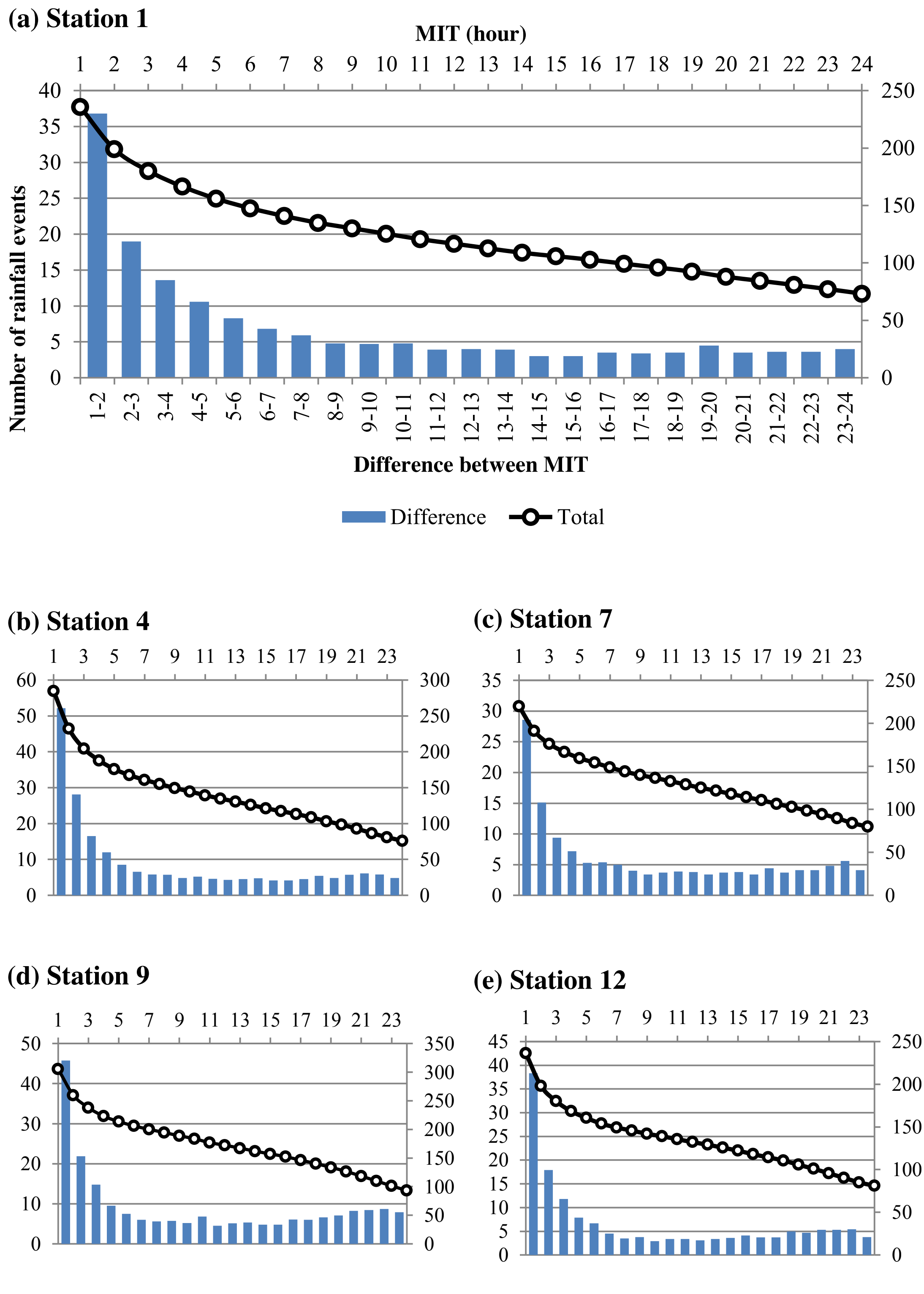


Figure 4.3: Total rainfall events across MIT of 1 to 24 hours and difference in number of total rainfall events between MIT.

Table 4.1: Stations and list of MIT that exceed the mean difference.

Station No.	Region	Mean Difference	Exceeded MIT (hour)
1	Northern	7.07	1, 2, 3, 4, 5
2	Northern	6.80	1, 2, 3, 4
3	Northern	6.64	1, 2, 3, 4
4	Eastern	9.07	1, 2, 3, 4
5	Eastern	7.82	1, 2, 3, 4
6	Eastern	11.9	1, 2, 3, 4, 5
7	Central	6.08	1, 2, 3, 4
8	Central	8.43	1, 2, 3, 4
9	Central	9.23	1, 2, 3, 4
10	Central	5.25	1, 2, 3, 4
11	Central	5.04	1, 2, 3, 4
12	Southern	6.75	1, 2, 3, 4
13	Southern	6.07	1, 2, 3, 4
14	Southern	8.89	1, 2, 3, 4, 5
15	Southern	5.72	1, 2, 3, 4
16	Southern	6.04	1, 2, 3, 4, 21, 22

The fluctuation of differences at long MIT as mentioned previously could also surpass the mean differences, such as the exceeded MIT (21 and 22 hours) of station 16 listed in Table 4.1. The proper way to evaluate the unsuitability of long MIT for separation of rainfall would be the analyses in the previous sections, where the method used to produce the result in Table 4.1 should only be applied for identification of unstable range of short MIT.

4.2.4 Regional and individual stations comparisons

The average number of yearly rainfall events separated by 1, 3, 6, 12 and 24 hours for each region are plotted for comparison (Figure 4.4). The eastern region has an outstanding number of rainfall events separated by MIT of 1 hour compared to other regions. A good reason for this is that eastern region is most influenced by the southeast monsoon. Although the central region is affected by the northeast monsoon, it appears that present of Sumatra might have reduced the impact of the monsoon on this region (see Figure 4.4). The northern region appears to have lowest number of rainfall events throughout the range of MIT compared in Figure 4.4, as it has the less influence of both monsoons due to its geographic location.

With the increment of MIT, it seems that the number of events for eastern region reduced significantly, and slightly surpassed by the central region later. Figure 4.3B shows that the percentage of 5 to 30 minutes events for station 4 (within the eastern region) is higher than all the other stations (for northern, central and southern regions) in Figure 4.2. Thus, it seems that the southeast monsoon has caused the increased number of short duration rainfall for the eastern region, which can be observed when a short MIT is applied.

The mean differences of number of yearly events between MIT are also compared for the four regions (Figure 4.5). Again, eastern region has an outstanding high differences compare to the other regions, which have similar values with each other. This significant difference of eastern region is due to its large number of rainfall events separated by short MIT that is shown in Figure 4.4. This observation agrees that a general guideline in term of a specific value, is not suitable to be used to analyze the unstable range of short MIT.

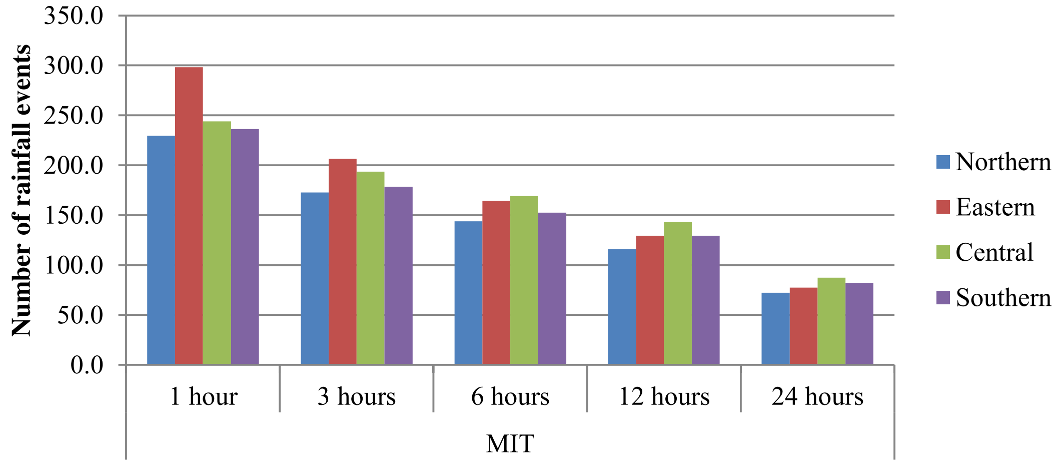


Figure 4.4: Annual number of events for northern, eastern, central and southern regions separated under 1, 3, 6, 12 and 24 hours of MIT.

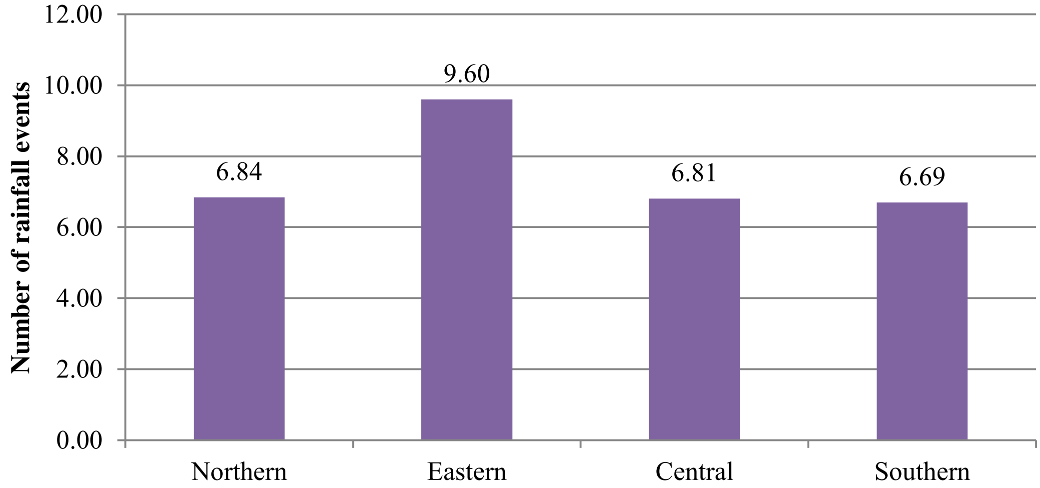


Figure 4.5: Mean differences for annual number of events between MIT for northern, eastern, central and southern regions.

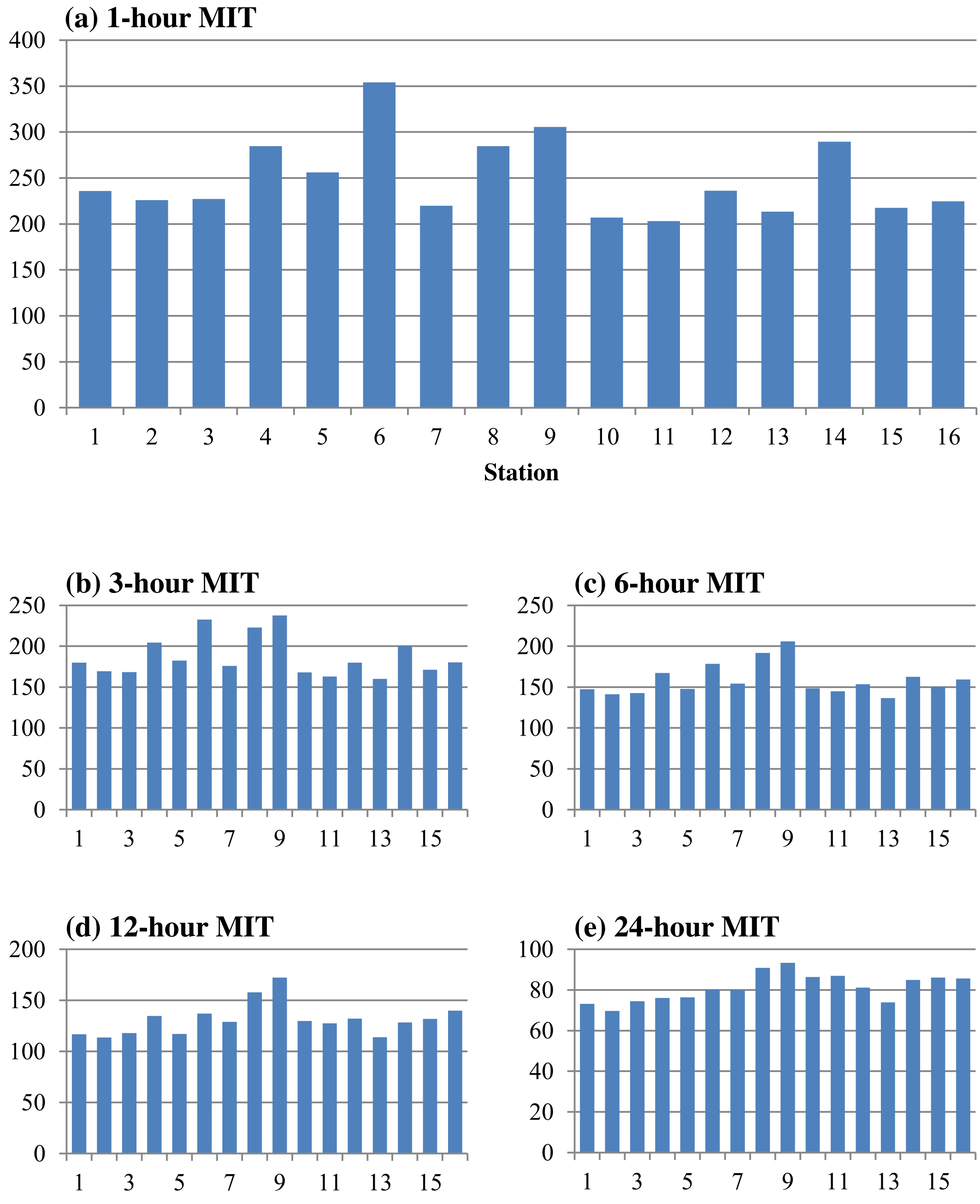


Figure 4.6: Annual number of events separated by different MIT for 16 stations.

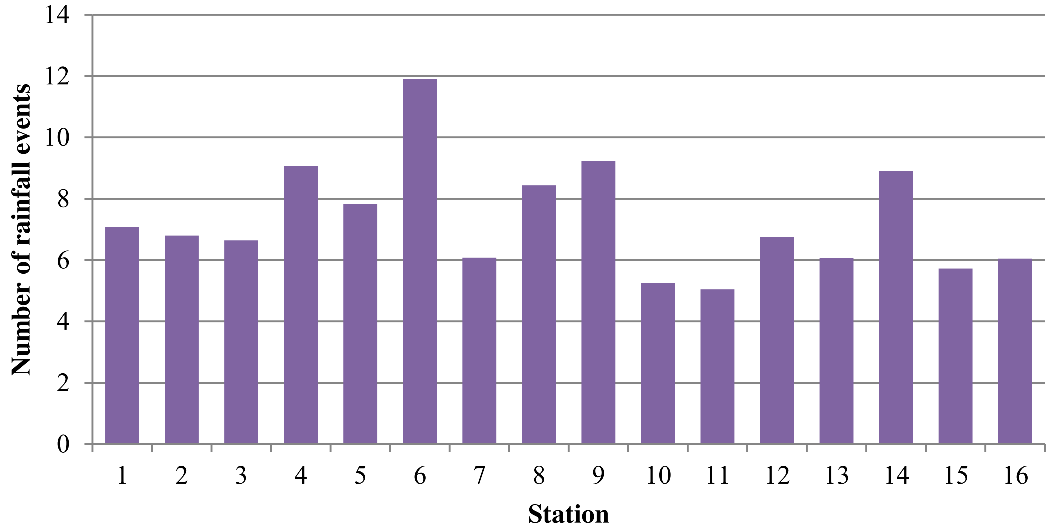


Figure 4.7: Mean differences in annual number of events between MIT for 16 stations.

After regional comparison, all 16 stations in this study are also compared to see the differences for these stations (Figure 4.6 and Figure 4.7) for more detailed observation. Based on Figure 4.6A, the number of events separated by MIT of 1 hour for station 4, 5, 6, 8, 9 and 14 has crossed the 250 line. Station 4, 5 and 6 are the stations within the eastern region, which shows obvious higher events than other regions. Although station 14 is within the southern region, it is nearest east coast in its region, and could be influenced by the northeast monsoon. Meanwhile, station 8 and 9 are located in the central region, but show a really outstanding measure compared to other stations within the central region (station 7, 10 and 11).

Located at the federal state and the most developed area of Malaysia, station 8 and 9 are experiencing rapid urban growths, which are believed to cause the rise of rainfall frequency. This supports the recent findings, which show that areas experiencing fast urban growth have received increases in heavy seasonal rainfall (Kishtawal et al., 2010). Urbanization is known to impact flood occurrences and cause increases in annual average rainfall (e.g. Franczyk and Chang, 2009; Hollis, 1975). This also shows that besides from the influence of monsoons in this study, urbanization or land-use also plays an important role in the characteristics of rainfall events of the study area. If station 8 and 9 were to be excluded from the central region, it will become the region with the lowest number of rainfall at 1-hour MIT (as station 7, 10 and 11 are among the stations with the least number of events).

The six stations (4, 5, 6, 8, 9 and 14) with highest number of rainfall events at 1-hour MIT, tends to reduce their difference with other station with the increment of MIT (Figure 4.6B-E); while station 8 and 9 remain to be the two stations with highest number of rainfall events. When combined with the observations in Figure 4.2 and Figure 4.4, they lead to two conclusions: the influence of monsoons are mainly on short duration events when separated by short MIT; while the impact of urbanization (station 8 and 9) gives a general increase in all rainfall events throughout the range of MIT applied in this study.

The mean differences of annual number of events between MIT are also compared (Figure 4.7) and appear to be very similar with Figure 4.6A, as the main contributors of the differences between these stations are the differences on annual number of events at short MIT, where the increment of MIT will lower these differences. This again supports the earlier finding regarding the inappropriateness of using a general value as threshold level for all stations to identify the unstable range of short MIT.

4.2.5 Conclusions

The analyses performed in this study show similar and consistent trends with the selected range of MIT (1 – 24 hours) for 16 stations in Peninsular Malaysia. It appears that a MIT that is as short as possible is desired as the longer the MIT; the more the consecutive rainfall events considered belonging to the same event (which results in the increase of long duration events, such as more than 24 hours). However, at short MIT, there are concerns on the independence of the separated rainfall events. By analyzing the differences of annual number of rainfall events, the range of the undesired rainfall events associated with their corresponding MIT can be identified.

The results from this analysis show that MIT of 5 hours and below is inappropriate to be used in Peninsular Malaysia, as for the application of rainfall extracting partial duration series, the condition of the rainfall events required to be independent and to include as many events as possible. The most suitable MIT for the climate condition in Peninsular Malaysia will be 6 hours (right after the unsuitable range of short MIT). For studies where specific MIT is selected, these analyses will help to identify the characteristics and conditions of the separated rainfall events.

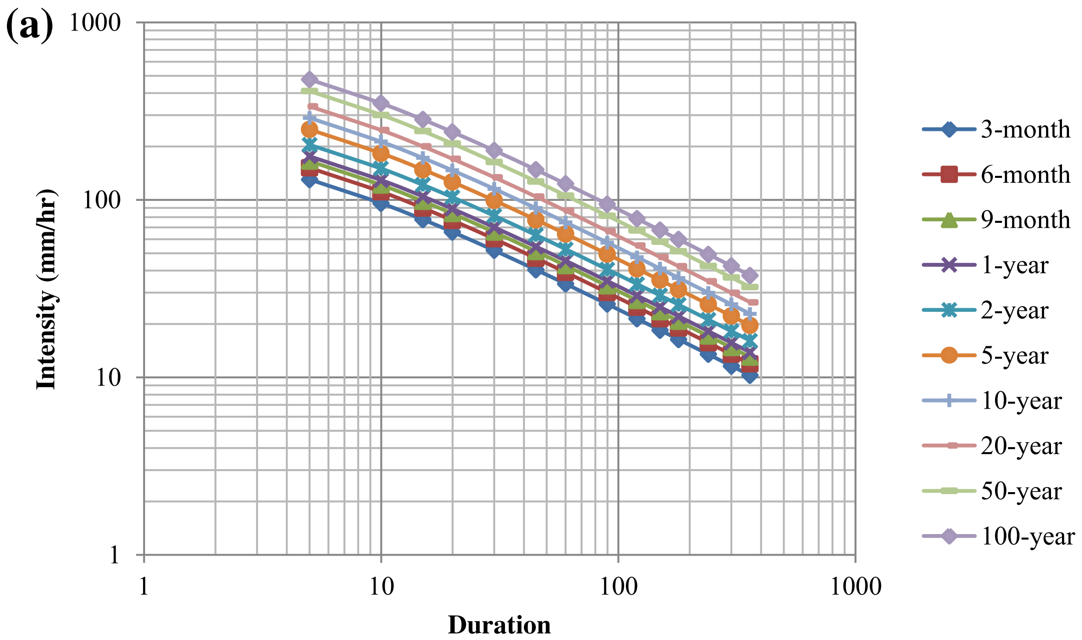
This study also discovered that the numbers of rainfall events are affected by monsoons and land-use (or urbanization in this case). Thus, the analyses conducted in this study can also be used to identify the variation of rainfall events and the influences of geographical parameters on the study area. It is found that the northeast monsoon has caused increases in number of rainfall events at east coast of Peninsular Malaysia and is observed with the application of short MIT. The impacts of urbanization and intensive land development in Kuala Lumpur (Station 8 and 9) could have caused the general increases of annual number of events.

4.3 Determination of best fitting distribution and data series

This segment of study compares and identifies the best approach (GEV/AMS or GPA/PDS) for derivation of IDF relationship in Peninsular Malaysia. By using MIT of 6 hours (result from section 4.2), rainfall PDS data can be extracted by using RainEMT. Extraction of rainfall AMS data is straight forward, and also done by using RainEMT. These two data series for all 60 selected rainfall stations are collected and used to derive their IDF relationship by using RainIDF. In order to compare the goodness of fit of GEV/AMS and GPA/PDS approach, RainIDF is used to identify their coefficient of L-moments and IDF relationship at the same time. Therefore, the necessity of software like RainIDF for this study can be seen here. Without a tool like RainIDF, this step might not seen to be reasonable to perform as the required time and effort is too high.

4.3.1 Derivation and comparison of GEV/AMS and GPA/PDS IDF curves

Using both GEV/AMS and GPA/PDS approaches, IDF relationship of all 60 selected rainfall stations has been derived. 3 chosen stations and their IDF relationship are compared (Figure 4.8, 4.9 and 4.10). The data series include durations of 5, 10, 15, 20, 30, 45, 60, 90, 120, 150, 180, 240, 300 and 360 minutes. By comparing IDF curves computed from annual maxima series with IDF curves of partial duration series, it seems that the rainfall intensity for short return period (e.g. 5 years and below) for partial duration series are slightly higher than annual maxima series throughout all the plotted durations. Note that the partial duration series used are around 3 events per year on average, compared to annual maxima series with 1 event per year. It is easier to derive IDF relationship with annual maxima series and GEV distribution as the extraction of annual maxima series is straightforward while extraction of partial duration series requires extra steps (e.g. separation of rainfall events and selecting threshold values). Although the choice of minimum inter-event time and threshold values can cause slight differences for partial duration series results, the most important step in preparation of data is to filter the invalid or problematic data (especially for short duration rainfall such as 5 and 10 minutes).



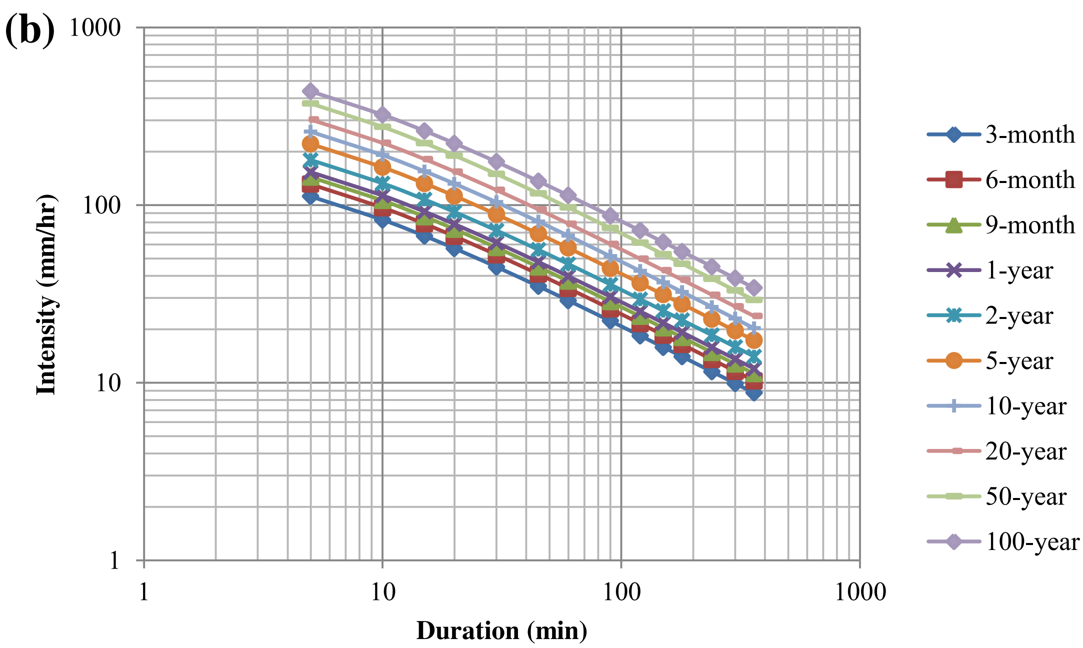
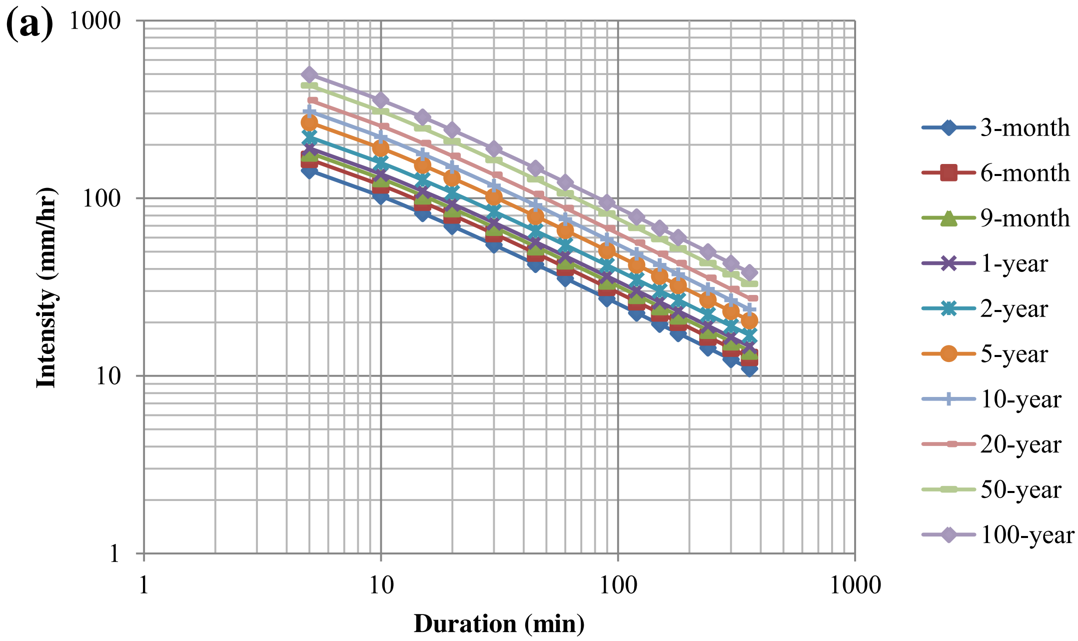


Figure 4.8: IDF curves derived from station 2330009 in Johor. (a) Partial duration series. (b) Annual maxima series.



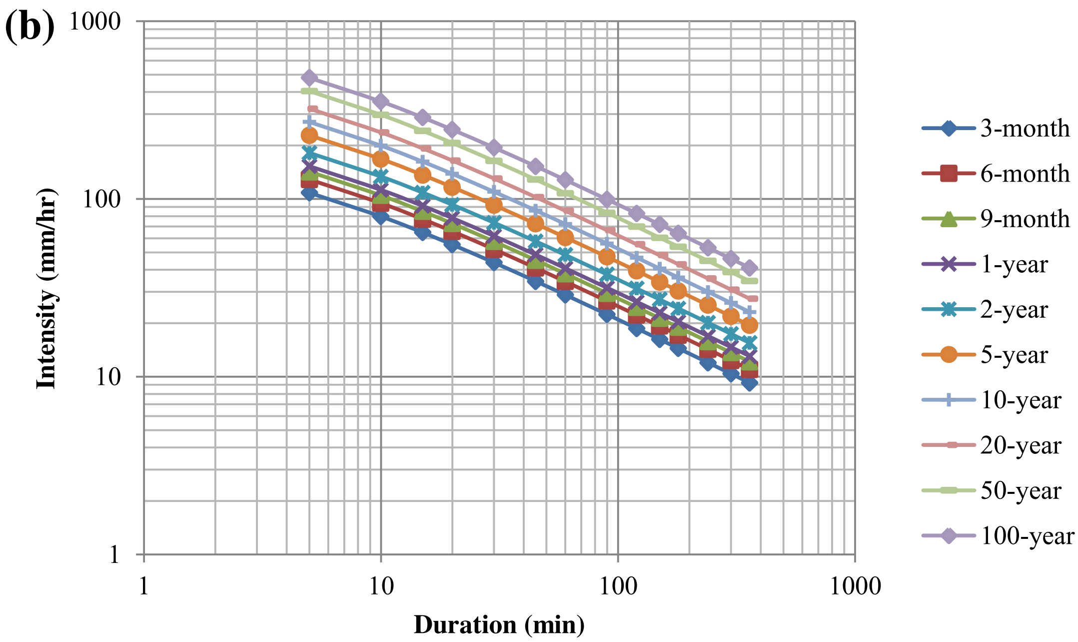
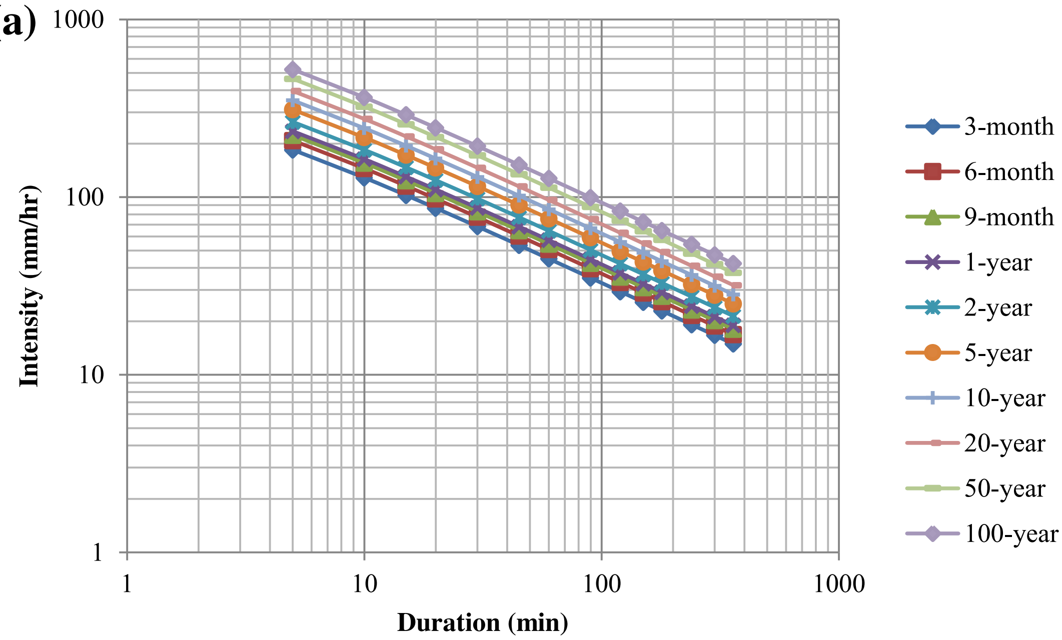


Figure 4.9: IDF curves derived from station 3628001 in Pahang. (a) Partial duration series. (b) Annual maxima series.

(a)



(b)

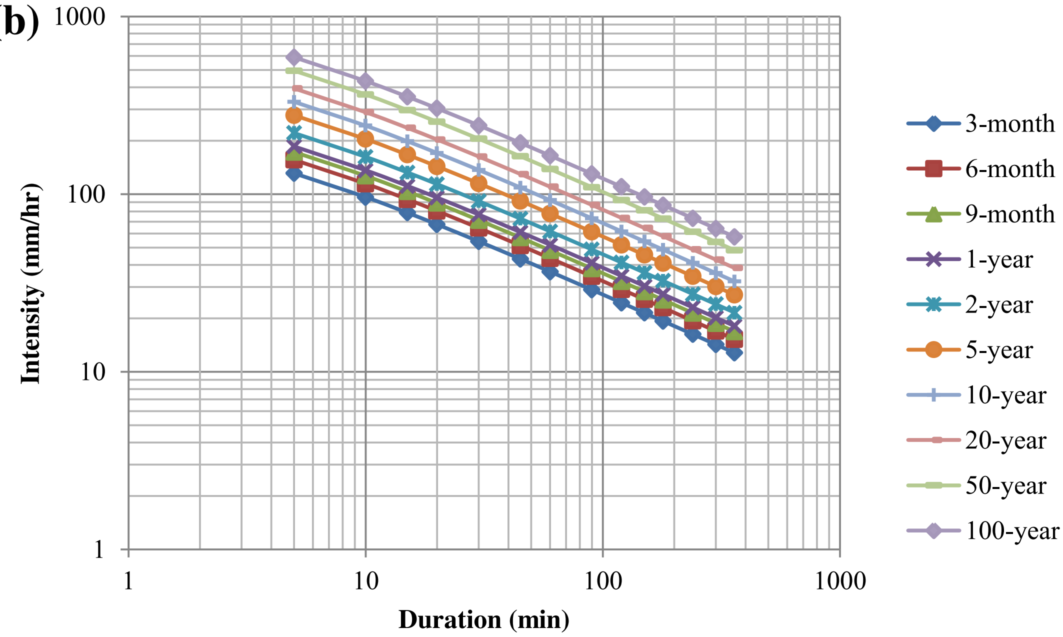


Figure 4.10: IDF curves derived from station 6019004 in Kelantan. (a) Partial duration series. (b) Annual maxima series.

4.3.2 L-moment ratio diagram of GEV/AMS and GPA/PDS

To choose between GEV/AMS and GPA/PDS approaches, *L*-moment ratio diagram as demonstrated by Hosking and Wallis (1997) is used to determine the good-fit of the AMS or PDS to their corresponding distributions. *L*-moment ratio diagrams for the 3 selected stations are plotted in Figure 4.11. By comparing the fitting of GEV/AMS (Figure 4.11A, C and E) with the fitting of GPA/PDS (Figure 4.11B, D and F), it is observed that GPA/PDS has a better fitting than GEV/AMS, as the GPA/PDS *L*-moments ratios and sample mean are closer to the population of *L*-skewness and *L*-kurtosis of the GPA distribution. In this case, the use of GPA/PDS approach is encouraged for the derivation of IDF relationship for these 3 stations. By plotting the mean *L*-moment ratio of all 60 selected rainfall stations, we can see that the fitting of PDS to GPA is better than the fitting of AMS to GEV (Figure 4.12). Since the *L*-moment ratio diagram provides a clear indication for the fitting of the data, additional goodness-of-fit tests (such as Anderson-Darling Test) are not needed in this study. The result shows that GPA/PDS approach is the most suitable approach for Peninsular Malaysia.

4.3.3 Conclusions

This study shows that GPA/PDS approach for derivation of IDF relationship is more desired than GEV/AMS approach for Peninsular Malaysia. The importance of rainfall IDF curves as design rainfall references in water resources engineering are increasing especially with the impact of climate change. RainIDF can help to speed up analysis work and thus, large-scale derivation of rainfall IDF relationship can be performed with ease (especially for partial duration series). If other IDF formula is preferred, one may use the distribution parameters and quantiles obtained via RainIDF to apply with the preferred IDF formula. Rainfall data with missing data and invalid values have to be filtered carefully before they are used to extract annual maxima or partial duration series, as these problematic data will affect the accuracy of the derived IDF curves.

On the other hand, if the desired return period is different from the supported return period, one may also calculate the quantiles of the desired return period manually based on the statistical parameters. Since RainIDF fits GEV distribution to annual maxima series and GPA distribution to partial duration series, one should study their suitability for the selected regions as their condition of fitting to data series might varied on different regions. Goodness-of-fit tests and *L*-moment ratio diagrams are among the widely used method for testing the goodness-of-fit of distributions.

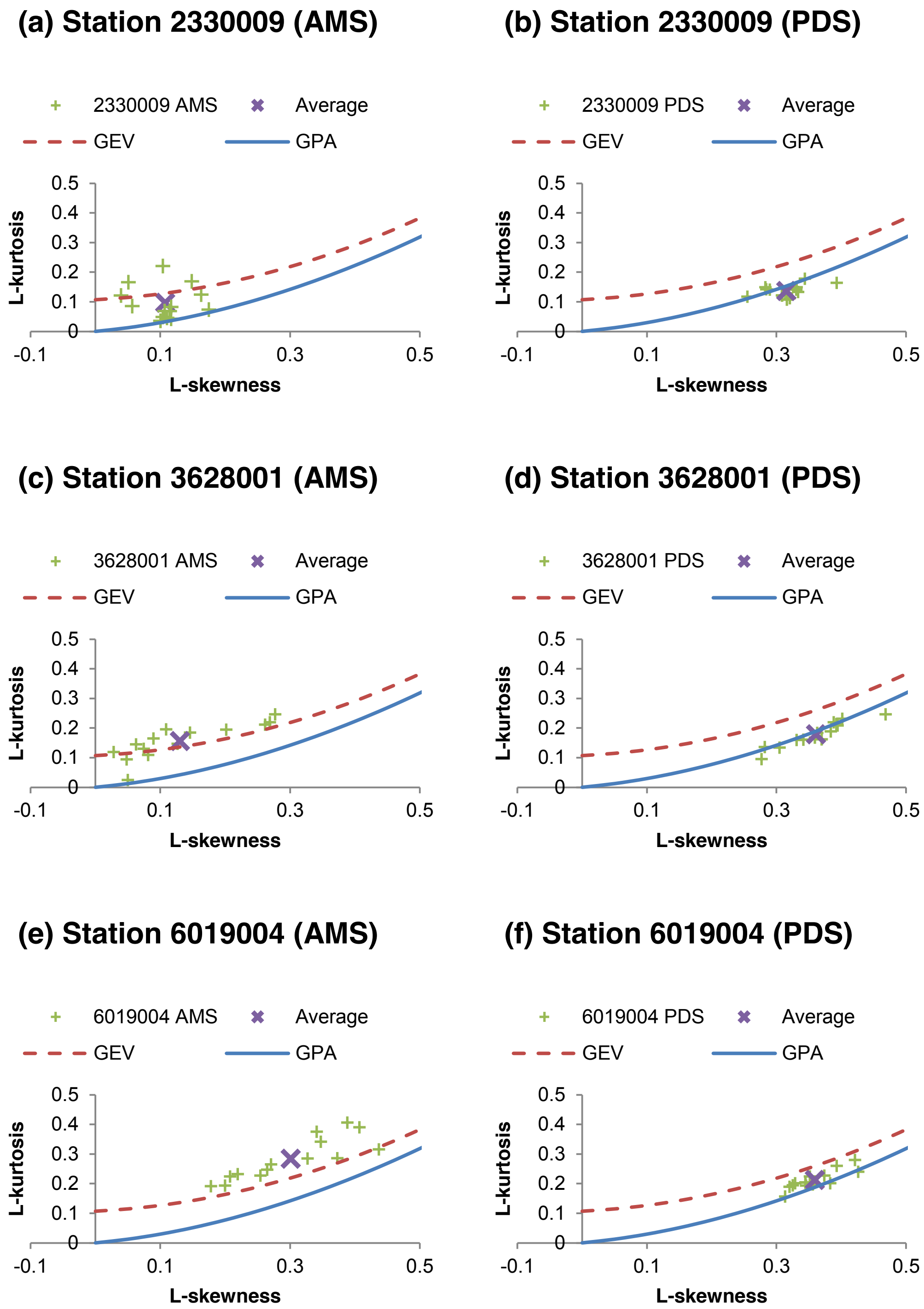


Figure 4.11: *L*-moment ratio diagrams for annual maxima series (AMS) and partial duration series (PDS) data obtained from 3 selected rainfall stations in Peninsular Malaysia.

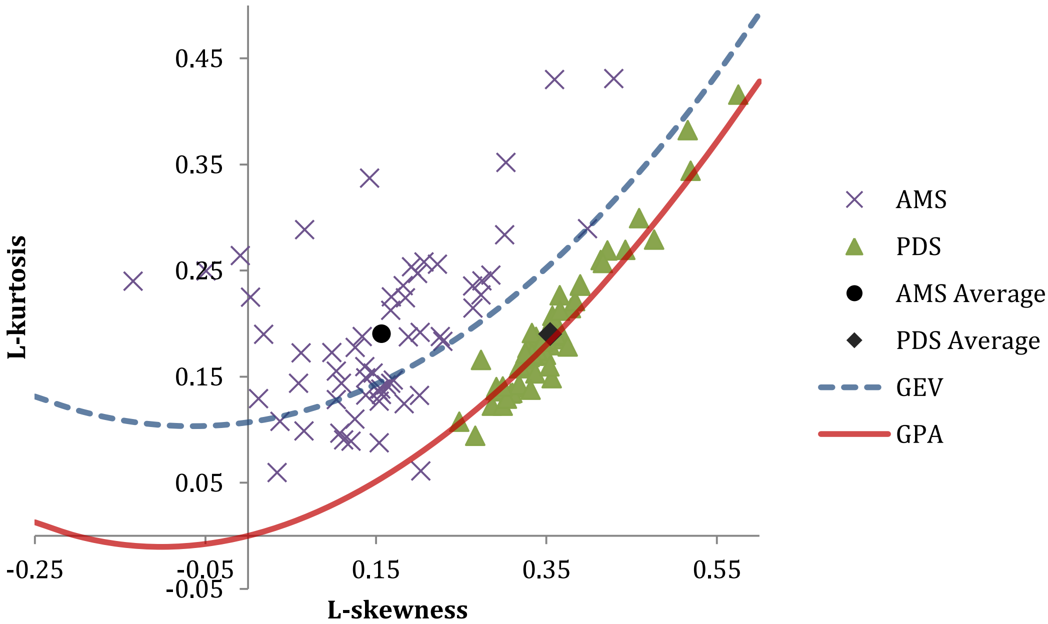


Figure 4.12. *L*-moment ratio diagram for annual maxima series (AMS) and partial duration series (PDS) data obtained from 60 rainfall stations in Peninsular Malaysia.

4.4 Presentation of design rainfall data for end-users with RainMap

All IDF relationship coefficients derived from RainIDF are then integrated into RainMap. As currently there is no convenient way to view location of rainfall station and their corresponding design rainfall dynamically, RainMap provides a huge leap in convenience of viewing design rainfall, location the nearest rainfall station for target project location. The main screen of RainIDF with all 60 rainfall stations can be seen on Figure 4.13. User can click on the pushpin on the map (powered by Bing Map – Microsoft) or select from the list of rainfall stations on the left. Moreover, users can zoom-in the dynamic Bing Map to locate the target rainfall station accurately (Figure 4.14). The information of the rainfall station such as coefficient of empirical IDF formula, station number and state are shown on the right side of the screen (Figure 4.13 and Figure 4.14). By clicking on the 'generate design rainfall' button, the design rainfall of the target rainfall stations will be displayed in a pop up window (Figure 4.15).

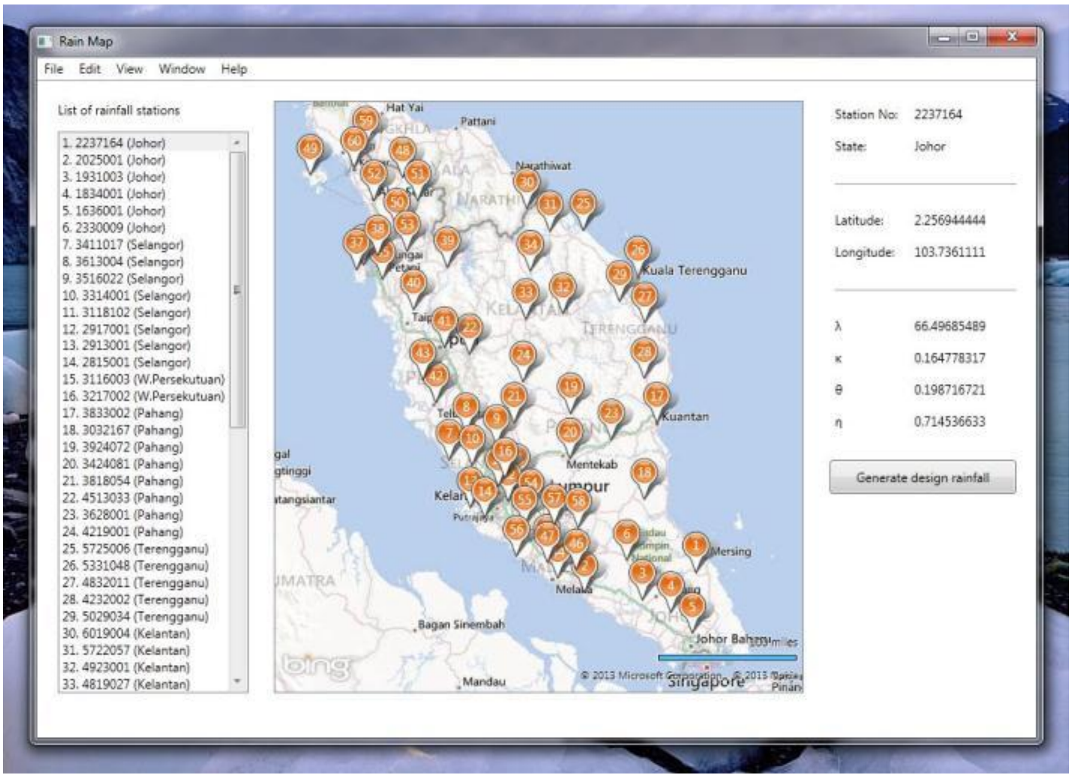


Figure 4.13: RainMap screenshot shows data and push pins of 60 rainfall stations.

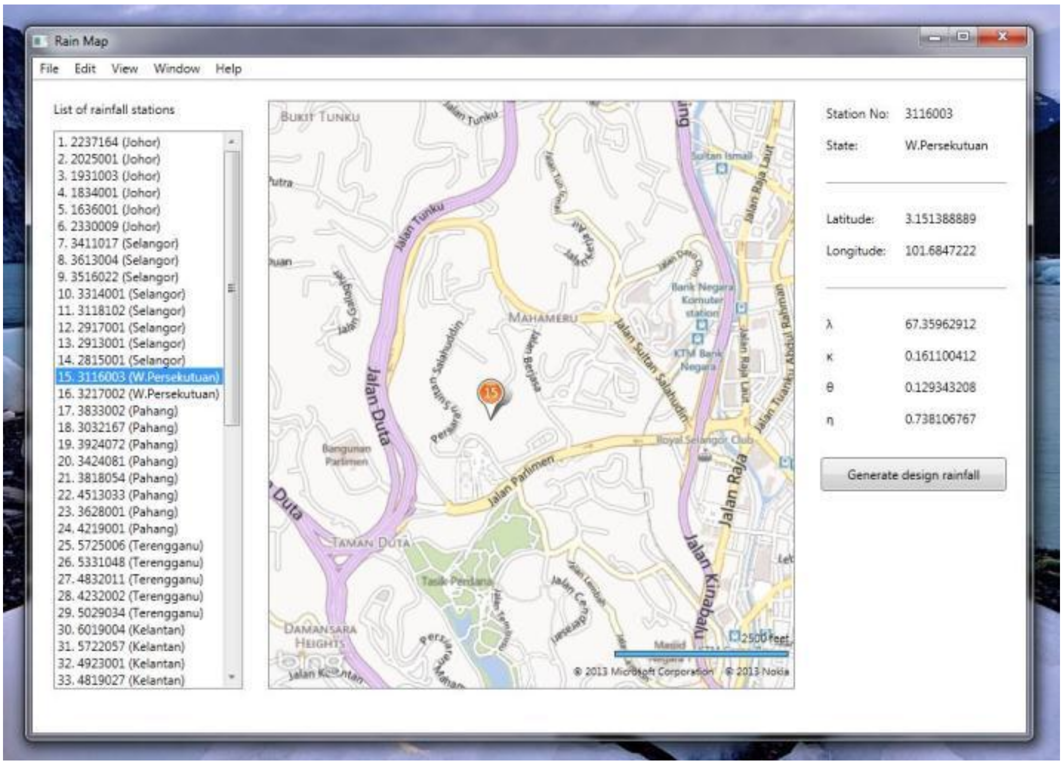
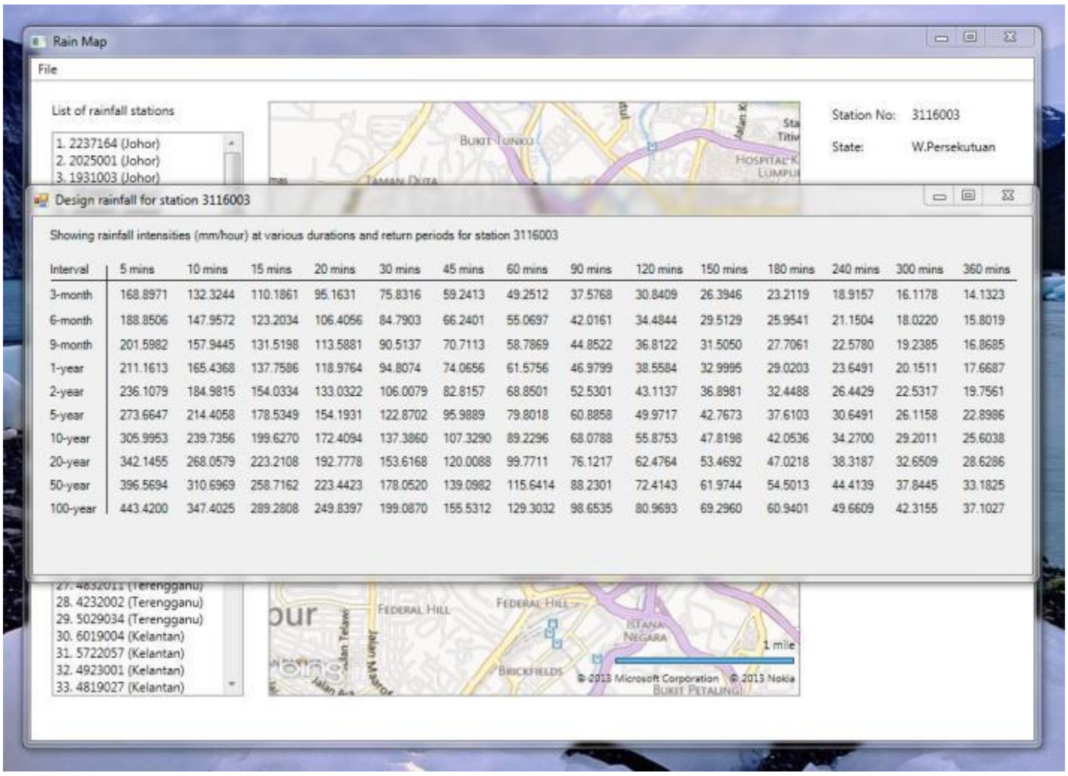


Figure 4.14: RainMap screenshot shows zoomed-in map view of a rainfall station.



Interval	5 mins	10 mins	15 mins	20 mins	30 mins	45 mins	60 mins	90 mins	120 mins	150 mins	180 mins	240 mins	300 mins	360 mins
3-month	168.8971	132.3244	110.1861	95.1631	75.8316	59.2413	49.2512	37.5768	30.8409	26.3946	23.2119	18.9157	16.1178	14.1323
6-month	188.8506	147.9572	123.2034	106.4056	84.7903	66.2401	55.0697	42.0161	34.4844	29.5129	25.9541	21.1504	18.0220	15.8019
9-month	201.5982	157.9445	131.5198	113.5881	90.5137	70.7113	58.7869	44.8522	36.8122	31.5050	27.7061	22.5780	19.2385	16.8685
1-year	211.1613	165.4368	137.7586	118.9764	94.8074	74.0656	61.5756	46.9799	38.5584	32.9995	29.0203	23.6491	20.1511	17.6687
2-year	236.1079	184.9815	154.0334	133.0322	106.0079	82.8157	68.8501	52.5301	43.1137	36.8981	32.4488	26.4429	22.5317	19.7561
5-year	273.6647	214.4058	178.5349	154.1931	122.8702	95.9889	79.8018	60.8858	49.9717	42.7673	37.6103	30.6491	26.1158	22.8986
10-year	305.9953	239.7356	199.6270	172.4094	137.3860	107.3290	89.2296	68.0788	55.8753	47.8198	42.0536	34.2700	29.2011	25.6038
20-year	342.1455	268.0579	223.2108	192.7778	153.6168	120.0088	99.7711	76.1217	62.4764	53.4692	47.0218	38.3187	32.6509	28.6286
50-year	396.5694	310.6969	258.7162	223.4423	178.0520	139.0982	115.6414	88.2301	72.4143	61.9744	54.5013	44.4139	37.8445	33.1825
100-year	443.4200	347.4025	289.2808	249.8397	199.0870	155.5312	129.3032	98.6535	80.9693	69.2960	60.9401	49.6609	42.3155	37.1027

Figure 4.15: RainMap screenshot shows design rainfall of a selected station.

CHAPTER 5

CONCLUSIONS

5.1 Introduction

This chapter contains the outcome of this research, which satisfies the objectives of this study. In general, this study has successfully produced solutions for the problems stated, such as identifying the best approach for derivation of IDF relationship and renewal of IDF curves in Peninsular Malaysia. It is important to ensure that the rainfall data is filtered from empty records and outliers before it is used for generation of IDF curves, as it is found that the quality of the rainfall data could directly affect the outcome. Moreover, the software products or packages developed in this study are reusable and will be very useful for future research considerations and commercial applications. The most important findings of this study are discussed in the following sections.

5.2 The most suitable minimum inter-event time (MIT) for separation of rainfall events in Peninsular Malaysia

There are different methods to produce individual rainfall events from rainfall data. The most used method by researchers around the globe is minimum inter-event time (MIT). However, the selection of MIT used to divide rainfall events has a huge impact on the number and independency of the rainfall events. The result of this study shows that the most suitable MIT for separation of rainfall events in Peninsular Malaysia is 6 hours. Therefore, it is recommended that future researchers should adopt MIT of 6 hours to separate rainfall events in Peninsular Malaysia. Researchers can use the rainfall characteristics analysis in this study to determine the most suitable MIT in other region (outside Peninsular Malaysia), as currently there are no other quantitative methods that serve the same purpose.

5.3 The most suitable approach for deriving rainfall IDF relationship in Peninsular Malaysia

After the GEV/AMS and GPA/PDS approaches are compared (by using L-moment ratio diagram), we have concluded that GPA/PDS is the most suitable approach for deriving rainfall IDF relationship in Peninsular Malaysia. The current approach used by government agency to derive IDF relationship in Malaysia is GEV/AMS. This study has shown the difference between IDF curves generated with GEV/AMS and GPA/PDS. The government agency should consider adopting GPA/PDS approach to derive rainfall IDF relationship, especially with increasing flood events in recent years.

5.4 The updated rainfall IDF curves for Peninsular Malaysia and usability of software packages for future study or application

The 60 updated rainfall IDF curves have been integrated into RainMap and can be used to display design rainfall and location of the rainfall station. The government agency can consider to use them as the latest design rainfall guideline for design of water resources structures. RainEMT is recommend for separation of rainfall events and extraction of event maximum or partial duration series. Meanwhile, RainIDF has been uploaded to Github and has been made available for researchers worldwide (http://github.com/kbchang/rainidf). RainEMT and RainIDF will ease the process of rainfall data extraction and plotting of rainfall IDF curves, and thus allow mass generation of rainfall IDF curves over the region. RainMap serves as a better way for viewing design rainfall and can be further developed into a platform that allows researchers to share their rainfall data, IDF curves and design rainfall worldwide.

REFERENCES

Adamowski, K., & Bougadis, J. (2003). Detection of trends in annual extreme rainfall. *Hydrological Processes, 17(18),* 3547-3560.

Ahmad, N. (2008). Characterization of convective rain in Klang Valley, Malaysia. Master Thesis (Hydrology and Water Resources), Universiti Teknologi Malaysia.

Ahmed, M. I., Sinclair, C. D., & Spurr, B.D. (1988). Assessment of flood frequency models using empirical distribution function statistics. *Water Resources Research, 24*, 1323 – 1328.

Arshad, M., Rasool, M. T. R., & Ahmed, M. I. (2002). Power study for empirical distribution function tests for generalized Pareto distribution. *Pakistan Journal of Applied Science, 2,* 1119 – 1122.

Balistrocchi, M., & Bacchi, B. (2011). Modelling the statistical dependence of rainfall event variables by a trivariate copula function. *Hydrol. Earth Syst. Sci. Discuss., 8(1),* 429 – 481.

Beguería, S. (2005). Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology, 303(1–4),* 215-230.

Ben-Zvi, A., & Azmon, B. (1997). Joint use of L-moment diagram and goodness-of-fit test: a case study of diverse series. *Journal of Hydrology, 198,* 245-259.

Ben-Zvi, A. (2009). Rainfall intensity-duration-frequency relationships derived from large partial duration series. *Journal of Hydrology, 367,* 104 – 114.

Bernard, M. M. (1932). Formulas for rainfall intensities of long durations. *Trans. ASCE, 96,* 592 – 624.

Branham, T. L., & Behera, P. K. (2010). Development of a Rainfall Statistical Analysis Tool for Analytical Probabilistic Models for Urban Stormwater Management Analysis. *World Environmental and Water Resources Congress 2010, edited* (pp. 3281–3290), American Society of Civil Engineers.

Cunnane, C. (1989). Statistical Distributions For Flood Frequency Analysis. *Operational Hydrology Report, no. 33*. World Meteorological Organization.

De Michele, C., Zenoni, E., Pecora, S. & Rosso, R. (2011). Analytical derivation of rain intensity–duration–area–frequency relationships from event maxima. *Journal of Hydrology, 399(3–4),* 385-393.

De Michele, C., & Salvadori, G. (2005). Some hydrological applications of small sample estimators of generalized Pareto and extreme value distributions. *Journal of Hydrology, 301,* 37 – 53.

DID (2000). Manual Saliran Mesra Alam (MSMA). Department of Irrigation and Drainage, Malaysia.

Dunkerley, D. (2008). Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site. *Hydrological Processes, 22(26),* 5024 – 5036.

Franczyk, J., & Chang, H. (2009). The effects of climate change and urbanization on the runoff of the Rock Creek basin in the Portland metropolitan area, Oregon, USA. *Hydrological Processes, 23(6),* 805-815.

Gumbel, E. J. (1941). The Return Period of Flood Flows. *The Annals of Mathematical Statistics, 12(2)*, 163 – 190. doi:10.1214/aoms/1177731747

Gumbel, E. J. (1954). Statistical theory of extreme values and some practical applications. *Applied Mathematics Series, vol. 33,* USDC, National Bureau of Standards.

Guo, Y., & Adams, B. J. (1998). Hydrologic analysis of urban catchments with event-based probabilistic models: 1. Runoff volume. *Water Resour. Res., 34(12),* 3421-3431.

Guo, Y., & Baetz, B. (2007). Sizing of Rainwater Storage Units for Green Building Applications. *Journal of Hydrologic Engineering, 12(2),* 197 – 205.

Haile, A. T., Rientjes, T., Habib, E., & Jetten, V. (2010). Rain event properties and dimensionless rain event hyetographs at the source of the Blue Nile River. *Hydrology and Earth System Sciences Discussions, 7(4),* 5805 – 5849.

Haile, A. T., Rientjes, T. H. M., Habib, E., Jetten, V., & Gebremichael, M. (2011). Rain event properties at the source of the Blue Nile River. *Hydrology and Earth System Sciences, 15(3),* 1023 – 1034.

Heneker, T. M., Lambert, M. F., & Kuczera, G. (2001). A point rainfall model for risk-based design. *Journal of Hydrology, 247(1–2),* 54 – 71.

Hollis, G. E. (1975). The effect of urbanization on floods of different recurrence interval. *Water Resour. Res., 11(3),* 431 – 435.

Hosking, J. R. M., & Wallis, J. R. (1997). Regional Frequency Analysis. Cambridge University Press, Cambridge.

ISSE (2011). Statistic of Weather and Climate Extremes: Background. University Corporation for Atmospheric Research. Retrieved on 29 December, 2011 from http://www.isse.ucar.edu/extremevalues/back.html.

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources, 25,* 1287 – 1304.

Kishtawal, C. M., Niyogi, D., Tewari, M., Pielke, R. A., & Shepherd, J. M. (2010). Urbanization signature in the observed heavy rainfall climatology over India. *International Journal of Climatology, 30(13),* 1908 – 1916.

Kisiel, C. C., Duckstein, L., & Fogel, M. M. (1971). Analysis of ephemeral flow in aridlands. *Journal of Hydraulics Division, 97,* 1699 – 1717.

Koutsoyiannis, D. & Baloutsos, G. (2000). Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences. *Natural Hazards, 22(1),* 31 – 51.

Koutsoyiannis, D., Demosthenes, K., & Manetas, A. (1998). A mathematical framework for studying rainfall intensity–duration–frequency relationships. *Journal of Hydrology, 206,* 118 – 135.

Kuo, Y.-M., Chu, H.-J., Pan, T.-Y., & Yu, H.-L. (2011). Investigating common trends of annual maximum rainfalls during heavy rainfall events in southern Taiwan. *Journal of Hydrology, 409(3–4),* 749-758.

Kyselý, J., Picek, J., & Beranová, R. (2010). Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold. *Global and Planetary Change, 72(1–2),* 55 – 68.

Madsen, H., Arnbjerg-Nielsen, K. & Mikkelsen, P. S. (2009). Update of regional intensity–duration–frequency curves in Denmark: Tendency towards increased storm intensities. *Atmospheric Research, 92(3),* 343-349.

Madsen, H., Mikkelsen, P. S., Rosbjerg, D., & Harremoës, P. (2002). Regional estimation of rainfall intensity-duration-frequency curves using generalized least squares regression of partial duration series statistics. *Water Resour. Res., 38(11),* 1239.

Madsen, H., Rasmussen, P. F., & Rosbjerg, D. (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resour. Res., 33(4),* 747-757.

Millington, N., Das, S., & Simonovic, S. P. (2011). The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models. *Water Resource Research Report no.*

077. Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, The University of Western Ontario.

Mohymont, B., Demarée, G. R., & Faka, D. N. (2004). Establishment of IDF-curves for precipitation in the tropical area of Central Africa – comparison of techniques and results. *Natural Hazards and Earth System Sciences, 4,* 375 – 387.

Overeem, A., Buishand, A., & Holleman, I. (2008). Rainfall depth-duration-frequency curves and their uncertainties. *Journal of Hydrology, 348(1–2),* 124-134.

Palynchuk, B., & Guo, Y. (2008). Threshold analysis of rainstorm depth and duration statistics at Toronto, Canada. *Journal of Hydrology, 348(3-4),* 535 – 545.

Pikand III, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics 3,* 119 – 131.

Powell, D., Khan, A., Aziz, N., & Raiford, J. (2007). Dimensionless Rainfall Patterns for South Carolina. *Journal of Hydrologic Engineering, 12(1),* 130-133.

Rasmussen, P. F., Ashkar, F., Rosbjerg, D., & Bobee, B. (1994). The POT method for flood estimation: a review. In Hipel, K.W. (Ed.), *Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Extreme Values: Floods and Droughts, vol. 1* (pp. 15 – 26). Kluwer, Dordrecht, NL.

Restrepo-Posada, P. J., & Eagleson, P. S. (1982). Identification of independent rainstorms. *Journal of Hydrology, 55(1–4),* 303-319.

Rowinski, P. M., Strupczewski, W. G., & Singh, V. P. (2001). A note on the applicability of log-Gumbel and log-logistic probability distributions in hydrological analyses: I. Known pdf. *Hydrological Sciences Journal, 47(1),* 107-122.

Shamsudin, S., Dan'azumi, S., & Aris, A. (2010). Effect of storm separation time on rainfall characteristics - a case study of Johor, Malaysia. *European Journal of Scientific Research, 45(2),* 162-167.

Shin, H. (2009). *Uncertainty assessment of quantile estimators based on the generalized logistic distribution* (Doctoral dissertation, Yonsei University). Retrieved from http://hydro-eng.yonsei.ac.kr

Smith, R. L. (1984). Threshold methods for sample extreme. In Tiago de Olivierra, J. (Ed.), *Statistical Extremes and Applications* (pp. 621– 638). Dordecht, Netherlands.

Stephens, M. A. (1986). Tests based on EDF statistics. In D'Agostino, R.B., Stephens, M.A. (Eds.), *Goodness-of-Fit Techniques* (pp. 97 – 193). Decker, New York.

Suhaila, J., Jemain, A. A., Hamdan M. F., & Wan Zin, W. Z. (2011). Comparing rainfall patterns between regions in Peninsular Malaysia via a functional data analysis technique. *Journal of Hydrology, 411(3–4),* 197-206.

Todorovic, P. (1978). Stochastic models of floods. *Water Resources Research, 14,* 345 – 356.

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research, 47,* 123 – 138. doi:10.3354/cr00953.

Van de Vyver, H., & Demarée, G. R. (2010). Construction of Intensity–Duration–Frequency (IDF) curves for precipitation at Lubumbashi, Congo, under the hypothesis of inadequate data. *Hydrological Sciences Journal, 55(4),* 555-564.

Villarini, G., Smith, J. A., Baeck, M. L., Vitolo, R., Stephenson, D. B., & Krajewski, W. F. (2011). On the frequency of heavy rainfall for the Midwest of the United States. *Journal of Hydrology 400(1–2),* 103 – 120.

Zalina, M. D., Desa, M. N. M., Nguyen, V.-Y.-V., & Kassim, A. H. M. (2002). Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science and Technology, 45(2),* 63 – 68.

LIST OF PUBLICATIONS

Chang, K. B., Lai, S. H., & Othman, F. (2013). RainIDF: automated derivation of rainfall intensity-duration-frequency relationship from annual maxima and partial duration series. *Journal of Hydroinformatics, 15(4),* 1224 – 1233. doi:10.2166/hydro.2013.192

  Journal of Hydroinformatics | 15.4 | 2013

RainIDF: automated derivation of rainfall intensity–duration–frequency relationship from annual maxima and partial duration series

Kian Boon Chang, Sai Hin Lai and Othman Faridah

ABSTRACT

RainIDF, a software tool for derivation of rainfall intensity–duration–frequency (IDF) relationship is developed as an Excel add-in by using Visual Basics for Applications (VBA). The tool is integrated with two of the most widely used statistical distributions for determination of IDF relationship: the generalized extreme value (GEV) distribution for annual maxima series, and the generalized Pareto (GPA) distribution for partial duration series (PDS). It provides automated distribution fitting for rainfall data in the form of annual maxima or PDS for multiple intervals, solving and plotting of rainfall IDF curves. RainIDF uses the Solver add-in function in Excel to solve the coefficients of the empirical IDF formula in one step. The methodology built into RainIDF is discussed and rainfall IDF relationships for several stations in Peninsular Malaysia are derived and compared. RainIDF is available for download on GitHub (http://github.com/kbchang/rainidf) as an Excel add-in.

Key words | annual maxima series, Excel, generalized extreme value distribution, generalized Pareto distribution, partial duration series, rainfall intensity–duration–frequency

Kian Boon Chang
Sai Hin Lai (corresponding author)
Othman Faridah
Department of Civil Engineering,
Faculty of Engineering Building,
University of Malaya,
50603 Kuala Lumpur,
Malaysia
E-mail: *laish@um.edu.my*

INTRODUCTION

The rainfall intensity–duration–frequency (IDF) relationship is an important tool for the determination of design rainfall in water resources structural design, urban stormwater management, flood modeling, etc. In some cases, depth–duration–frequency (DDF) and intensity–duration–area–frequency (IDAF) relationships are used, which serve the same purpose as the IDF relationship. Generally, there are two types of rainfall data series that are widely applied for derivation of IDF relationship: annual maxima series (AMS) and partial duration series (PDS). There are different methods or approaches to derive rainfall IDF relationship (e.g. Koutsoyiannis & Baloutsos 2000; Overeem *et al*. 2008; Palynchuk & Guo 2008; Ben-Zvi 2009; Madsen *et al*. 2009; Van de Vyver & Demarée 2010; De Michele *et al*. 2011), where the selection of good fitting probabilistic distributions for the target region is very important.

RainIDF is developed by using Visual Basics for Applications (VBA) and can be installed as an Excel add-in. Basically, input data in the form of annual maxima or PDS for multiple intervals are inserted into an Excel worksheet and the RainIDF add-in will fit the data with the corresponding probabilistic distribution (generalized extreme value (GEV) or generalized Pareto (GPA) distribution), list out all statistical parameters and optimization procedures to obtain the empirical IDF formula. Besides, it also takes the advantage of Excel's chart plotting functionality to plot the IDF curves automatically based on the derived empirical IDF formula. The output return periods or annual recurrence intervals (ARI) for the IDF relationship derived are 3-month, 6-month, 9-month, 1-year, 2-year, 5-year, 10-year, 20-year, 50-year and 100-year. For other return periods or ARI, it has to be performed manually based on the parameters of the fitted distribution.

doi: 10.2166/hydro.2013.192

APPENDIX A

60 selected rainfall stations in Peninsular Malaysia

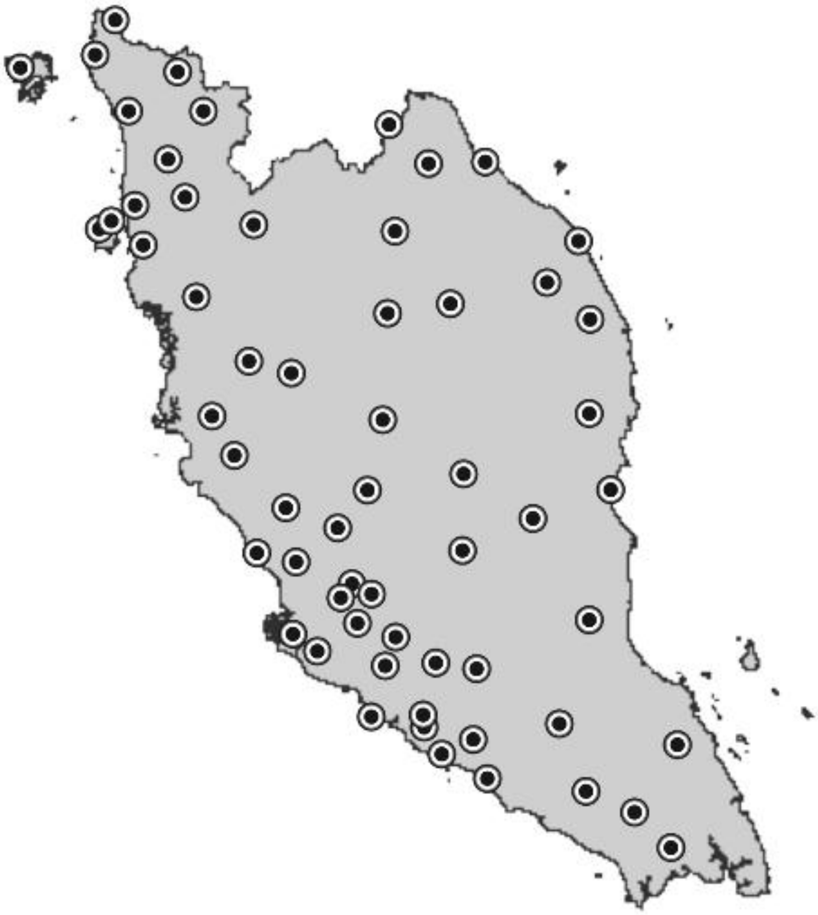


Figure A1. Location of 60 selected rainfall stations in Peninsular Malaysia.

Table A1. Information of 60 selected rainfall stations.

Station No	Latitude	Longitude	Record	Excluded Year	Total Full Year
2237164	02 15 25	103 44 10	June 1970 - Oct 2011	1970, 1974, 1975, 1976, 1978, 1981, 1982, 2000, 2006, 2007, 2011	31
2025001	02 03 05	102 34 40	Aug 1974 - Sept 2011	1974, 1992, 1994, 1995, 1996, 2005, 2006, 2011	30
1931003	01 58 25	103 10 45	Sept 1982 - Oct 2011	1982, 2005, 2006, 2007, 2011	25
1834001	01 50 45	103 28 30	March 1989 - Oct 2011	1989, 2006, 2011	20
1636001	01 37 50	103 41 50	Sept 1980 - Oct 2011	1980, 1989, 1991, 1992, 1994, 2004, 2005, 2006, 2011	23
2330009	02 23 05	103 01 00	June 1970 - Sept 2011	1970, 1974, 1975, 1992, 1993, 1994, 1995, 2005, 2006, 2007, 2011	31
3411017	03 25 25	101 10 23.9	June 1970 - Dec 2011	1970, 1977, 1978, 1992, 1995, 1996, 1997, 2007	34
3613004	03 41 53	101 20 60	June 1970 - Dec 2011	1970, 1973, 1975, 1978, 1979, 1981, 1996, 2005	34
3516022	03 34 33	101 39 56	June 1970 - Dec 2011	1970, 1975, 1977, 1978, 1979, 1991, 2006	35
3314001	03 22 08	101 24 43.9	Jan 1974 - Dec 2011	1974, 1977, 1979, 1995, 1996, 2005	32
3118102	03 10 25	101 52 20	July 1970 - Dec 2011	1970, 1971, 1972, 1973, 1977, 1979, 1982, 2006	34
2917001	02 59 46	101 47 8.9	April 1975 - Dec 2011	1975, 1989, 1998, 2006, 2007, 2008	31
2913001	02 55 50	101 23 35	Dec 1973 - Dec 2011	1973, 1975, 1976, 1977, 1978, 1989, 1992, 1996, 1997, 1998, 2008, 2009	27
2815001	02 49 35	101 32 30	June 1970 - Dec 2011	1970, 1986, 1989, 1990	38
3116003	03 09 05	101 41 05	Dec 1992 - Oct 2011	1992, 2005, 2011	17
3217002	03 14 10	101 45 10	Dec 1972 - Oct 2011	1972, 2005, 2011	37
3833002	03 48 30	103 19 45	May 1985 - Nov 2011	1985, 1989, 1993, 1994, 1996, 1997, 2008	20
3032167	03 01 00	103 11 55	Aug 1981 - Nov 2011	1981, 1982, 1996, 1997, 1999, 2000, 2005, 2006, 2007, 2008	21
3924072	03 54 15	102 26 00	June 1970 - Nov 2011	1970, 1993, 1996, 1997	38
3424081	03 26 20	102 25 35	June 1970 - Nov 2011	1970, 1973, 1975, 1976, 1978, 1979, 1980, 1997, 2004, 2005, 2006, 2007,	28

				2008, 2009	
3818054	03 48 20	101 50 50	July 1970 - Nov 2011	1970, 1975, 1976, 1977, 1978, 1979, 1982	35
4513033	04 31 00	101 23 00	July 1975 - Dec 2010	1975, 1987, 1989, 1999, 2004, 2008	30
3628001	03 38 00	102 51 20	July 1975 - Nov 2011	1975, 1976, 1977, 1979, 1980, 1981, 1986, 1993, 1997, 2006	27
4219001	04 14 00	101 56 25	July 1974 - Nov 2011	1974, 1984, 1985, 1986, 1989, 1996, 2002, 2004, 2005	29
5725006	05 47 50	102 33 55	July 1970 - Nov 2011	1970, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1994, 1995, 1996, 1997, 2002, 2006	26
5331048	05 19 05	103 08 00	June 1970 - Oct 2011	1970, 1971, 1972, 1973, 1983, 1992, 1996, 1998, 2005, 2006, 2011	31
4832011	04 50 35	103 12 15	Dec 1985 - Oct 2011	1985, 1987, 1988, 1989, 1998, 2000, 2005, 2006, 2011	18
4232002	04 16 15	103 11 55	Dec 1985 - Oct 2011	1985, 1993, 1996, 1997, 2005, 2006, 2011	20
5029034	05 04 00	102 56 30	July 1971 - Nov 2011	1971, 1978, 1982, 1985, 1986, 1988, 1990, 1992, 1994, 1998, 2005	30
6019004	06 01 25	101 58 45	June 1970 - Oct 2011	1970, 1974, 1975, 1977, 1978, 1979, 1980, 1981, 1989, 1990, 1991, 1997, 1999, 2005, 2007, 2011	24
5722057	05 47 15	102 13 10	June 1970 - Oct 2011	1970, 1973, 1974, 1976, 1980, 1981, 1982, 1985, 1986, 1987, 1988, 1989, 1990, 1992, 1998, 2000, 2001, 2002, 2008, 2011	22
4923001	04 56 15	102 21 10	Nov 1974 - Oct 2011	1974, 1975, 1981, 1984, 1988, 1989, 1992, 1993, 2006, 2011	28
4819027	04 52 45	101 58 10	July 1971 - Nov 2011	1971, 1975, 1978, 1987, 1988	36
5320038	05 22 40	102 00 55	Sept 1971 - Oct 2011	1971, 1983, 1986, 1987 ,1990, 2000, 2004, 2011	33
5204048	05 17 38	100 28 50	Jan 1988 - Oct 2011	2011	23
5402002	05 26 25	100 17 10	July 1975 - Oct 2011	1975, 2011	35
5302001	05 23 30	100 12 45	July 1970 - Nov 2011	1970	41

5504035	05 32 05	100 25 50	July 1970 - Oct 2011	1970, 1971, 1976, 1996, 2011	37
5411066	05 25 00	101 09 15	June 1972 - Nov 2011	1972, 1973, 1974, 1975, 1976, 1978, 1979, 1980, 1982, 1983, 1984, 2006, 2007, 2008	26
4908018	04 58 45	100 48 15	July 1970 - Nov 2011	1970, 2001, 2003, 2005, 2006, 2007, 2008, 2009	34
4511111	04 35 20	101 07 30	May 1972 - Nov 2011	1972, 1973, 1975, 2003, 2005, 2006, 2007	33
4010001	04 01 00	101 02 10	July 1970 - Nov 2011	1970, 1977, 1978, 1979, 1980, 2007	36
4209093	04 15 20	100 54 00	July 1970 - Nov 2011	1970, 1971, 1972, 1974, 2005, 2007	36
2223023	02 12 00	102 18 00	April 1994 - Nov 2011	1994, 1995, 1998, 2000, 2001	13
2421003	02 26 20	102 11 10	April 1994 - Nov 2011	1994, 1995, 1996, 1997, 1998, 2001, 2002, 2008	10
2224038	02 17 20	102 29 30	July 1970 - Nov 2011	1970	41
2321006	02 21 50	102 11 35	May 1974 - Nov 2011	1974, 1990, 2006, 2007, 2008	33
6306031	06 20 35	100 41 25	July 1970 - Nov 2011	1970, 1982, 2004, 2005, 2006, 2007, 2008	35
6397111	06 21.3 47	99 43.9 03	Sept 1972 - Nov 2011	1972, 1974, 1975, 1976, 1977, 1978, 1979, 1995, 1996, 1997, 1999, 2006, 2007	27
5806066	05 48 50	100 37 55	June 1970 - Nov 2011	1970, 1976, 1978, 1993	39
6108001	06 06 20	100 50 50	Dec 1974 - Nov 2011	1974, 1979, 1981, 2006, 2007	33
6103047	06 06 20	100 23 30	July 1970 - Nov 2011	1970, 1971, 2005	39
5507076	05 35 00	100 44 10	Dec 1977 - Nov 2011	1977, 1979, 1981, 1982, 1992, 2004, 2006, 2007	27
2820011	02 54 44	102 01 10	June 1995 - Nov 2011	1995, 1997, 1998, 1999	13
2719001	02 44 15	101 57 20	June 1970 - Oct 2011	1970, 1975, 1976, 1991, 2011	37
2418034	02 25 40	101 52 15	June 1995 - Oct 2011	1995, 1997, 1998, 1999, 2010, 2011	11
2722002	02 45 20	102 15 50	July 1970 - Oct 2011	1970, 1971, 1972, 1976, 1991, 1992, 2008, 2009, 2011	33
2725083	02 43 10	102 30 45	June 1970 - Sept 2011	1970, 1971, 1973, 1974, 1975, 1976, 1984, 1986, 1991, 1992, 1993, 2011	30
6603002	06 39 25	100 18 35	Dec 1974 - Nov 2011	1974, 1975, 1976, 2004, 2005, 2006	32
6401002	06 26 45	100 11 15	Nov 1974 - Dec 2011	1974, 1975, 1976, 2005, 2006	33

APPENDIX B

Sample of extracted AMS and PDS data

Table B1. AMS data for Station 2330009 at Johor.

Year	5 min	10 min	15 min	20 min	30 min	45 min	60 min	90 min	120 m	150 m	180 m	240 m	300 m	360 m
1971	3.6	7.2	10.8	14.4	18.6	25.9	32.3	45.5	52.5	55.5	58.5	64	70.5	81.3
1972	18.5	23.4	26.1	31.4	34	49.8	60.9	67.6	69.1	69.5	69.5	69.5	72.5	73.4
1973	16	25.1	34.9	42.4	45.7	45.7	52.9	64.7	67.5	67.5	67.5	67.5	67.5	67.5
1976	14.2	23.4	26.8	30.2	36.8	41.5	45.7	51.3	55.1	57.7	57.7	57.7	57.7	59.2
1977	12.9	21.1	28.4	37.6	41.8	46.1	49.1	53.1	53.1	53.1	53.1	53.1	53.1	53.1
1978	15	19.5	23.1	30.8	40.2	41	45.8	55.2	63.7	68.8	80	109	124.7	138.8
1979	12.9	23.8	30.7	32.7	36.6	37.9	39.1	40.2	43.2	50.1	56.1	68.1	80.1	92.1
1980	15.1	25.8	33.8	35.9	40.1	44.5	49.5	65.7	81.9	90.1	90.7	91.5	92.8	96.4
1981	10.5	13.4	16	22	30	45	56.2	76	95.8	100.2	100.3	100.3	100.3	100.9
1982	10.8	21.6	32.4	43.2	47.3	49.4	51.5	58.7	68.5	77.5	78	85.2	99.6	107.7
1983	12.2	24.4	26.8	28.8	33	43.6	53.2	56.6	69	80.4	90.6	90.6	90.6	90.6
1984	6.8	11.2	15.2	19.2	27.2	39.2	51.2	65.5	75.7	85.9	96.1	111	119.4	127.8
1985	25.8	35	35	35	46.2	56.1	63.9	84.1	99.7	109	109.5	110.1	110.1	110.1
1986	34.8	40.5	40.5	40.5	44.2	59.5	73.2	75.8	77.6	79.4	85	85.5	85.5	85.5
1987	27.5	27.5	27.5	33.2	48.7	63.2	72.7	83.5	93.6	93.6	94.2	95.1	95.1	95.1
1988	10.3	17.8	24.2	27	27	39.1	46.3	55.9	55.9	63.5	65.2	65.9	65.9	68.2
1989	20	20	20	20	26.8	34.9	37.4	37.4	37.4	37.6	37.6	37.6	38.3	44.3
1990	10.5	16.5	22.5	30	35.3	35.3	39.6	50.7	51.3	52.8	53.3	53.3	53.8	58.2
1991	24.3	48.5	49	49.5	50.5	52	53.5	56.5	56.9	57.3	57.3	57.8	61.1	61.1
1996	24	24.2	24.4	24.8	29.4	40	40.1	40.1	40.1	40.1	40.3	42.4	46.2	47.2
1997	22.6	23.9	25.8	26.8	28.8	34	34.3	39.8	46.4	50.2	57.9	73.5	89.1	103.2
1998	27.3	30	37.5	49.7	54.5	61.7	68.9	83.3	101.3	111	124.2	137	137.8	140.2
1999	31.4	33	34.6	36.2	39.4	39.6	41.3	51.5	61.7	68.4	68.4	69.9	69.9	70.4
2000	9.5	12.8	19.2	23.7	32.4	42.4	46.7	53.4	58.7	62.7	65.9	66.4	66.4	66.9
2001	12.7	15.5	18.3	21.1	26.7	35.1	39.7	39.7	39.7	42.3	46.5	52.5	56.1	56.1
2002	9.5	11	12.5	14	17	22.5	28.8	40.6	44	53	58.3	58.3	58.3	58.3

2003	22.2	41.7	51.8	54	58.4	65	71.6	84.8	98	104.8	110.8	122.8	134.8	142.7
2004	29	29.2	29.4	29.6	29.9	31.5	42	50.2	57.3	57.3	62.1	74.1	86.1	98.1
2008	11.7	22.2	29.6	36.5	45.4	51.7	59.3	75.5	81	84.6	86.5	99	116.1	124.1
2009	12.8	25	34.2	39.7	51.7	68.5	73.2	74.7	76.9	78.1	78.4	78.7	78.7	78.7
2010	14.8	28.4	37.2	45.5	62.4	79	92.4	98.2	99.1	100.4	101.7	105.3	107.7	109.3

Table B2. PDS data applied with arbitrary thresholds for Station 2330009 at Johor.

5 min	10 min	15 min	20 min	30 min	45 min	60 min	90 min	120 min	150 min	180 min	240 min	300 min	360 min
10.1	15.1	17.6	22.7	27.6	32.8	37.9	43.2	45.6	47.6	50.2	53.1	55.2	55.2
10.1	15.2	17.6	22.7	27.7	33.1	38.1	43.2	45.6	47.6	50.3	53.1	56	55.9
10.1	15.2	17.6	22.8	27.8	33.2	38.2	43.3	45.7	47.8	50.7	53.1	56.1	56
10.1	15.2	17.7	23.2	27.9	33.2	38.4	43.4	45.8	48.1	50.7	53.3	56.4	56.1
10.1	15.2	17.7	23.6	28	33.3	38.4	44	45.9	48.1	51	53.4	56.7	56.4
10.3	15.3	17.7	23.7	28.1	33.3	38.4	44.3	46.1	48.1	51.2	53.5	57.2	56.7
10.3	15.3	18	23.8	28.2	33.4	38.7	44.5	46.1	48.2	51.4	53.6	57.2	57.2
10.3	15.3	18	23.9	28.2	33.6	38.8	44.8	46.4	48.3	51.7	53.7	57.5	57.2
10.3	15.3	18.1	24.1	28.2	33.7	38.9	45.5	46.4	48.4	52	54.3	57.6	57.5
10.5	15.4	18.1	24.2	28.6	33.7	38.9	45.5	46.4	48.6	53.1	54.9	57.7	57.5
10.5	15.4	18.3	24.4	28.8	33.9	39	45.5	46.6	49.1	53.1	55	58.2	58.2
10.5	15.5	18.3	24.4	28.8	34	39.1	45.6	46.7	49.2	53.3	55.2	58.2	58.2
10.5	15.5	18.3	24.5	28.9	34	39.1	45.7	47	49.4	53.4	55.6	58.3	58.2
10.6	15.5	18.5	24.8	29.2	34.1	39.1	47	47.1	49.8	53.6	56.4	58.4	58.3
10.8	15.6	18.5	24.8	29.4	34.2	39.5	47.1	47.4	50.1	53.7	56.7	58.7	58.4
10.8	15.8	18.6	24.8	29.4	34.2	39.6	47.3	47.4	50.2	53.7	57.2	59	58.6
10.8	15.8	18.6	25	29.9	34.3	39.6	47.4	47.6	50.2	54.2	57.2	59.4	59.1
10.9	15.9	18.6	25.1	29.9	34.4	39.6	47.5	47.8	50.7	54.6	57.5	59.7	59.2
10.9	16	18.7	25.1	30	34.5	39.7	47.6	47.8	50.7	54.9	57.7	59.9	59.7
10.9	16.1	18.9	25.5	30	34.6	39.7	47.8	47.9	50.7	55.2	57.7	60.3	59.8
11	16.1	19.1	25.7	30.1	34.7	39.8	47.8	48	50.9	55.4	57.8	60.3	59.9
11	16.1	19.2	25.7	30.1	34.9	40.1	48.1	48.1	51	56.1	58.2	60.7	60.3
11	16.5	19.2	26	30.5	35.1	40.8	48.2	48.1	51.2	56.4	58.3	60.9	60.7
11.1	16.5	19.2	26.1	31.1	35.1	40.9	48.3	48.2	51.4	56.7	58.4	61.1	61.1
11.1	16.6	19.3	26.4	31.1	35.1	41	48.4	48.6	52.8	57.2	58.7	61.7	61.7

11.2	16.6	19.3	26.5	31.3	35.1	41	48.7	49.2	53	57.3	58.9	62.1	62.6
11.2	16.7	19.4	26.5	31.5	35.2	41.2	49.2	49.8	53.1	57.3	59.4	62.2	62.8
11.4	16.9	19.5	26.8	32.1	35.3	41.3	50.2	49.8	53.1	57.5	59.9	62.8	63.5
11.7	16.9	19.6	27	32.2	35.3	41.4	50.3	50	53.2	57.7	60.3	63.4	65.4
11.7	17.1	19.8	27.5	32.3	35.7	41.8	50.7	50.7	53.4	57.9	60.7	64.7	65.6
11.7	17.4	19.8	27.5	32.4	35.8	41.8	51.2	51	53.6	58.2	61.4	65.4	65.6
11.8	17.6	19.9	27.7	32.6	36	42	51.3	51	53.7	58.3	61.7	65.6	65.9
11.9	17.8	20	27.7	32.7	36	42	51.4	51.2	54.3	58.4	62.5	65.6	65.9
12.2	18	20.2	27.9	32.7	36.1	42.2	51.5	51.3	54.9	58.5	64	65.9	66.4
12.2	18	20.5	28.1	32.9	36.1	42.4	51.6	51.4	55.2	58.7	65.2	66.4	66.7
12.2	18	20.5	28.8	32.9	36.2	42.4	51.7	51.5	55.4	59.3	65.4	66.9	66.9
12.3	18.4	20.7	28.8	33	36.2	42.5	52.6	52	55.5	59.6	65.6	67.5	67.5
12.3	18.5	20.7	28.8	33	36.3	43.1	52.6	52.5	55.8	59.8	65.9	68.8	67.6
12.5	18.5	20.8	28.8	33	36.9	43.2	52.8	52.5	55.9	59.9	66.2	69.1	68.2
12.6	18.5	20.8	28.9	33	36.9	43.5	53.1	53.1	56.1	60.5	66.2	69.9	69.4
12.6	18.6	21.1	29.1	33.1	36.9	44	53.1	53.1	56.7	62.1	66.3	70.6	69.6
12.7	18.9	21.2	29.2	34	37.1	44.5	53.3	53.2	57.2	64	66.4	71	70.3
12.7	18.9	21.3	29.4	34	37.4	44.9	53.4	53.4	57.3	64.4	67.5	71.4	70.4
12.8	19	21.4	29.5	34.9	37.5	45	53.7	53.4	57.3	64.4	68.1	71.8	70.5
12.9	19.1	21.6	29.6	34.9	37.9	45.3	53.7	53.7	57.5	65.2	69.5	72.5	70.6
12.9	19.5	21.9	30	35	38	45.5	53.8	53.9	57.5	65.4	69.9	75.1	71
13	19.5	22	30	35.1	38	45.6	53.9	54	57.6	65.6	70.3	77.2	71.4
13	19.5	22.4	30.2	35.3	38.3	45.7	54.5	54.9	57.6	65.9	70.6	77.2	71.8
13.1	19.8	22.5	30.4	35.4	38.8	45.7	54.9	55.1	57.7	67.5	71	77.8	73.4
13.6	19.8	22.5	30.5	35.8	38.8	45.8	55.2	55.1	58	67.6	72.3	78.1	75.1
14	20	22.8	30.8	35.9	39.1	46.1	55.4	55.2	58.7	68.2	73.5	78.2	76.8
14.1	20.1	22.8	30.8	35.9	39.1	46.1	55.7	55.9	58.7	68.3	74.1	78.4	77.2
14.1	20.2	23.1	31.1	36.1	39.2	46.2	55.9	56	59.4	68.4	74.2	78.5	77.2

14.2	20.6	23.1	31.4	36.6	39.2	46.3	55.9	56.3	61.1	68.4	74.6	78.7	77.7
14.2	20.7	23.1	31.4	36.6	39.2	46.7	56.5	56.3	61.8	69.4	74.8	79.5	77.8
14.7	21	23.4	31.6	36.8	39.6	46.8	56.6	56.5	62.7	69.5	75.1	80.1	78.1
14.8	21.1	24.1	32.7	38.2	40	47.4	56.8	56.9	63.2	69.6	77.2	80.9	78.5
15	21.2	24.2	32.9	38.4	40.1	48.2	57.5	57.1	63.4	70.3	78.1	82.5	78.7
15.1	21.6	24.4	33.2	39.4	40.1	49.1	58.2	57.2	63.5	70.6	78.5	82.9	79.5
15.6	21.7	24.4	34	39.6	40.4	49.5	58.7	57.3	64	70.8	78.7	85.5	80.7
15.9	21.8	24.4	35	39.7	40.7	49.7	59.9	57.5	64.1	73.6	79.5	86.1	81.8
16	22	24.5	35.2	39.8	40.8	50.8	61	58.7	64.6	75.5	80.6	86.5	82.2
16.1	22.2	24.9	35.4	40.1	41	51.1	62.7	59.3	65.3	77.2	82.9	88	82.9
16.1	23	25	35.9	40.2	41.4	51.2	63	59.9	65.4	78	84.4	89.1	85.5
16.6	23.4	25.2	36.2	40.3	41.4	51.5	63.2	60.4	65.4	78.1	85.2	90.4	86.7
18.1	23.4	25.3	36.5	40.8	41.4	51.5	64.4	60.6	65.6	78.1	85.5	90.6	87.4
18.4	23.8	25.6	37.6	40.8	41.5	51.7	64.7	61.1	66.8	78.4	86.5	91.4	89.3
18.5	23.8	25.7	38.4	41.4	42.4	52.9	64.9	61.7	67	79.5	88	92.8	90.6
18.5	23.9	25.8	38.7	41.8	42.9	53.2	65.5	62.6	67.5	80	90	95.1	91.4
18.5	24	25.8	39.7	43	42.9	53.2	65.7	63.2	68.3	80	90.6	99.6	92.1
19	24.2	25.9	40.5	43	43.6	53.5	67.6	63.7	68.4	80.8	91.4	100.3	92.8
19.5	24.4	26	41.4	43.2	44.5	53.9	67.8	65	68.5	85	91.5	100.7	95.1
20	25	26.1	41.6	43.8	45	54	73.2	65.4	68.8	86.1	93.6	106.8	96.4
20.6	25.1	26.1	42.4	44.2	45.7	56.1	73.3	65.6	69.5	86.5	95.1	107.1	98.1
22.2	25.2	26.3	43.2	45.4	45.7	56.2	74.6	65.9	72.5	90.6	99	107.7	100.7
22.6	25.8	26.5	45.5	45.7	45.7	56.5	74.7	66.4	77.2	90.7	100.3	110.1	100.9
23.6	26	26.8	49.5	46.2	46	57	75.5	67.5	77.5	90.8	100.7	116.1	103.2
24	27.5	26.8	49.7	47.3	46.1	59.3	75.8	68.5	77.7	94.2	105.3	119.4	107.7
24.3	27.5	27.1	54	48.7	48.1	59.4	76	69	78.1	96.1	109	119.8	109.3
25.8	28.4	27.5		49.3	48.2	60.6	79.8	69.1	79	100.3	110.1	124.7	110.1
27.3	29.2	27.5		50.5	48.7	60.9	83.3	73.4	79	100.7	111	131.1	110.6
27.5	30	28.4		51.7	48.9	61.3	83.5	75.7	79.4	101.7	118.9	134.8	115.2

27.5	30.5	28.9		52.3	49.2	63.9	84.1	76.9	80.4	104.4	119.1	137.8	119.8
29	32.2	29.2		54.5	49.2	64	84.8	77.2	84.6	109.5	122.8		124.1
31.4	33	29.4		58.4	49.4	68.9	98.2	77.2	85.5	110.8	137		127.8
34.8	35	29.6		62.4	49.6	71.6		77.6	85.9	116.5			136.1
	40.5	30.1			49.8	72.7		78.8	89.6	124.2			138.8
	41.7	30.5			50.5	73.2		79.6	90.1				140.2
	48.5	30.7			50.6	73.2		81	93.2				142.7
		31.1			50.9	92.4		81.9	93.6				
		31.8			51.4			82.1	95				
		32.4			51.7			88.4	100.2				
		32.7			52			93.6	100.4				
		33.8			54.1			95.8	104.8				
		34.2			56.1			97.7	109				
		34.6			59.5			98	109.3				
		34.9			61.1			99.1	111				
		35			61.7			99.7					
		37.2			63.2			101.3					
		37.5			65								
		40.4			68.5								
		40.5			79								
		49											
		51.8											

APPENDIX C1

Source project of RainEMT in Microsoft Access

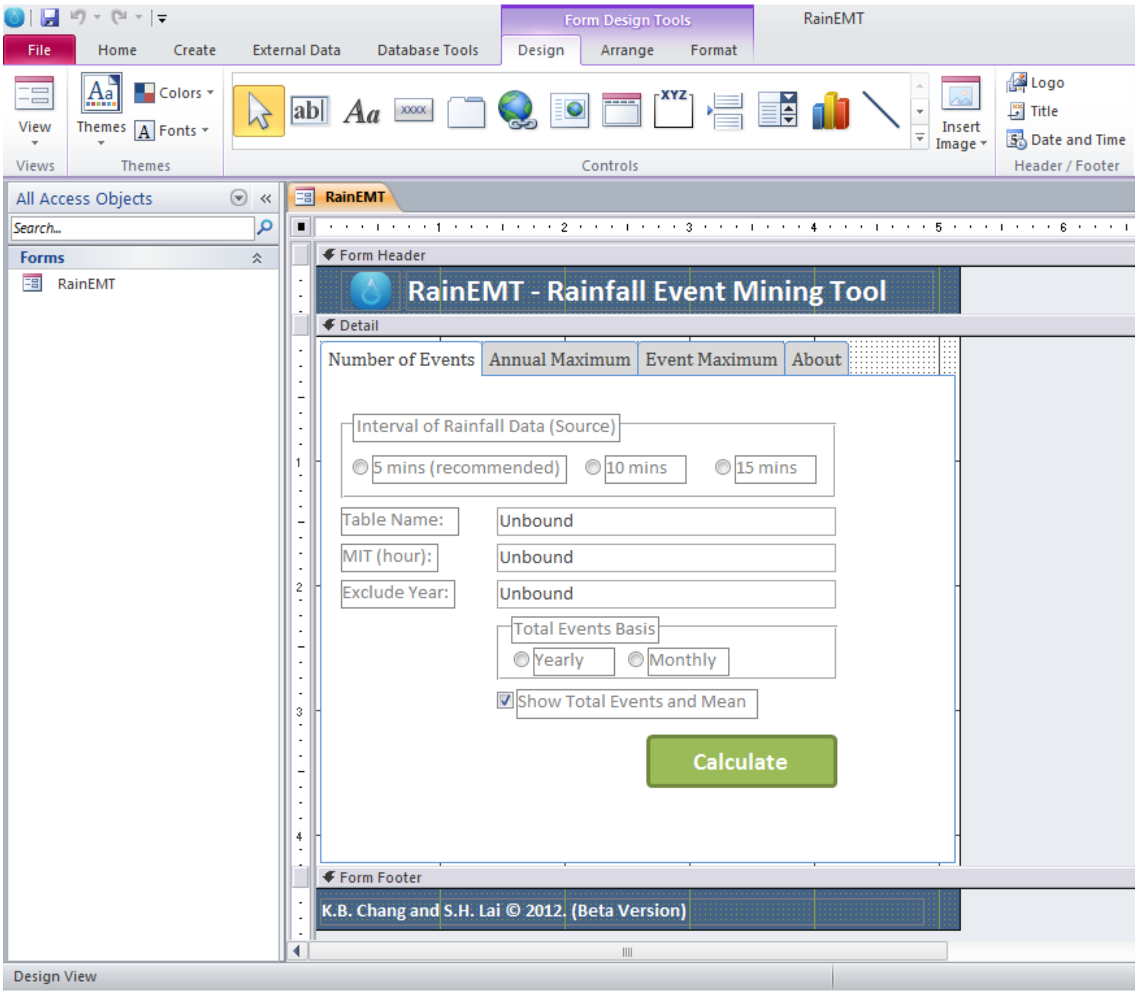


Figure C1. Interface design and development of RainEMT.

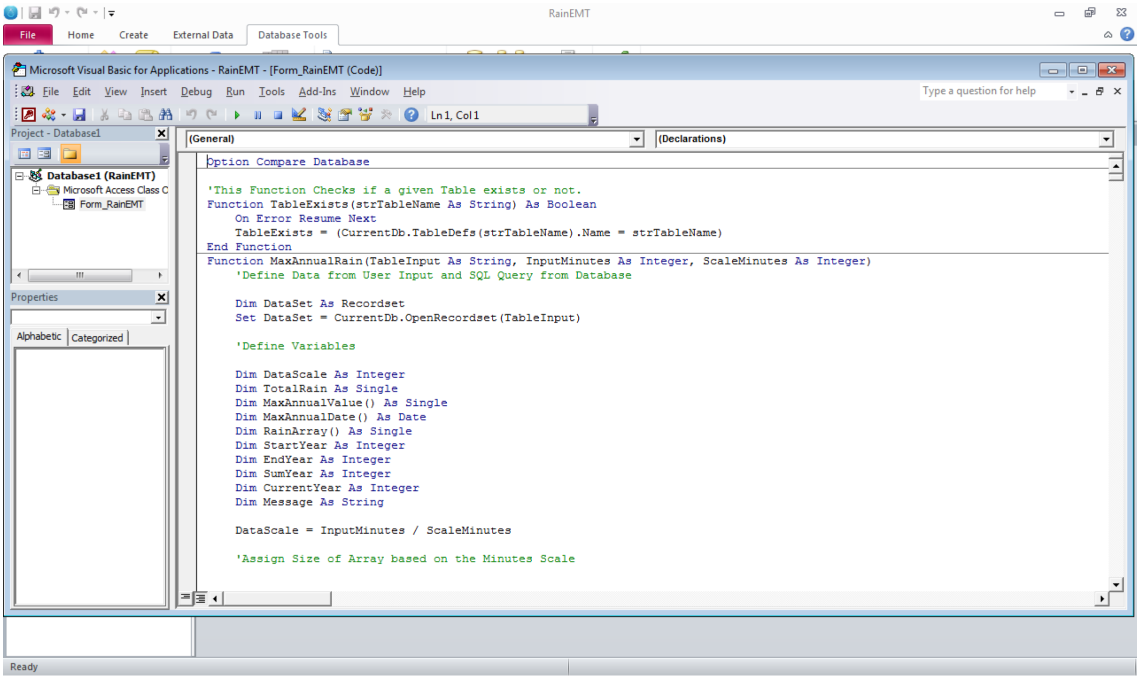


Figure C2. Source code (VBA) of RainEMT: Checking function for table existent and defining source of data.

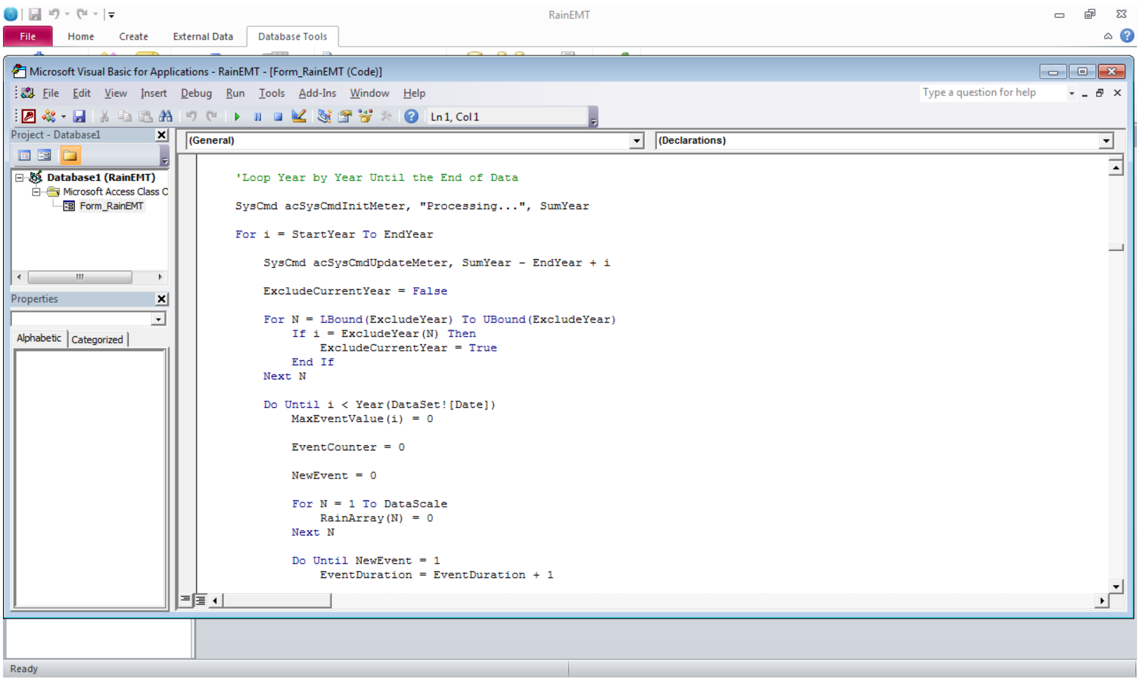


Figure C3. Source code (VBA) of RainEMT: Looping rainfall data from the first year until the last year.

APPENDIX C2

Source project of RainIDF in Microsoft Excel

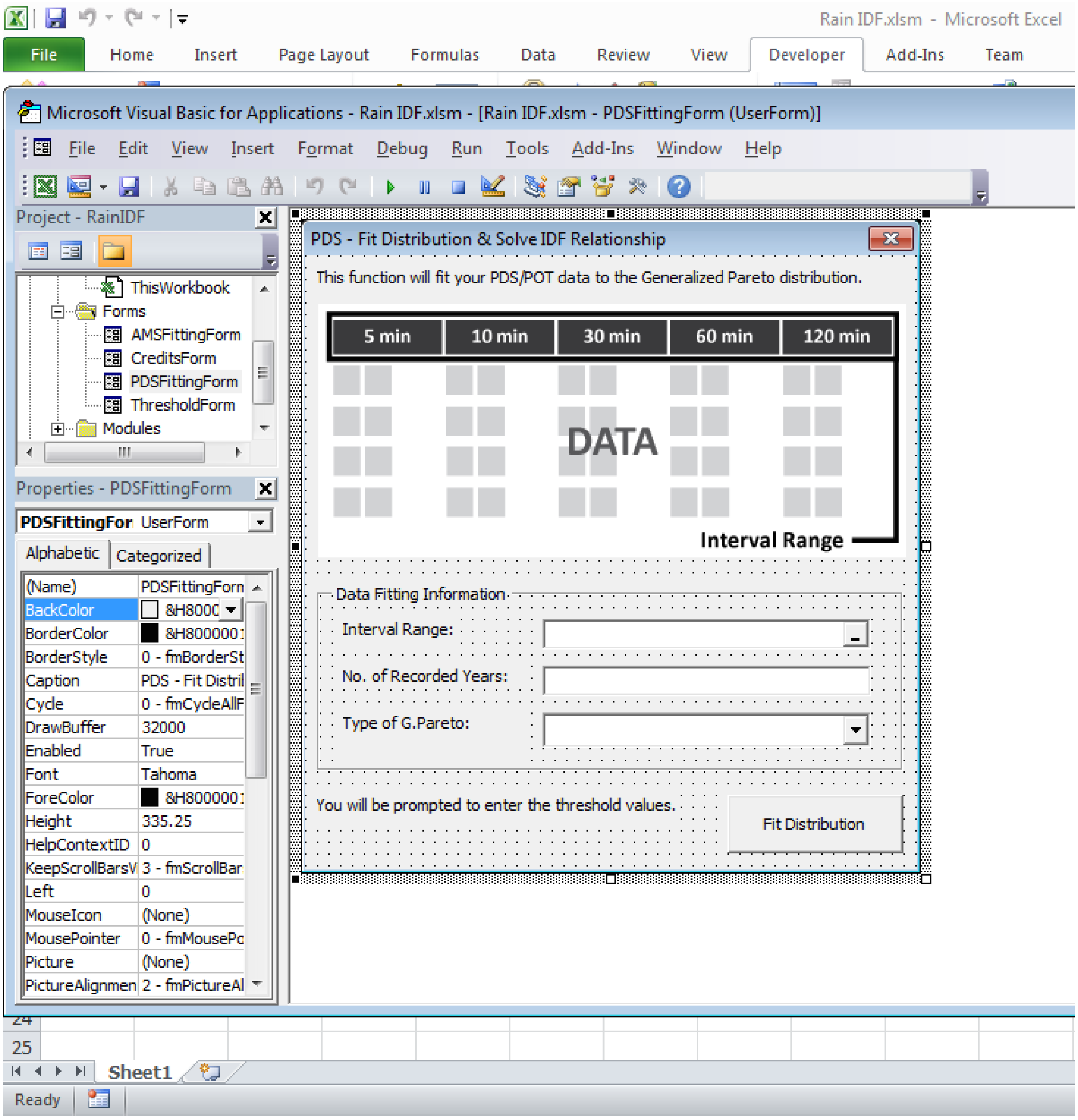


Figure C4. Interface design and development of RainIDF.

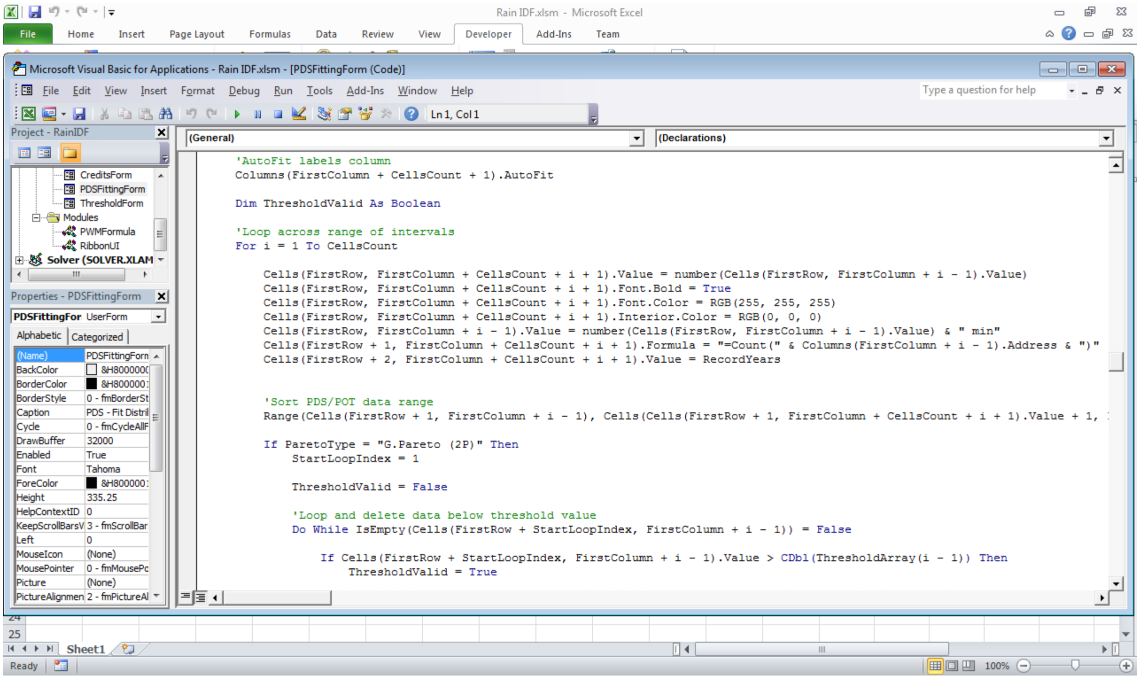


Figure C5. Source code (VBA) of RainIDF: Looping across range of intervals identified from data.

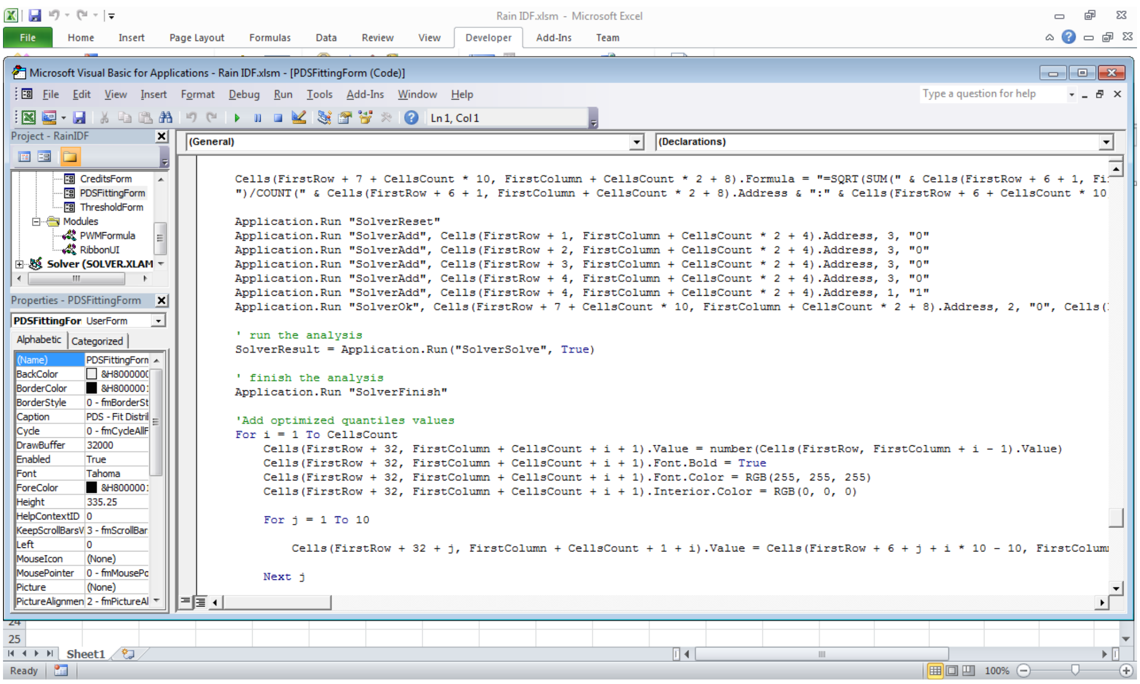


Figure C6. Source code (VBA) of RainIDF: Calling solver add-in function to perform one-step least squares method to solve empirical IDF formula.

APPENDIX C3

Source project of RainMap in Microsoft Visual Studio

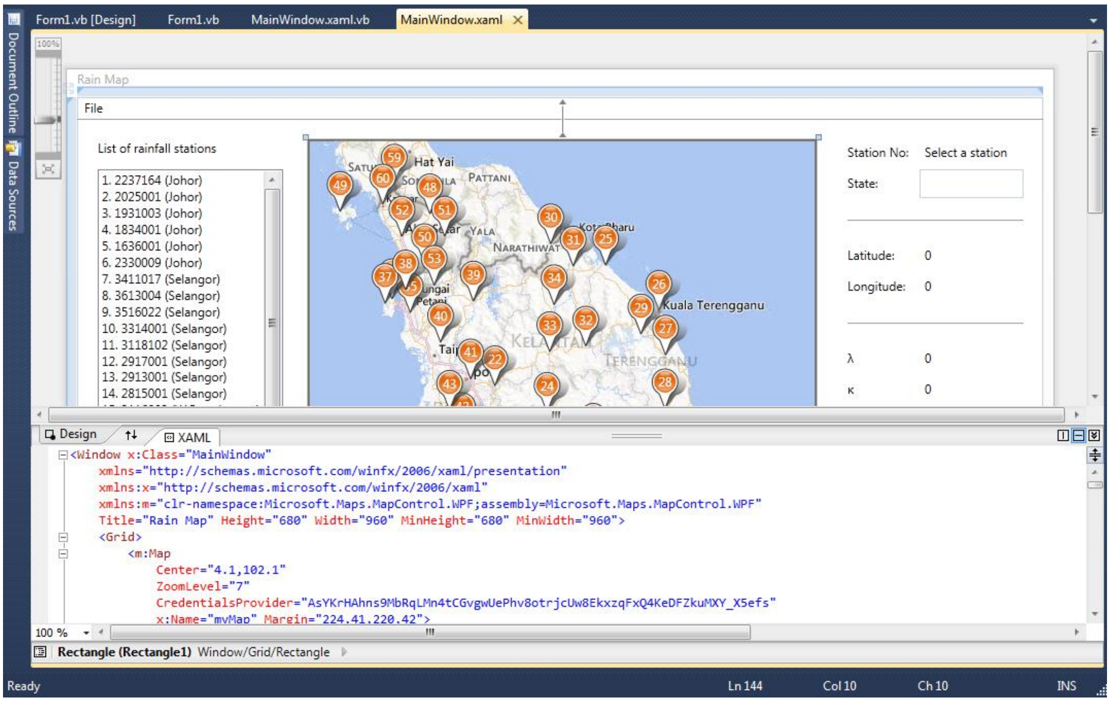


Figure C7. Interface design and development of RainMap.

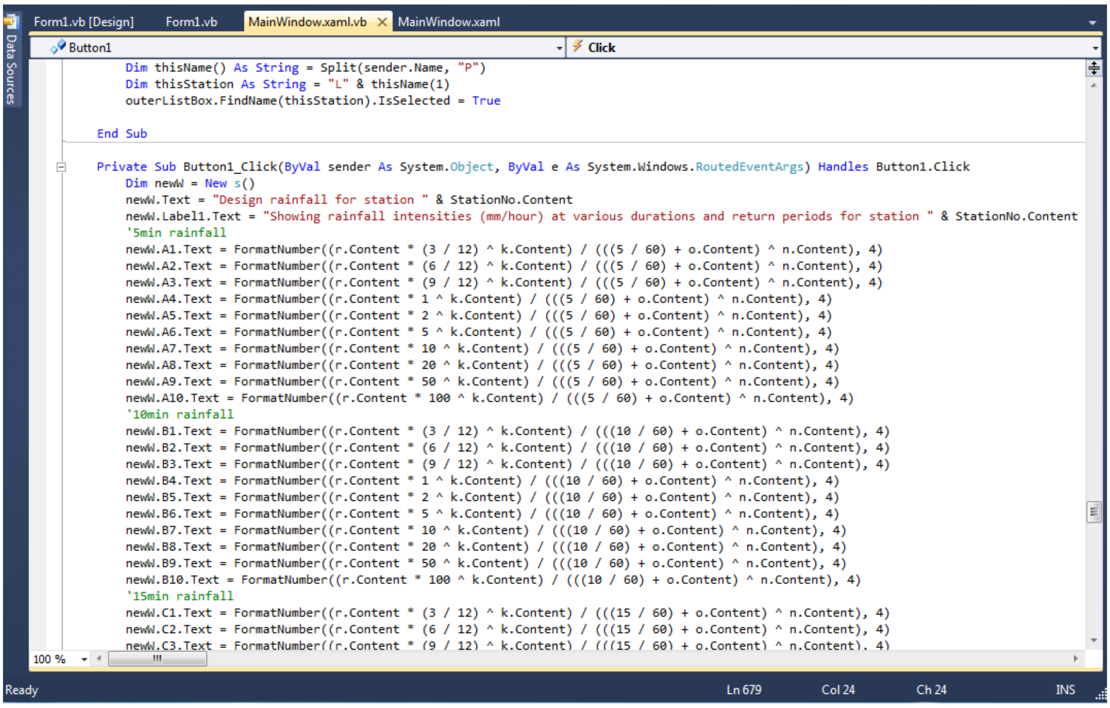


Figure C8. Source code (Visual Basic) of RainMap: Calculation of design rainfall.

APPENDIX D

L-moment ratio data of 60 selected stations

Table D1. Average *L*-moment ratio of PDS and AMS data.

Station	PDS		AMS	
	τ_3 (L-skewness)	τ_4 (L-kurtosis)	τ_3 (L-skewness)	τ_4 (L-kurtosis)
2237164	0.316929374	0.139928291	0.012461963	0.129280581
2025001	0.354096566	0.179545238	0.155916231	0.138473787
1931003	0.301389862	0.136212593	0.125975203	0.17767759
1834001	0.308225675	0.13413849	0.120890866	0.089266048
1636001	0.329073365	0.159374084	0.162222113	0.142336737
2330009	0.310045263	0.1346473	0.107799095	0.096652735
3411017	0.378784587	0.214531093	0.274648827	0.240223076
3613004	0.318058889	0.142379488	0.138107081	0.132565019
3516022	0.35645283	0.148744007	0.183476485	0.12444284
3314001	0.359351847	0.19385604	0.184793104	0.224408463
3118102	0.41372555	0.25989561	0.142724738	0.337004884
2917001	0.325160667	0.168505088	0.103608272	0.155339384
2913001	0.515928459	0.382452314	0.429385368	0.43082528
2815001	0.326316497	0.166664921	0.098540962	0.172598961
3116003	0.357041545	0.202774273	0.191800072	0.253563879
3217002	0.266589401	0.094293883	0.034282431	0.059663566
3833002	0.337983173	0.187642843	0.168232812	0.225214372
3032167	0.337390397	0.153239613	0.065798287	0.098880375
3924072	0.330649843	0.168467509	0.103540669	0.128497442
3424081	0.286303861	0.122792536	0.037550623	0.107976097
3818054	0.327852094	0.176930036	0.134289923	0.187683682
4513033	0.575346464	0.415995324	0.359816738	0.430180434
3628001	0.349506255	0.170757496	0.138150254	0.148858798
4219001	0.317544811	0.158016353	0.154042438	0.126748536
5725006	0.389867934	0.236992628	0.273303196	0.227019069
5331048	0.357015153	0.207313002	0.201914025	0.19162692
4832011	0.332995495	0.191270788	0.167399933	0.212429388
4232002	0.335837883	0.169600763	0.152618249	0.137230017
5029034	0.519377341	0.344155099	0.28518986	0.24580907
6019004	0.421671662	0.268926096	0.301327194	0.28381115
5722057	0.33511961	0.169982548	-0.049138429	0.249702167
4923001	0.369204248	0.185463916	0.167386332	0.14381515
4819027	0.318866986	0.16044218	0.109770092	0.143775771
5320038	0.339595176	0.174722507	0.188010642	0.187493611
5204048	0.442774563	0.269481011	0.264209621	0.214771398
5402002	0.334113795	0.170995728	0.170478814	0.146780829
5302001	0.368817816	0.188710566	0.125614356	0.109938717
5504035	0.291598142	0.14026345	0.059405151	0.143872735
5411066	0.476492649	0.279277296	0.398606423	0.289474411
4908018	0.331590873	0.137613686	0.018615459	0.189997449

4511111	0.383867546	0.221224226	0.222239121	0.2560757
4010001	0.365702913	0.226851847	0.182143843	0.235766001
4209093	0.389593591	0.235799703	0.225874021	0.187929436
2223023	0.299017411	0.122612876	0.003156153	0.224868738
2421003	0.358071971	0.204383306	0.066417401	0.288537638
2224038	0.350732237	0.184706704	0.146840956	0.153506885
2321006	0.353813287	0.159816347	0.154035137	0.087679056
6306031	0.361225939	0.189614708	0.228659596	0.183222181
6397111	0.366227619	0.21210382	0.199137906	0.247119078
5806066	0.345857866	0.18306985	0.137224053	0.159599304
6108001	0.416329337	0.257056175	0.263493947	0.235467722
6103047	0.290138859	0.137477959	0.062551013	0.172317238
5507076	0.303720643	0.129294307	0.112258588	0.090424927
2820011	0.332996203	0.171140175	0.206437221	0.258175846
2719001	0.325532152	0.171763952	-0.008943554	0.264120939
2418034	0.273383554	0.165924287	0.20291038	0.061212657
2722002	0.45893517	0.299324487	0.302874392	0.351915403
2725083	0.24790931	0.107544148	-0.134562668	0.240056079
6603002	0.298884906	0.140744394	0.156566092	0.133301306
6401002	0.375036571	0.178358662	0.20133227	0.132161675
AVERAGE	0.354361061	0.190096794	0.156556856	0.190656137

Table D2. *L*-moment ratio plotting data of GEV and GPA distributions.

GEV		GPA	
τ3	τ4	τ3	τ4
-0.25	0.131338928	-0.25	0.012761133
-0.2	0.118256568	-0.2	-0.000349744
-0.15	0.109244956	-0.15	-0.008012301
-0.1	0.104338135	-0.1	-0.010398579
-0.05	0.103577663	-0.05	-0.007674526
0	0.10701	0	0
0.05	0.114684309	0.05	0.012471234
0.1	0.126650673	0.1	0.029591501
0.15	0.142958722	0.15	0.051219219
0.2	0.163656677	0.2	0.077218896
0.25	0.188790803	0.25	0.107461133
0.3	0.218405279	0.3	0.141822621
0.35	0.252542477	0.35	0.180186144
0.4	0.291243659	0.4	0.222440576
0.45	0.334550085	0.45	0.268480884
0.5	0.382504531	0.5	0.318208125
0.55	0.435153228	0.55	0.371529449
0.6	0.492548206	0.6	0.428358096

These data are used for plotting of *L*-moment ratio diagram (Figure 4.12).

APPENDIX E

IDF curves derived using GEV/AMS and GPA/PDS models

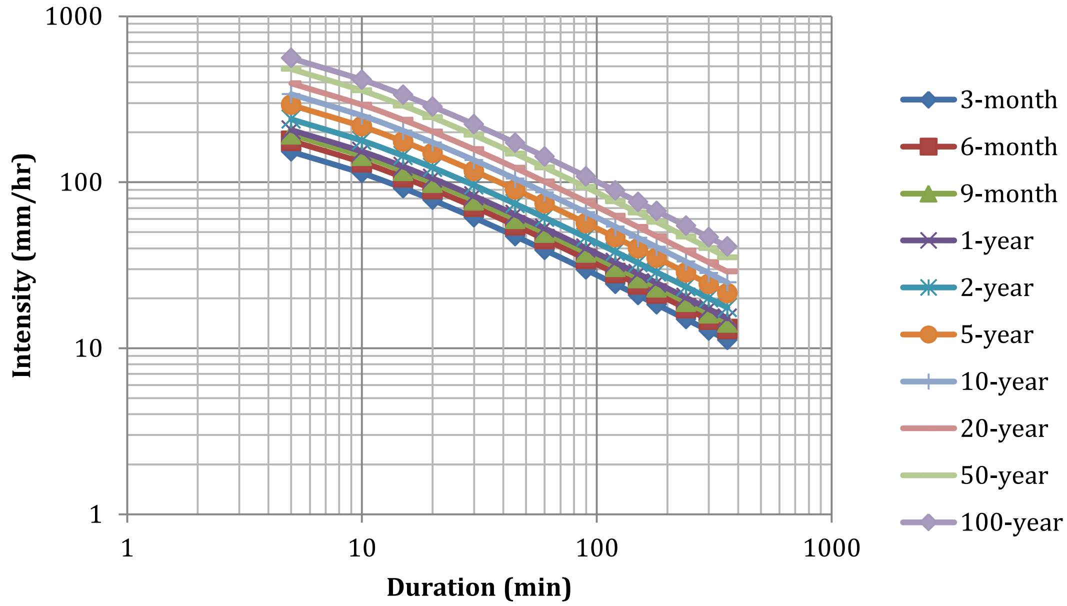


Figure E1. IDF curve based on GEV/AMS model for station 2025001

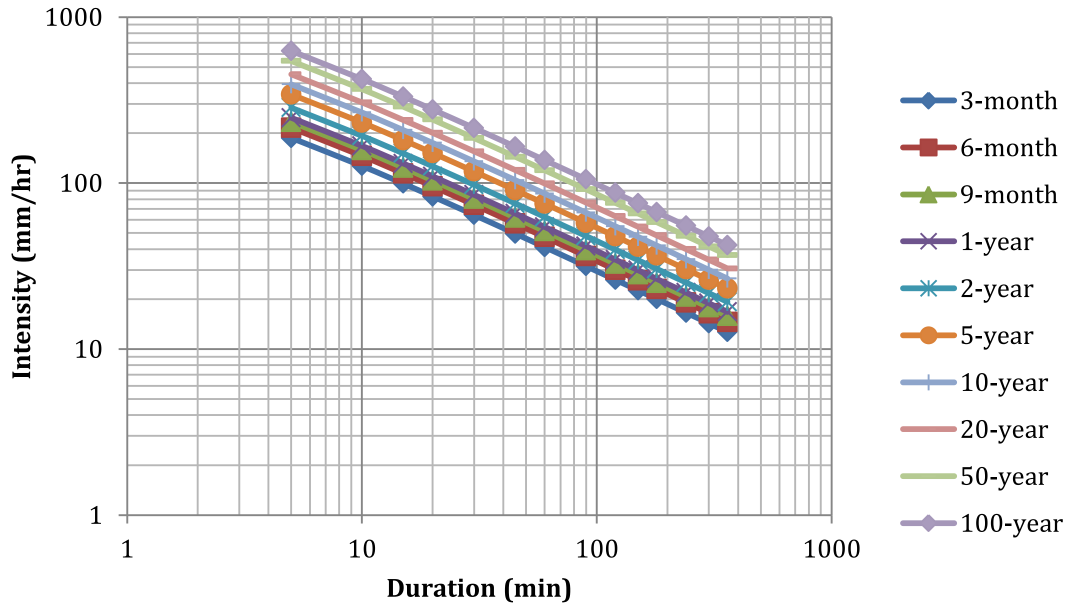


Figure E2. IDF curve based on GPA/PDS model for station 2025001

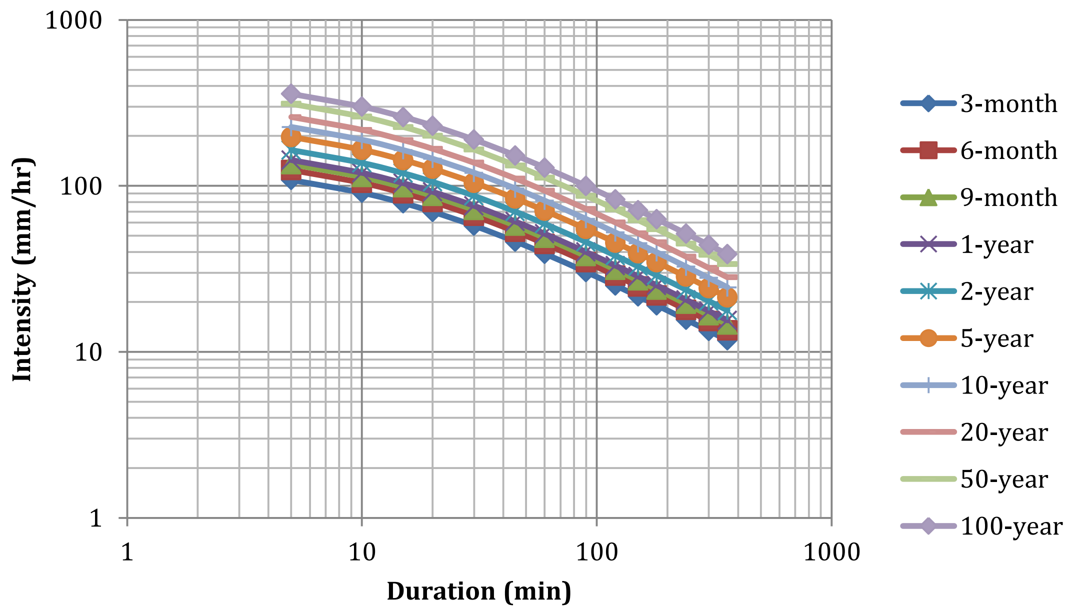


Figure E3. IDF curve based on GEV/AMS model for station 2237164

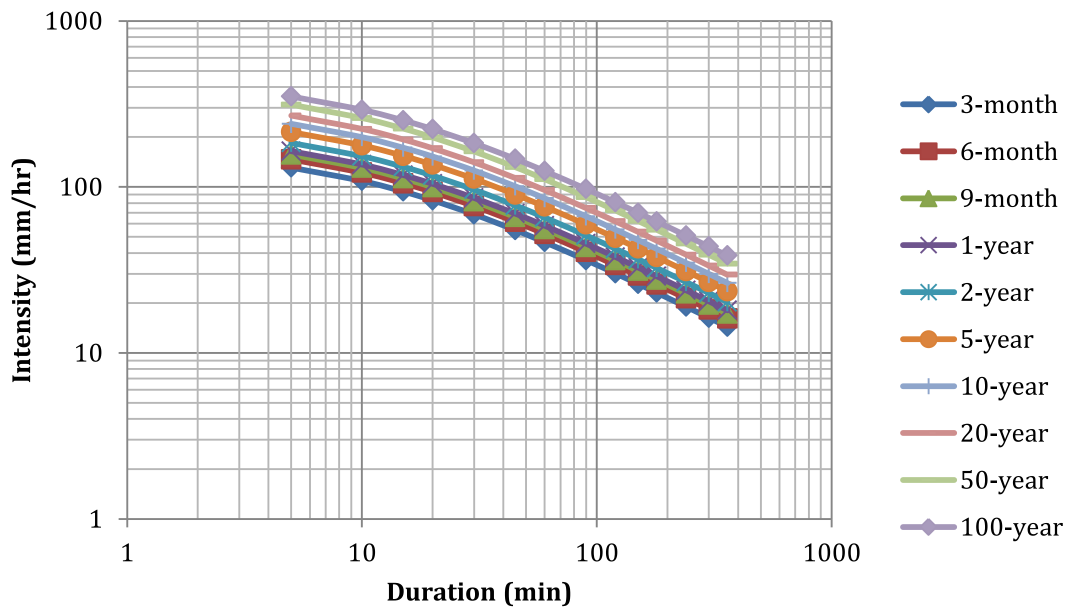


Figure E4. IDF curve based on GPA/PDS model for station 2237164

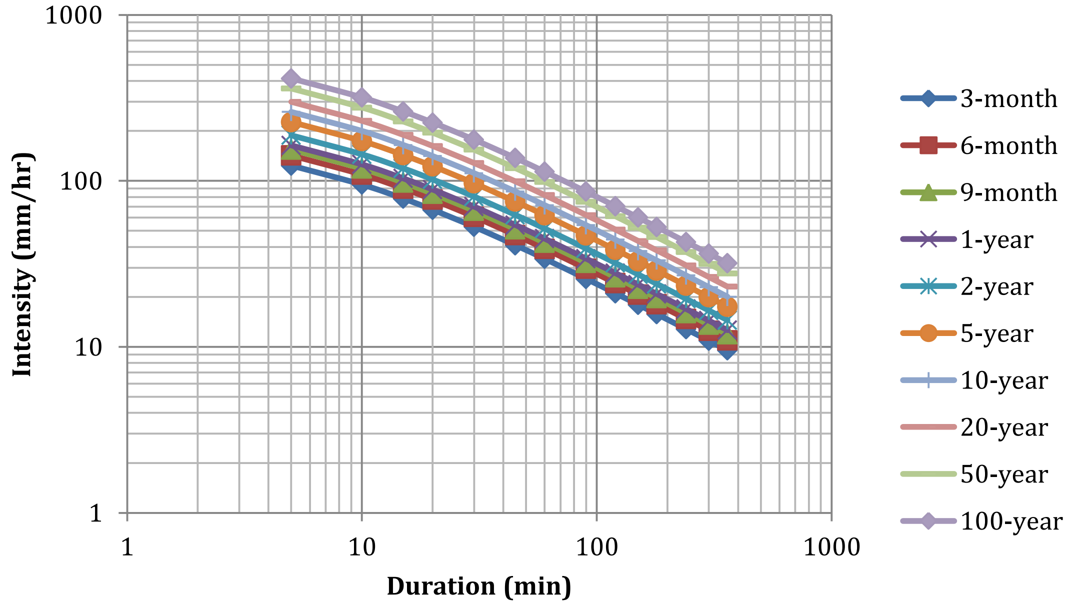


Figure E5. IDF curve based on GEV/AMS model for station 5507076

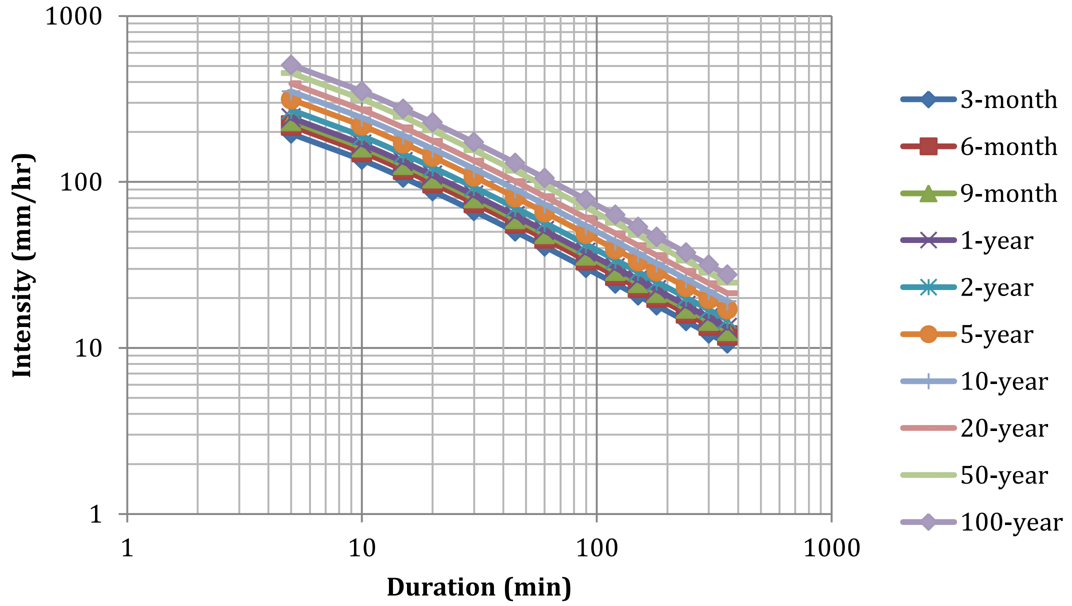


Figure E6. IDF curve based on GPA/PDS model for station 5507076

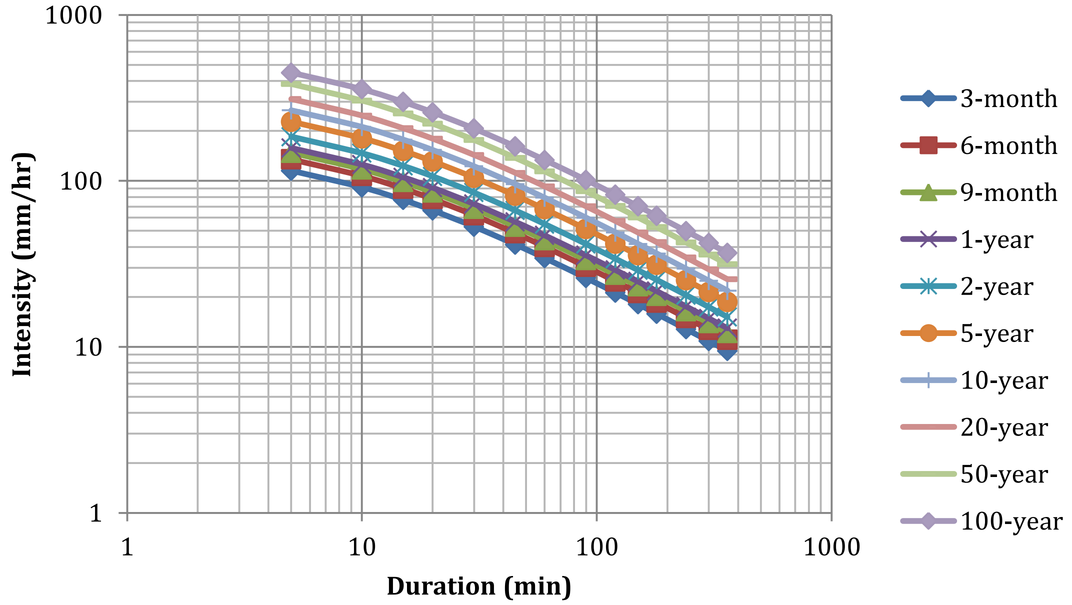


Figure E7. IDF curve based on GEV/AMS model for station 5320038

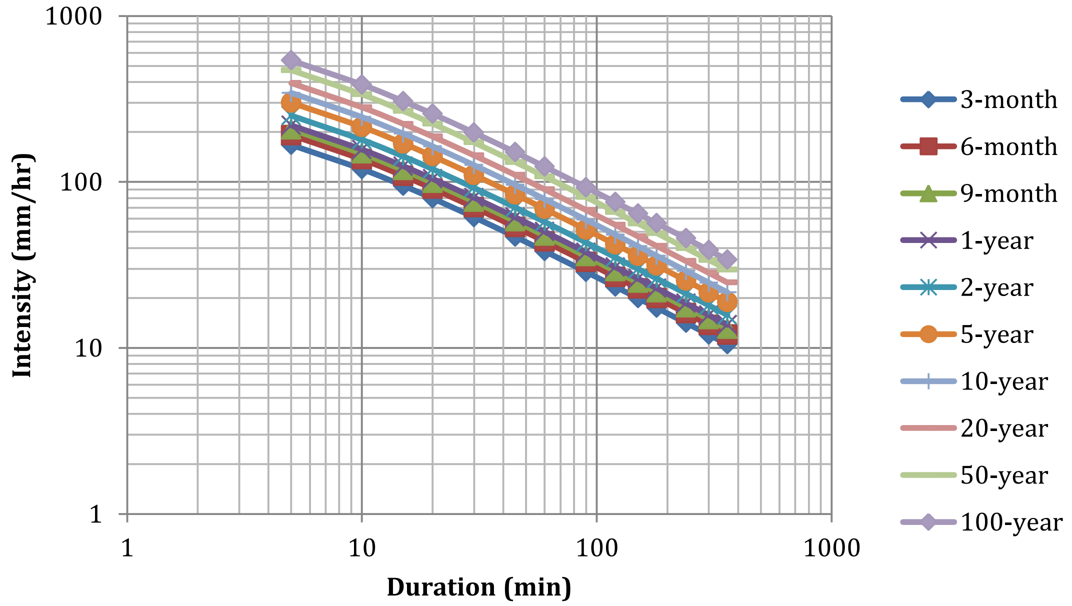


Figure E8. IDF curve based on GPA/PDS model for station 5320038

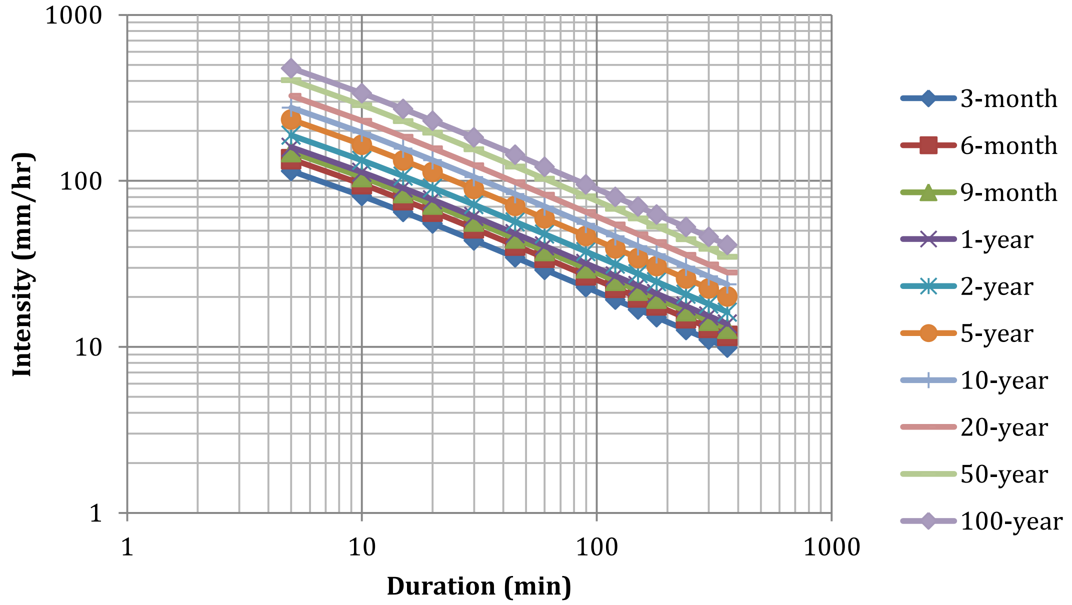


Figure E9. IDF curve based on GEV/AMS model for station 2321006

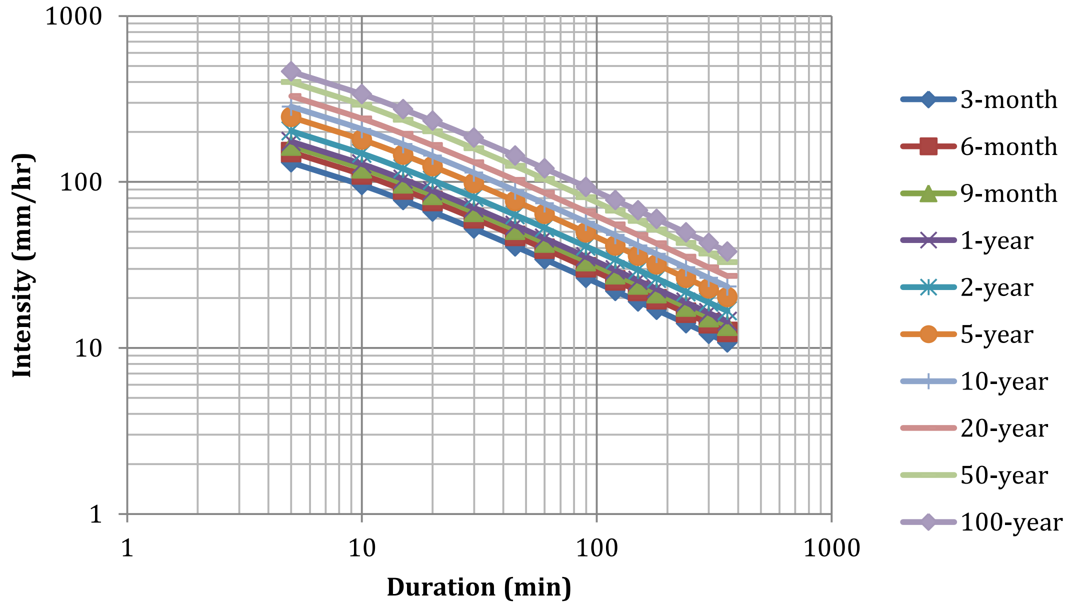


Figure E10. IDF curve based on GPA/PDS model for station 2321006

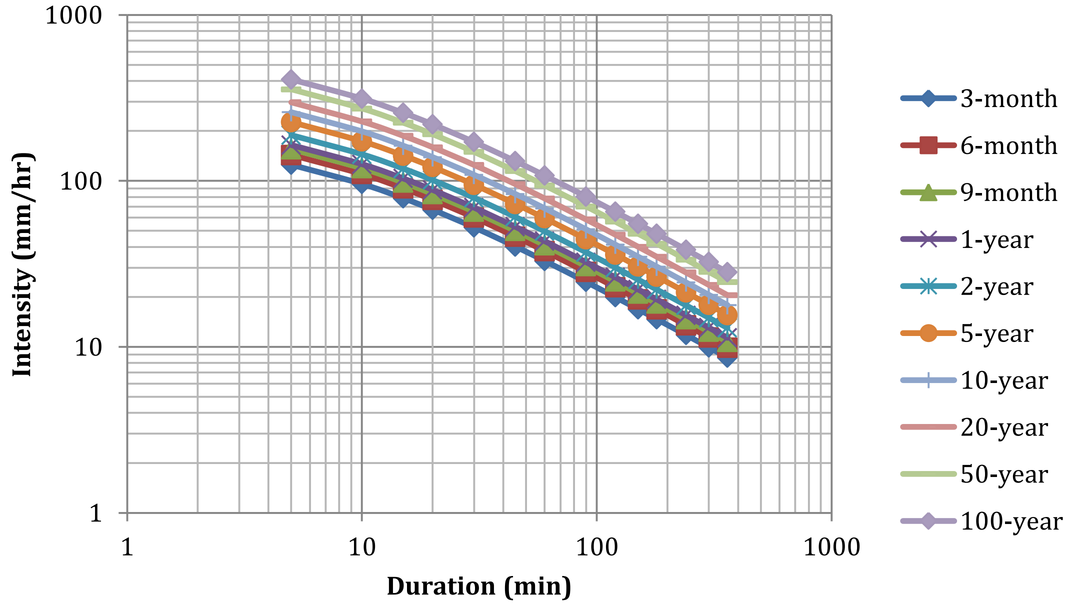


Figure E11. IDF curve based on GEV/AMS model for station 2719001

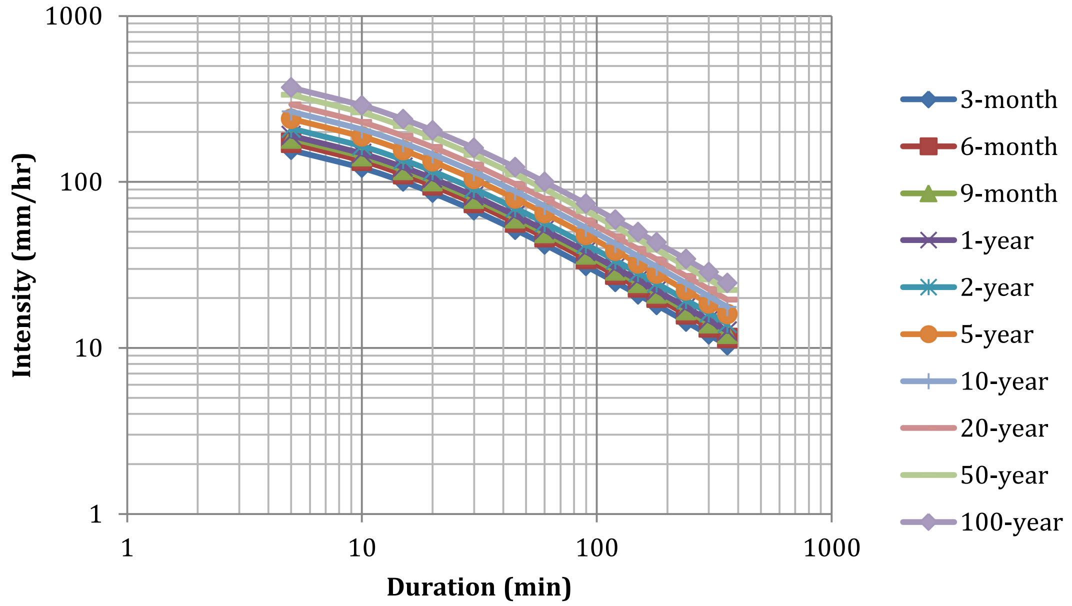


Figure E12. IDF curve based on GPA/PDS model for station 2719001

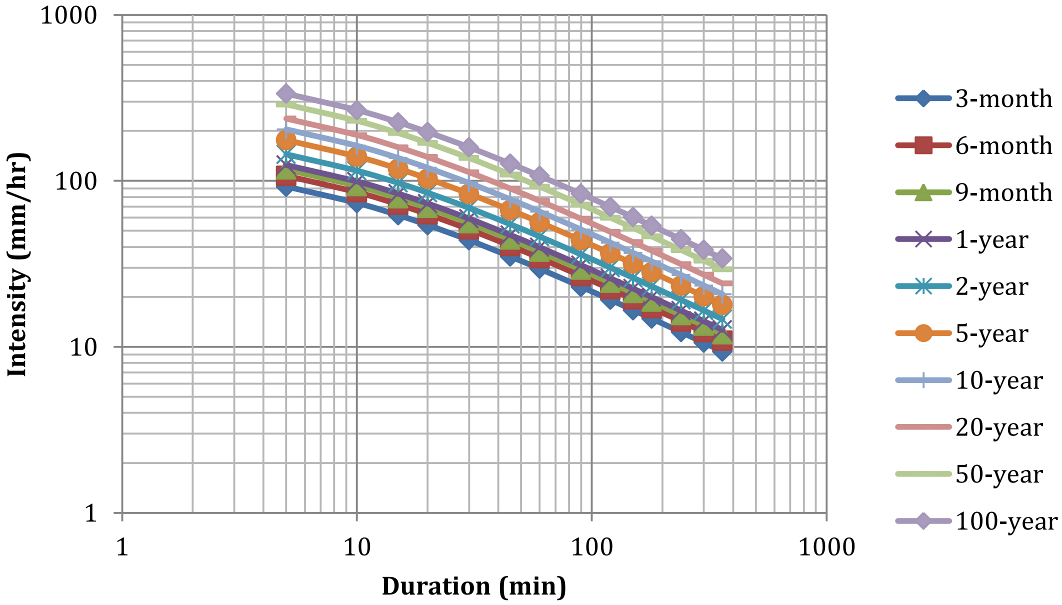


Figure E13. IDF curve based on GEV/AMS model for station 3818054

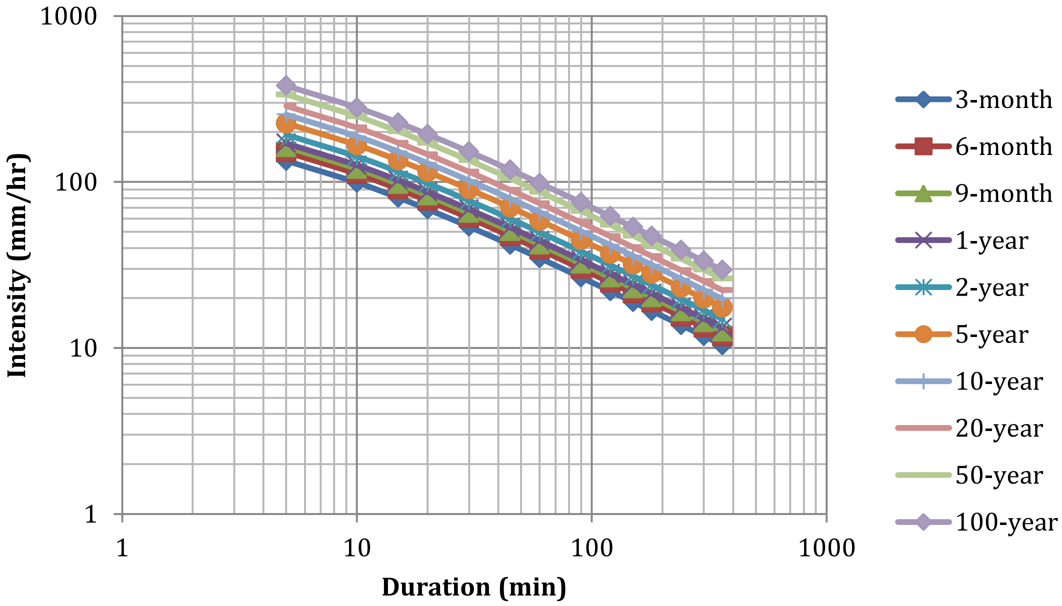


Figure E14. IDF curve based on GPA/PDS model for station 3818054

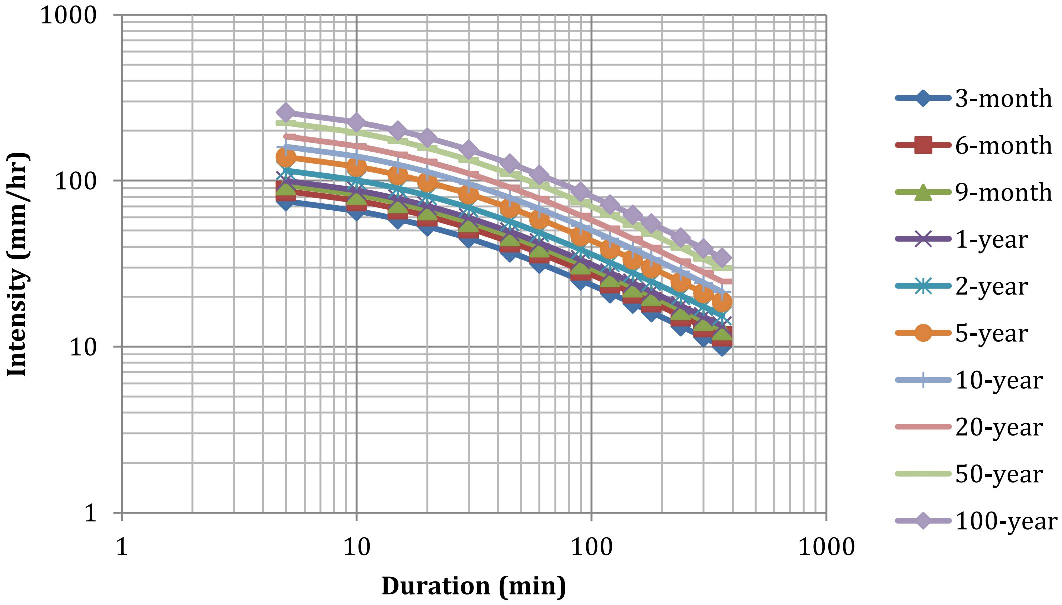


Figure E15. IDF curve based on GEV/AMS model for station 3424081

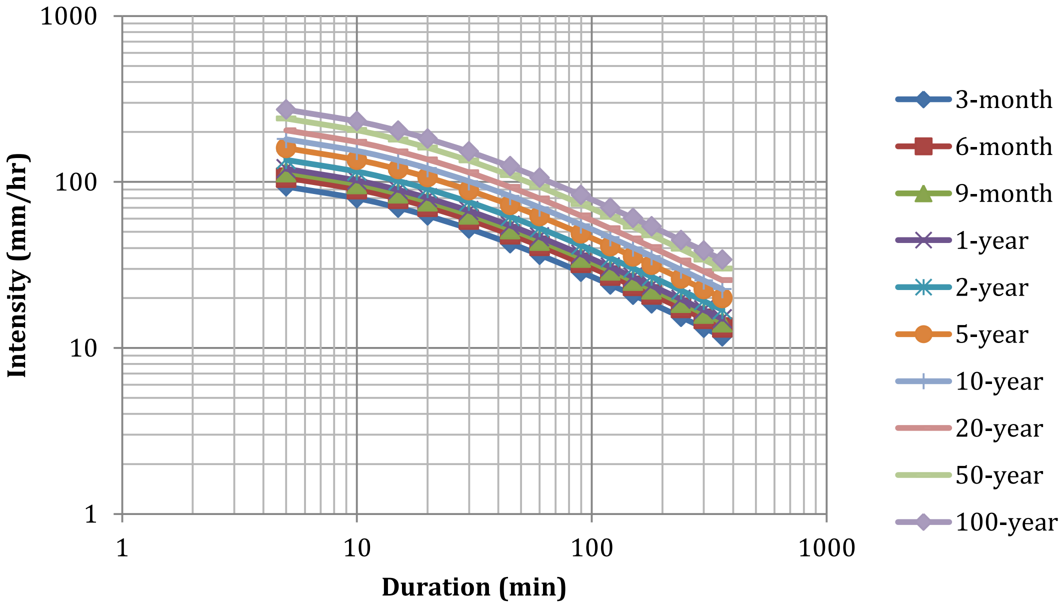


Figure E16. IDF curve based on GPA/PDS model for station 3424081

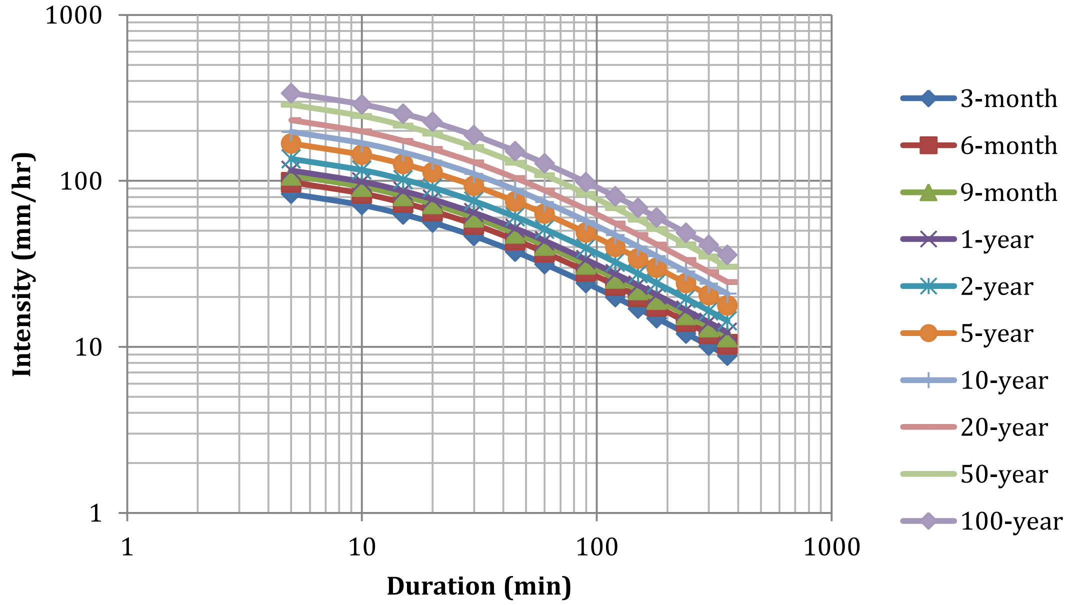


Figure E17. IDF curve based on GEV/AMS model for station 3924072

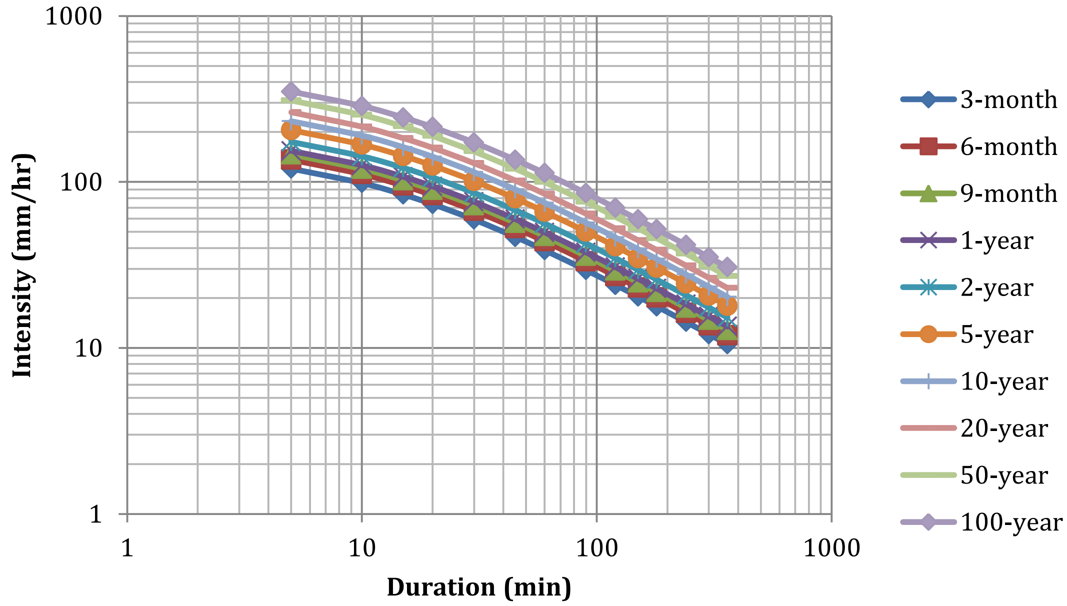


Figure E18. IDF curve based on GPA/PDS model for station 3924072

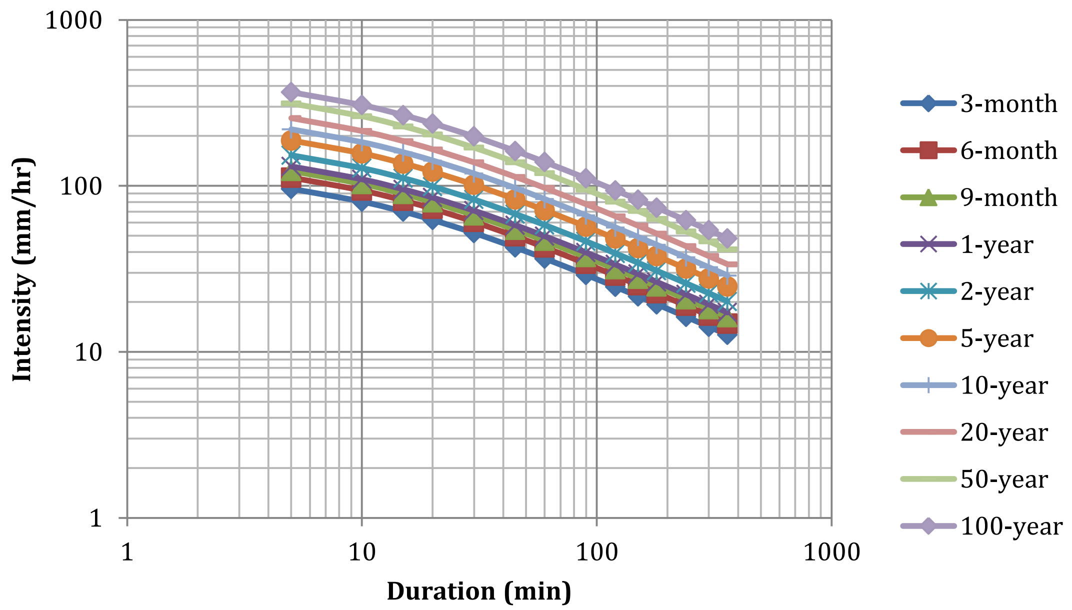


Figure E19. IDF curve based on GEV/AMS model for station 5302001

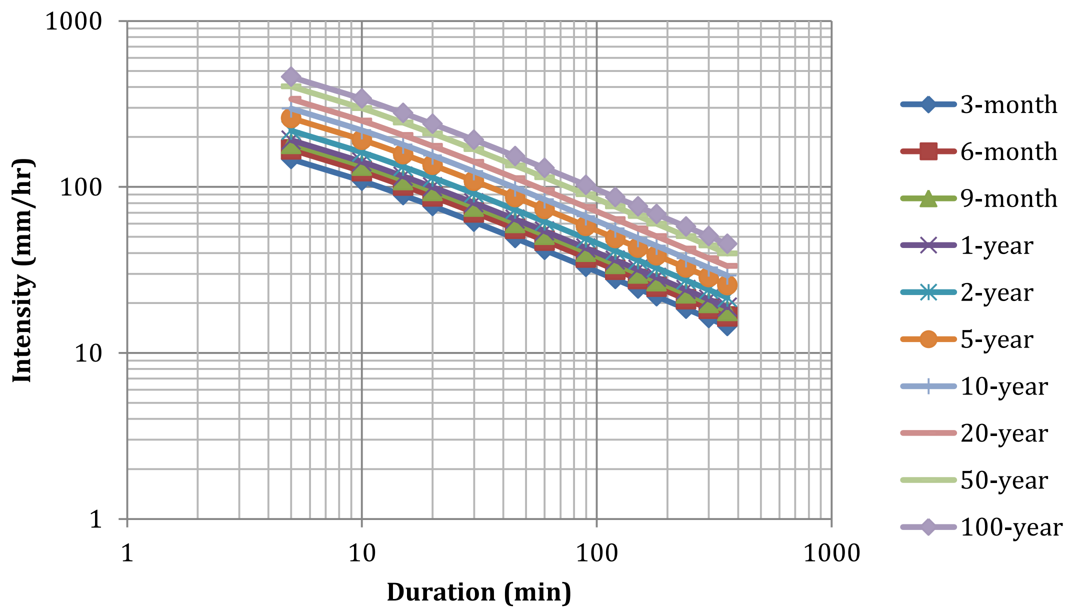


Figure E20. IDF curve based on GPA/PDS model for station 5302001

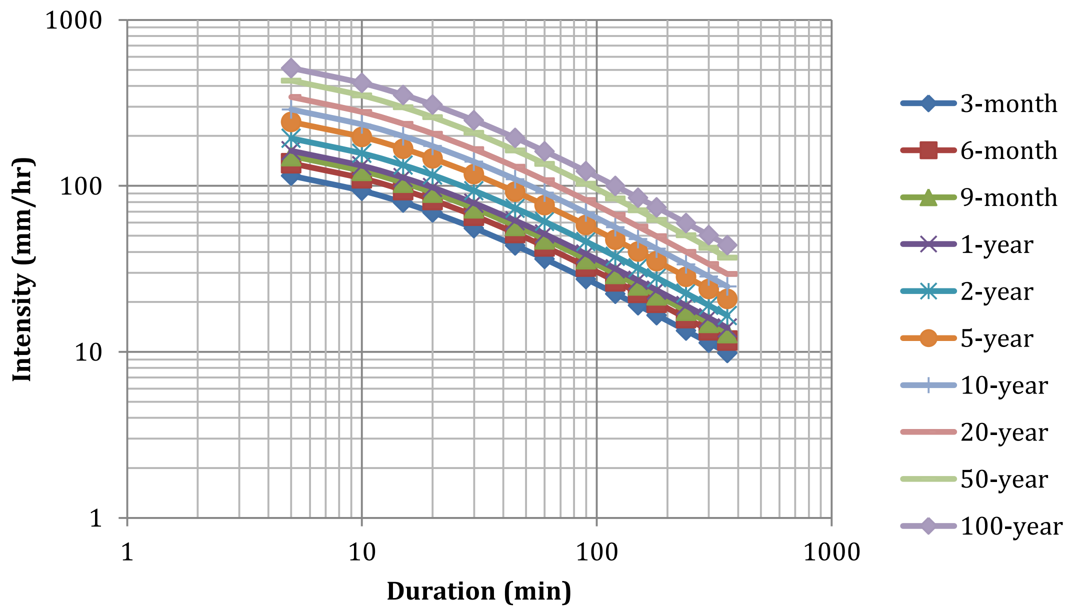


Figure E21. IDF curve based on GEV/AMS model for station 4511111

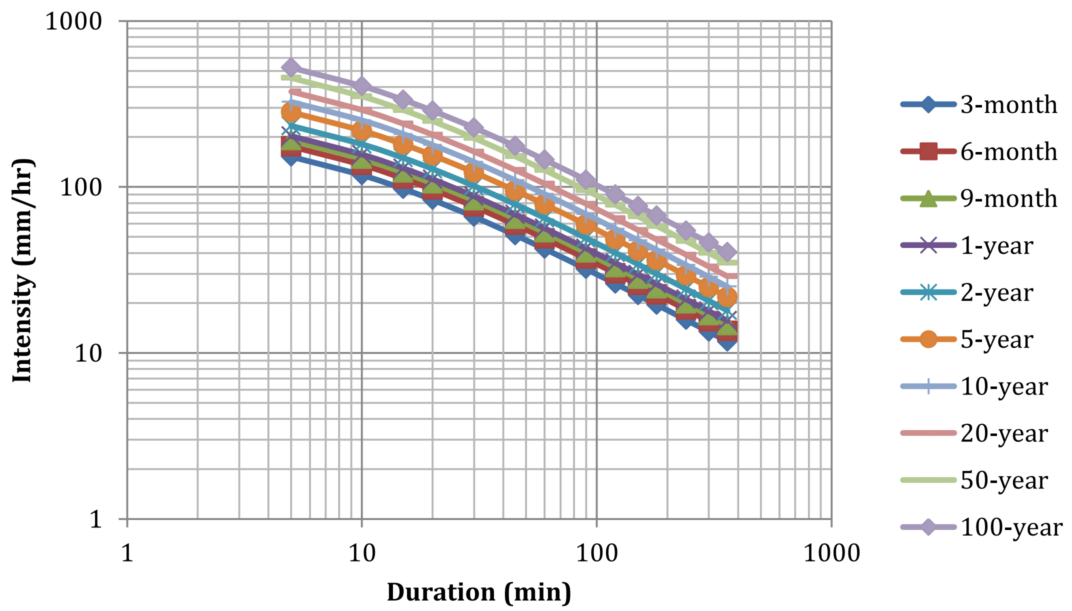


Figure E22. IDF curve based on GPA/PDS model for station 4511111

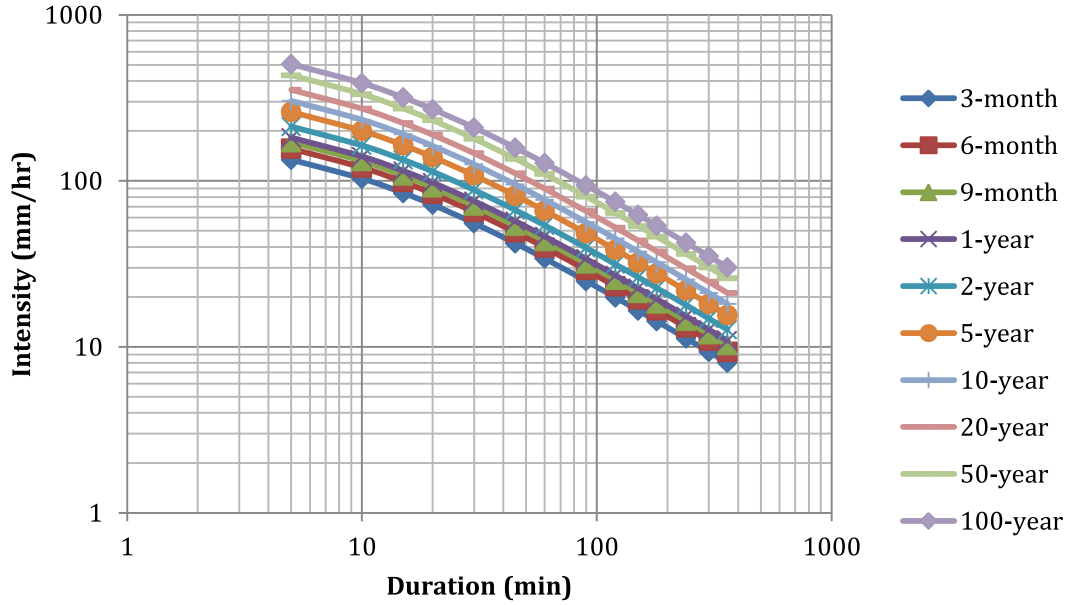


Figure E23. IDF curve based on GEV/AMS model for station 4010001

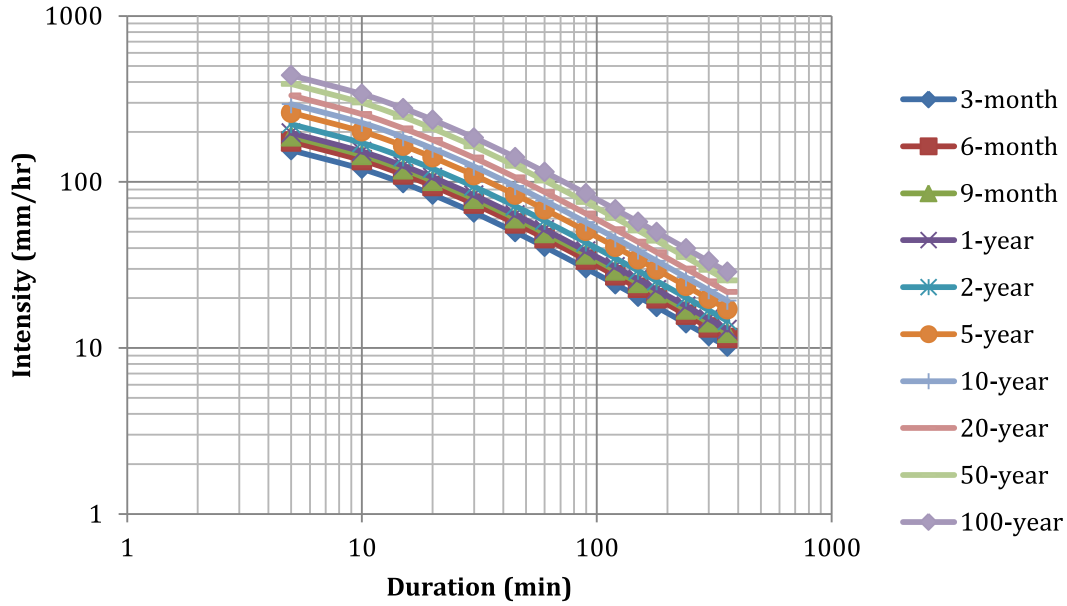


Figure E24. IDF curve based on GPA/PDS model for station 4010001

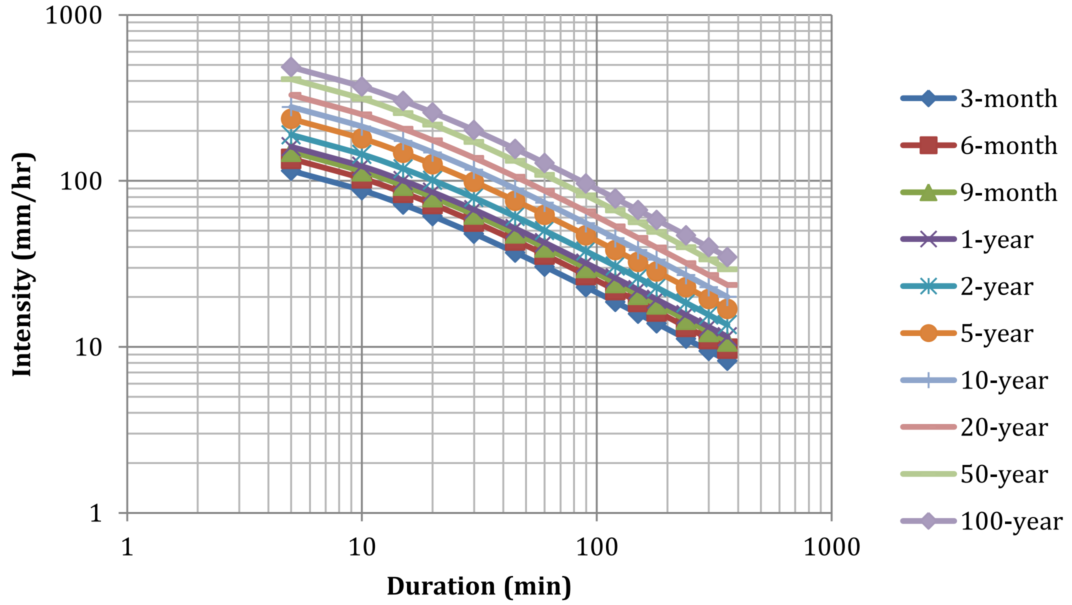


Figure E25. IDF curve based on GEV/AMS model for station 6603002

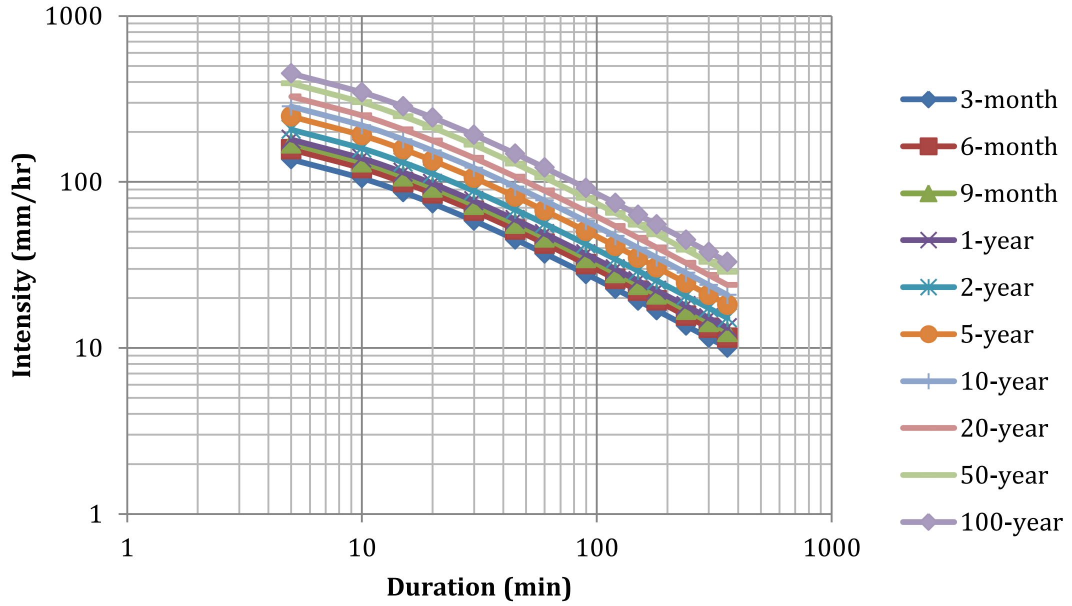


Figure E26. IDF curve based on GPA/PDS model for station 6603002

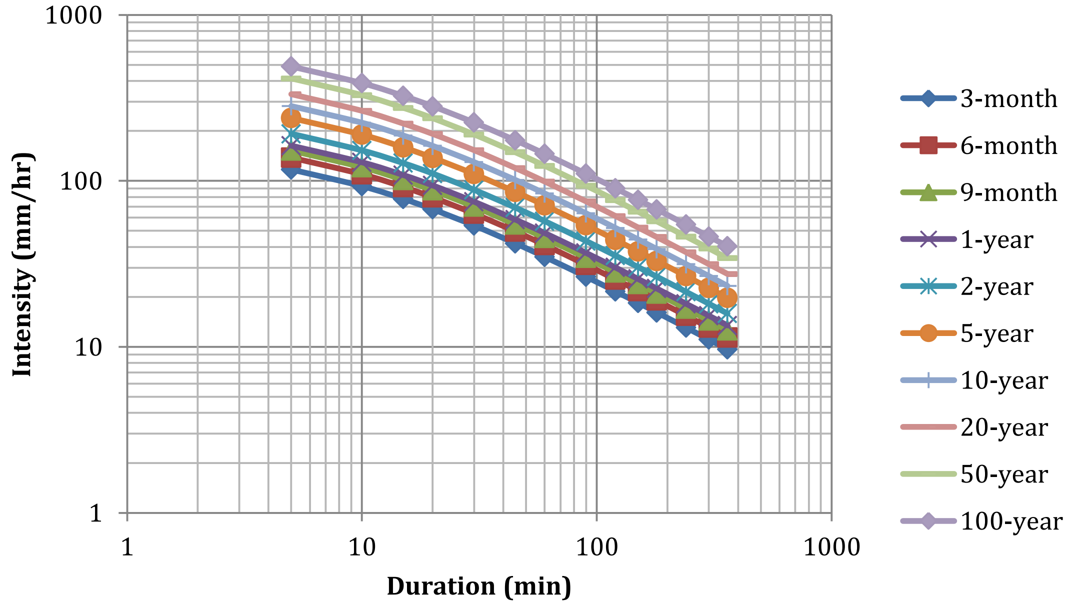


Figure E27. IDF curve based on GEV/AMS model for station 3516022

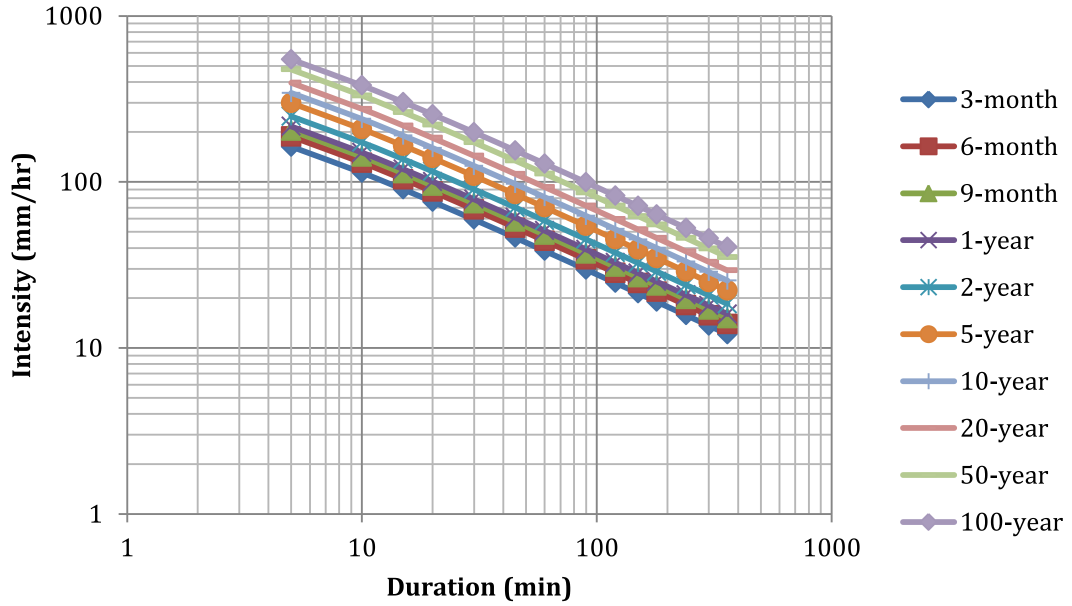


Figure E28. IDF curve based on GPA/PDS model for station 3516022

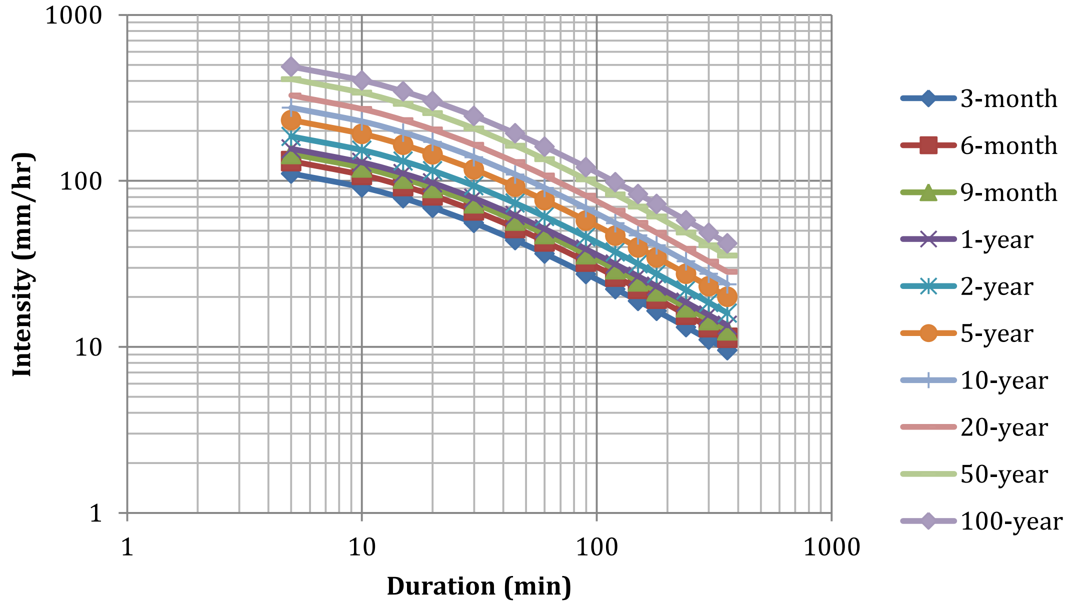


Figure E29. IDF curve based on GEV/AMS model for station 3411017

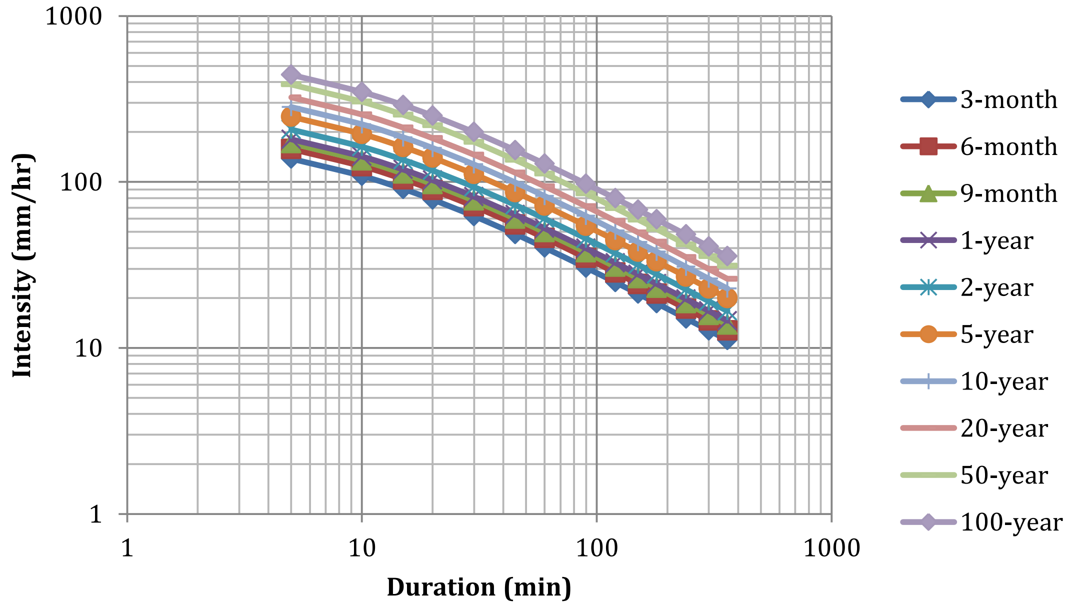


Figure E30. IDF curve based on GPA/PDS model for station 3411017

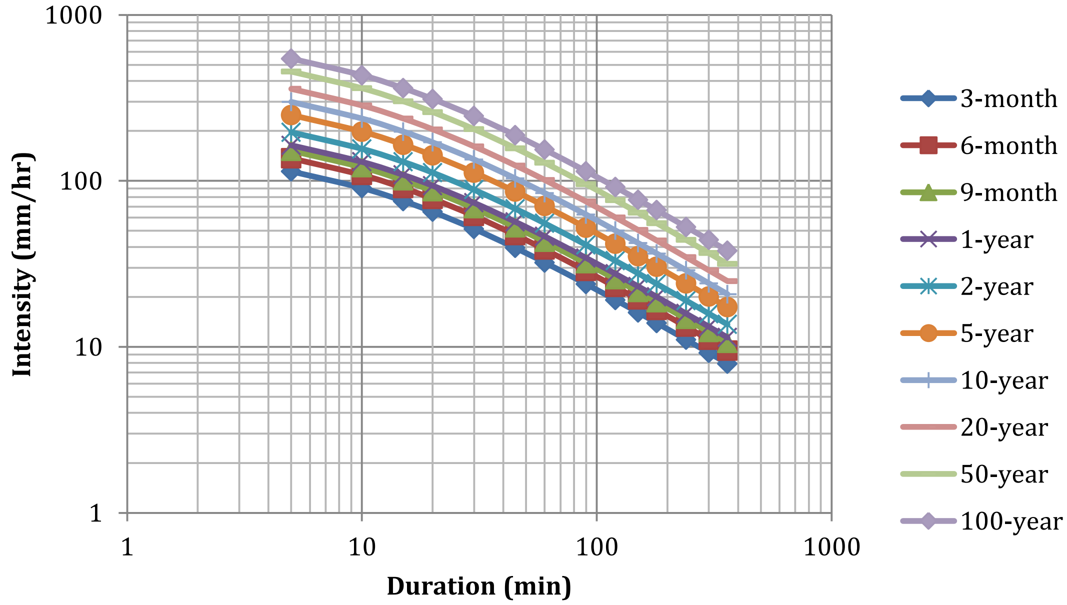


Figure E31. IDF curve based on GEV/AMS model for station 3118102

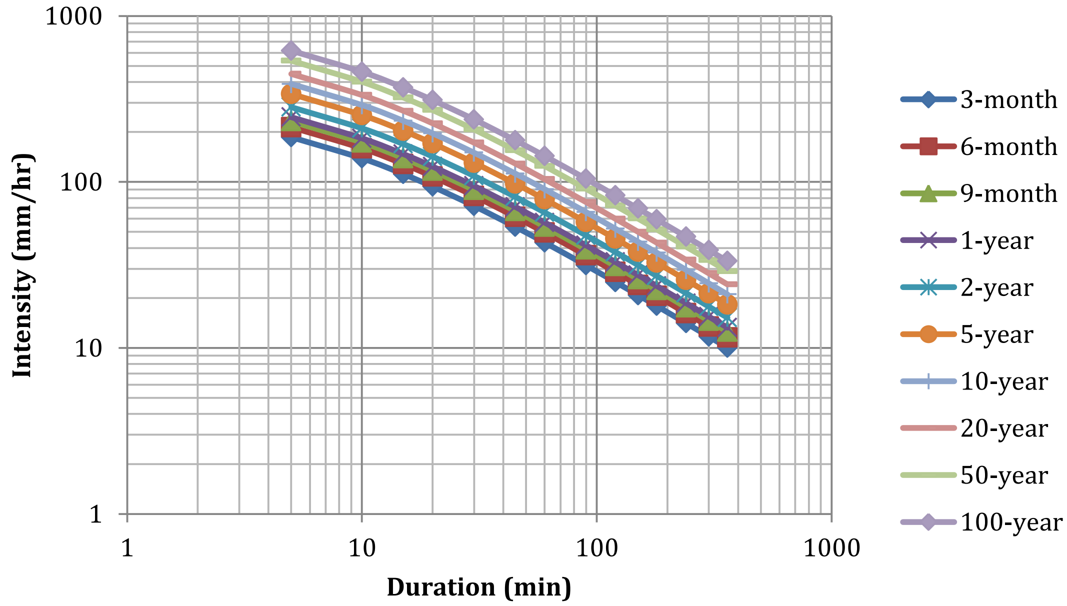


Figure E32. IDF curve based on GPA/PDS model for station 3118102

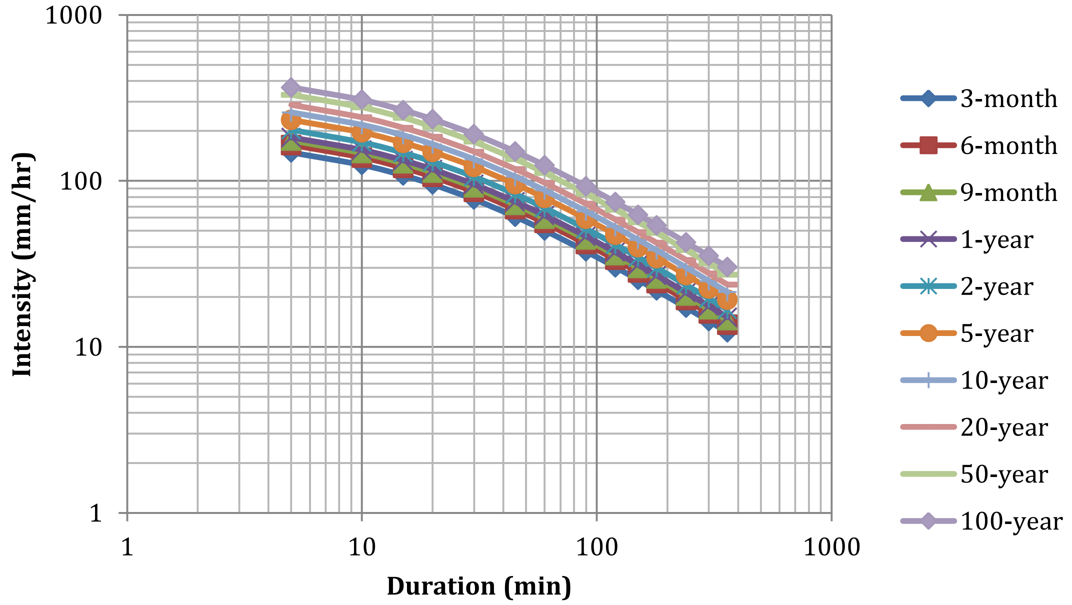


Figure E33. IDF curve based on GEV/AMS model for station 3613004

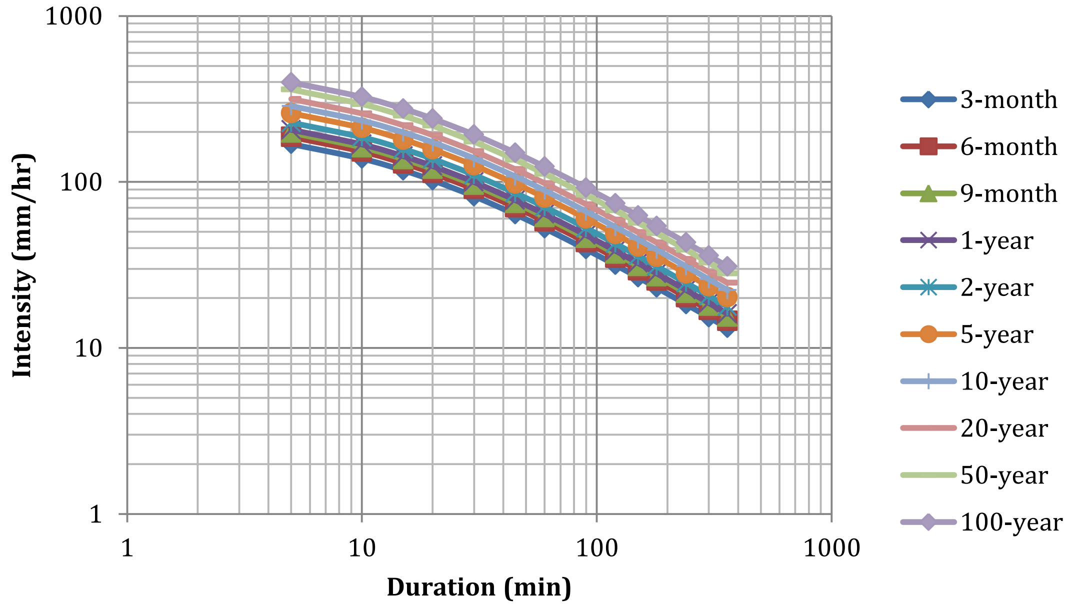


Figure E34. IDF curve based on GPA/PDS model for station 3613004

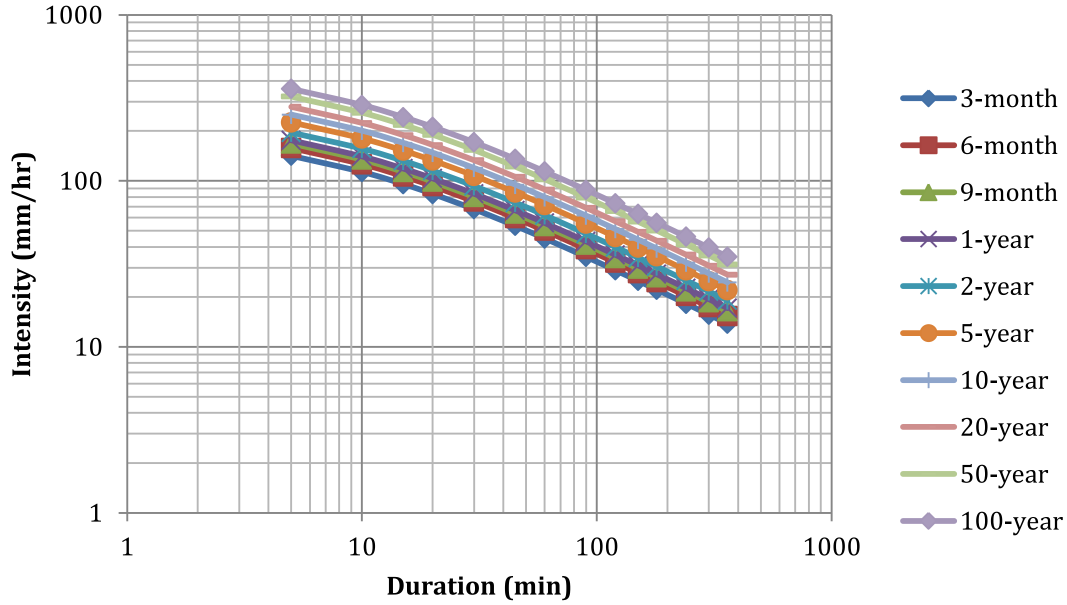


Figure E35. IDF curve based on GEV/AMS model for station 4232002

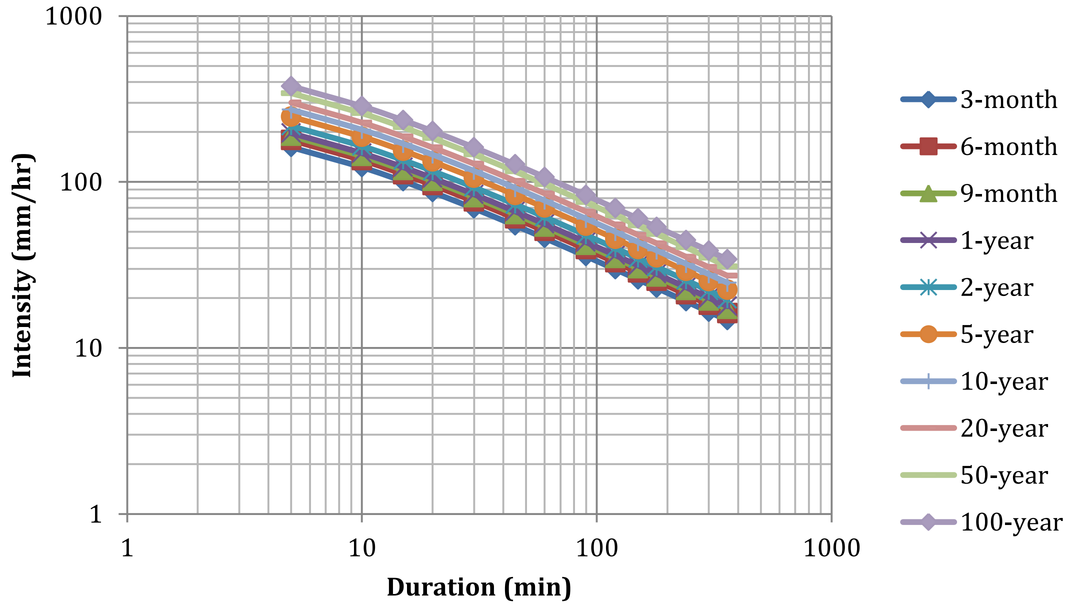


Figure E36. IDF curve based on GPA/PDS model for station 4232002

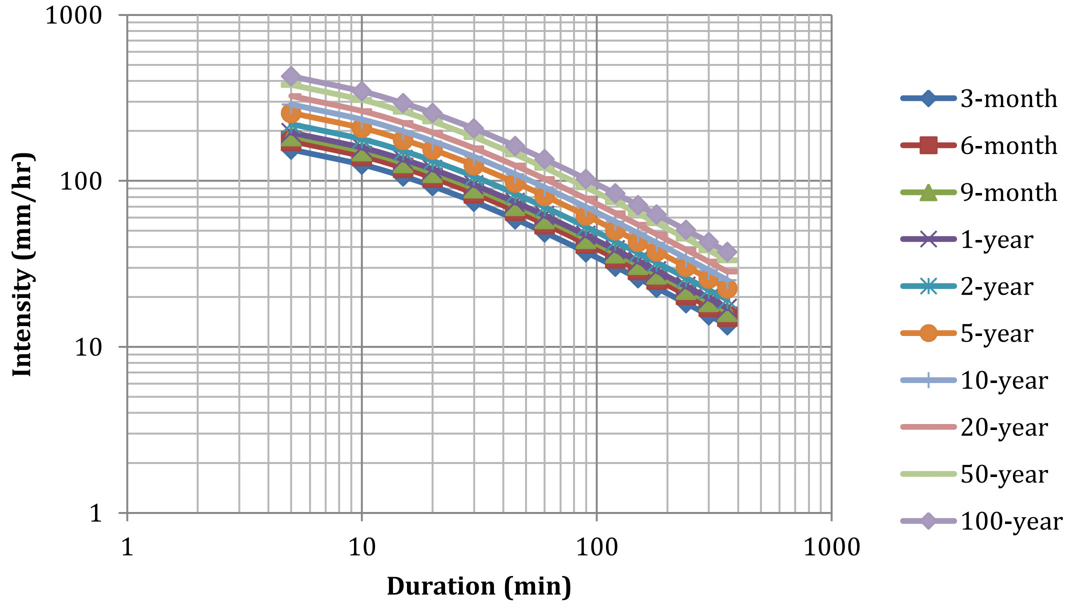


Figure E37. IDF curve based on GEV/AMS model for station 3116003

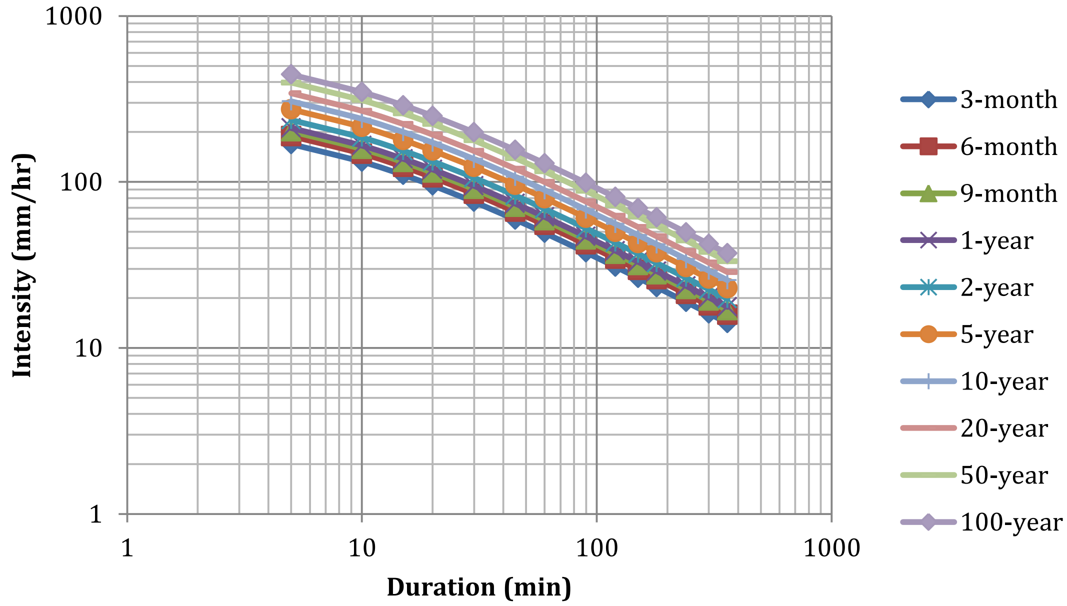


Figure E38. IDF curve based on GPA/PDS model for station 3116003

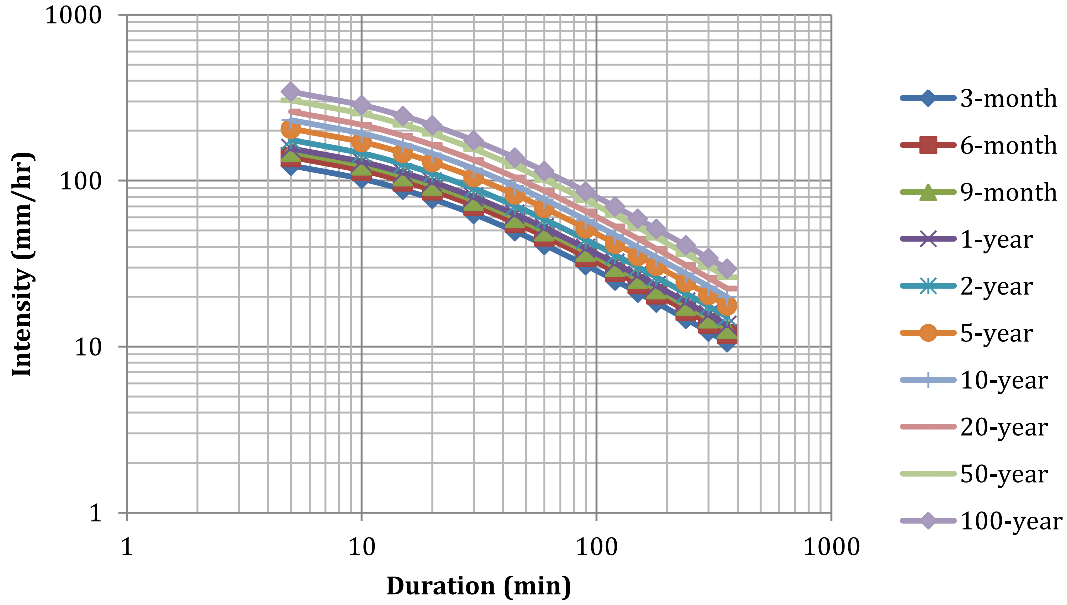


Figure E39. IDF curve based on GEV/AMS model for station 3217002

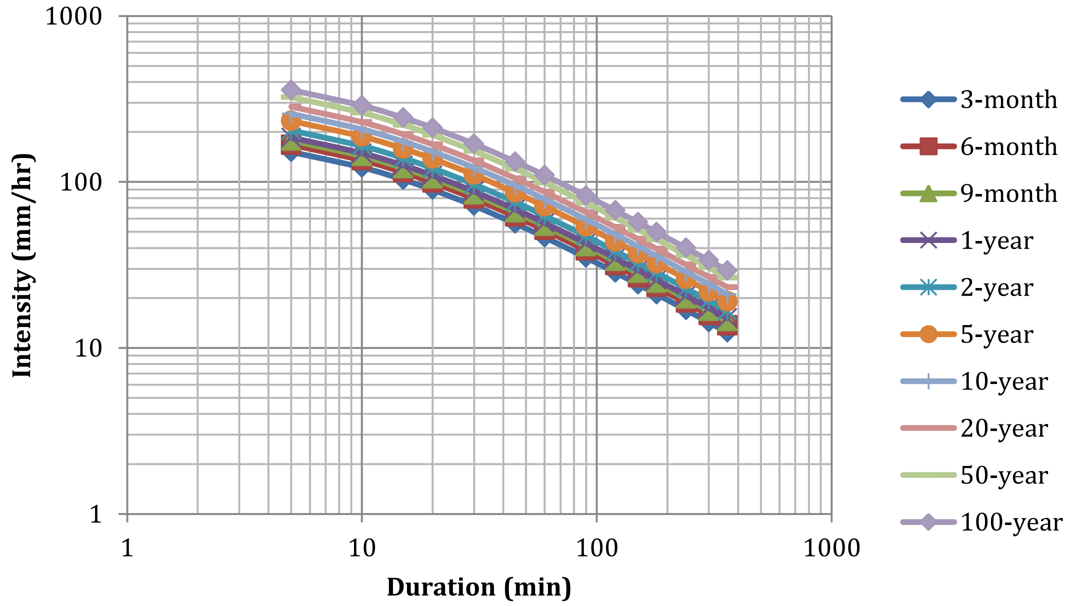


Figure E40. IDF curve based on GPA/PDS model for station 3217002