

University of Southern Queensland
FACULTY OF HEALTH, ENGINEERING AND SCIENCES
SCHOOL OF CIVIL ENGINEERING AND SURVEYING

**Power Law flood frequency analysis of selected
Queensland stream gauges**

A dissertation submitted by

Denika Moes

in fulfilment of the requirements of

ENG4111/4112 Engineering Research Project

towards the degree of

Bachelor of Engineering: Civil

Submitted: 29th October 2015

Abstract

Flood Frequency Analysis is key to the prediction of frequency and magnitude of extreme flood events in given catchments. Floods can cause millions of dollars damage to communities and since it is important to have adequate flood protection measures, it is desirable to obtain accurate estimates of flood recurrence.

The aim of this project is to investigate the suitability of the Power Law frequency model as a more accurate way of the peak discharges of flood events. Using two goodness of fit tests, the Chi-Squared test and the R-Squared test, the Power Law relationship has been tested against the more conventional methods used in Australian, the Log Pearson type 3 distribution and the exponential distribution for ten stream gauge stations located throughout Queensland.

The results of the analysis of the ten stream gauges found that generally the Log Pearson type 3 distribution was more accurate in predicting the peak discharges of the observed historical flows for sites of which the floods are expected to occur in intervals greater than ten years. The Power Law frequency model however produces a more conservative estimate for the return period of the larger floods and hence increasing the estimated likelihood of severe floods.

Therefore the use of the Power Law relationship as procedure for flood frequency analysis for the extreme events would create more conservative infrastructure designs and land use restrictions.

University of Southern Queensland

Faculty of Health, Engineering and Sciences

ENG4111/ENG4112 Research Project

Limitations of Use

The Council of the University of Southern Queensland, its Faculty of Health, Engineering & Sciences, and the staff of the University of Southern Queensland, do not accept any responsibility for the truth, accuracy or completeness of material contained within or associated with this dissertation.

Persons using all or any part of this material do so at their own risk, and not at the risk of the Council of the University of Southern Queensland, its Faculty of Health, Engineering & Sciences or the staff of the University of Southern Queensland.

This dissertation reports an educational exercise and has no purpose or validity beyond this exercise. The sole purpose of the course pair entitled “Research Project” is to contribute to the overall education within the student’s chosen degree program. This document, the associated hardware, software, drawings, and other material set out in the associated appendices should not be used for any other purpose: if they are so used, it is entirely at the risk of the user.

University of Southern Queensland

Faculty of Health, Engineering and Sciences

ENG4111/ENG4112 Research Project

Certification of Dissertation

I certify that the ideas, designs and experimental work, results, analyses and conclusions set out in this dissertation are entirely my own effort, except where otherwise indicated and acknowledged.

I further certify that the work is original and has not been previously submitted for assessment in any other course or institution, except where specifically stated.

Denika Moes

0061019261

Acknowledgements

This project was carried out under the supervision of Dr. Ian Brodie of the University of Southern Queensland. Without his supervision and guidance, this project would not have been possible

I would like to thank my family and friends, who have been there for me through every step and supported me in everything I have done and also the residents of Concannon College for their encouragement and constant interest throughout the process.

DENIKA MOES

University of Southern Queensland

October 2015

Table of Contents

Abstract	i
Acknowledgements	iv
Table of Contents	v
List of Figures	ix
List of Tables	xi
Chapter 1 – Introduction	1
1.1 Introduction	1
1.2 Background	1
1.3 Project Aim.....	4
1.4 Project Objectives.....	4
1.5 Justification	5
1.6 Consequential Effects.....	6
1.6.1 Safety.....	6
1.6.2 Cost	7
1.7 Overview of Dissertation.....	8
Chapter 2 – Literature Review	10
2.1 Chapter Overview.....	10
2.2 Types of Data Series.....	10
2.2.1 Annual Flood Series	11

2.2.2	Partial Series.....	13
2.3	Probability Distributions	16
2.3.1	Log Pearson type 3 Distribution	20
2.4	Power Law Flood Frequency Analysis.....	23
2.5	Goodness of Fit	25
2.5.1	Chi-Squared Test, x^2	27
2.5.2	R-squared test, R^2	27
2.6	Poor Fits	29
2.7	Summary	30
Chapter 3 – Methodology		31
3.1	Chapter Overview.....	31
3.2	Resource Analysis	31
3.3	Station Selection.....	32
3.4	Annual Series and Log Pearson 3 Distribution	35
3.4.1	Outliers.....	38
3.5	Partial Series and Exponential Distribution	39
3.6	Power Law Relationship	41
3.7	Applying Goodness of Fit tests	42
3.7.1	Chi-Squared	42
3.7.2	R-Squared.....	44
3.7.3	Graphical Comparison	45
3.8	Summary	46

Chapter 4 – Results	47
4.1 Chapter Overview.....	47
4.2 Stream Gauge Results	47
4.2.1 Coen River at Racecourse (922101B).....	48
4.2.2 Fisher Creek at Nerada (112002A)	53
4.2.3 Herbert River at Abergowrie (116006B)	57
4.2.4 Carmila Creek at Carmila (126003A).....	61
4.2.5 Isis River at Bruce Highway (137201A).....	65
4.2.6 Mary River at Miva (138001A)	69
4.2.7 Caboolture River at Upper Caboolture (142001A).....	73
4.2.8 Stanley River at Peachester (143303A)	77
4.2.9 Coomera River at Army Camp (146010A).....	81
4.2.10 Condamine River at Elbow Valley (422394A).....	84
4.3 Summary	89
Chapter 5 – Discussion	90
5.1 Chapter Overview.....	90
5.2 Summary of Results	91
5.3 Exponential Distribution	93
5.4 Log Pearson 3 Distribution.....	94
5.5 Power Law Model	96
5.6 Limitations.....	98
5.7 Summary	102

Chapter 6 – Conclusion	103
6.1 Chapter Overview.....	103
6.2 Project Conclusions	103
6.3 Further Work	105
List of References	107
List of Appendices	112
Appendix A Project Specification	113
Appendix B Frequency Factors K_T for Log Pearson 3 Distribution	114
Appendix C Tables used for identifying high and low outliers	116
Appendix D Chi-Square Distribution Table.....	118

List of Figures

Figure 2.3.1 Typical Flood Frequency Curve (Sivandran 2002)	20
Figure 2.5.1 Examples of different values of R^2 (GraphPad Software n.d.).....	28
Figure 3.3.1 The location of the ten stream gauge stations used in this analysis.....	34
Figure 4.2.1 Coen River stream gauge model predicted peak discharges at Racecourse fitted to the average recurrence intervals	50
Figure 4.2.2 Coen River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data	52
Figure 4.2.3 Fisher Creek stream gauge model predicted peak discharges at Nerada fitted to the average recurrence intervals	54
Figure 4.2.4 Fisher Creek stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data.....	56
Figure 4.2.5 Herbert River stream gauge model predicted peak discharges at Abergowrie fitted to the average recurrence intervals	58
Figure 4.2.6 Herbert River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data.....	60
Figure 4.2.7 Carmila Creek stream gauge model predicted peak discharges at Carmila fitted to the average recurrence intervals	62
Figure 4.2.8 Carmila Creek stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data.....	64
Figure 4.2.9 Isis River stream gauge model predicted peak discharges at the Bruce Highway fitted to the average recurrence intervals.....	66
Figure 4.2.10 Isis River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data	68

Figure 4.2.11 Mary River stream gauge model predicted peak discharges at Miva fitted to the average recurrence intervals.....	70
Figure 4.2.12 Mary River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data.....	72
Figure 4.2.13 Caboolture River stream gauge model predicted peak discharges at Upper Caboolture fitted to the average recurrence intervals	74
Figure 4.2.14 Caboolture River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data.....	76
Figure 4.2.15 Stanley River stream gauge model predicted peak discharges at Peachester fitted to the average recurrence intervals	78
Figure 4.2.16 Stanley River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data.....	80
Figure 4.2.17 Coomera River stream gauge model predicted peak discharges at Army Camp fitted to the average recurrence intervals.....	82
Figure 4.2.18 Coomera River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data.....	84
Figure 4.2.19 Condamine River stream gauge model predicted peak discharges at Elbow Valley fitted to the average recurrence intervals	86
Figure 4.2.20 Condamine River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data.....	88
Figure 5.5.1 Graphical comparison of the Power Law model and the Log Pearson 3 distribution discharge predictions for the largest event compared to the observed largest event, for the ten stream gauge stations over Queensland.	97

List of Tables

Table 2.3.1 Plotting Position Formulas (Shabri 2002).....	18
Table 3.3.1 Summary of the final stream gauge stations used in this analysis	33
Table 4.2.1 Coen River results of Chi-Squared test	51
Table 4.2.2 Coen River analysis of R-Squared values	51
Table 4.2.3 Fisher Creek results of Chi-Squared test.....	55
Table 4.2.4 Fisher Creek analysis of R-Squared values.....	55
Table 4.2.5 Herbert River results of Chi-Squared test	59
Table 4.2.6 Herbert River analysis of R-Squared values	59
Table 4.2.7 Carmila Creek results of Chi-Squared test.....	63
Table 4.2.8 Carmila Creek analysis of R-Squared values.....	63
Table 4.2.9 Isis River results of Chi-Squared test.....	67
Table 4.2.10 Isis River analysis of R-Squared values	67
Table 4.2.11 Mary River results of Chi-Squared test.....	71
Table 4.2.12 Mary River analysis of R-Squared values.....	71
Table 4.2.13 Caboolture River results of Chi-Squared test.....	75
Table 4.2.14 Caboolture River analysis of R-Squared values.....	75
Table 4.2.15 Stanley River results of Chi-Squared test	79
Table 4.2.16 Stanley River analysis of R-Squared values.....	79
Table 4.2.17 Coomera River results of Chi-Squared test.....	83
Table 4.2.18 Coomera River analysis of R-Squared values.....	83
Table 4.2.19 Condamine River results of Chi-Squared test	87
Table 4.2.20 Condamine River analysis of R-Squared values	87

Table 5.2.1 Best fitting model of each stream gauge station found from the goodness of fit tests	92
Table 5.6.1 The best fitting distribution for the annual and partial flood data series according to the Chi-Squared test performed using Mathwave's Easyfit software	101

Chapter 1 – Introduction

1.1 Introduction

This report outlines the background, objectives, methodologies, results and conclusions pertaining to the investigation of flood frequency analysis models. In particular, identifying the effectiveness of a simple Power Law relationship model for predicting the peak discharges of selected floods opposed to the current conventional in use today. This investigation is based on historical flood data collected from ten stream gauge stations across Queensland. This project aims to supplement previous research findings, as well as provide engineering guidance on the use of the Power Law relationship for the predictions of peak flood discharges.

1.2 Background

Floods are one of the worst natural disasters and each year can cause millions of dollars of damage and loss of human lives. A flood is considered to be an unusually high stage of the river. Surface runoff invariably produces a stream rise, but it does not necessarily cause a flood, the difference being in magnitude only. It is important to differentiate rigidly between surface runoff and a flood, a flood is commonly defined as being an

unusually or abnormally high stage of the river or when at the stage at which the stream channel becomes filled and overflows its banks and inundates the adjacent lands. Although it is true that the latter condition usually accompanies floods, it is not an essential characteristic as streams flowing through deep ravines, gorges or canyons would never be subject to floods. Streams are commonly recognised as being in flood when their stage is unusually high.

The study of floods and flood flows demonstrate there is a succession of floods of ever-varying size. There is a flood that is expected to be exceeded every year, while there is also a greater flood that may be expected to occur, on an average, once in ten years, not at regular ten year intervals, but say ten times in a century. A great flood again may be expected as often as once in a century and there is assumed to be floods that occur only at intervals of several centuries.

If a flood of a certain magnitude occurs approximately once in a hundred years, there is a one percent chance or a one chance in a hundred that that flood will occur during a year. A flood with a magnitude that is likely to be exceeded on average of once every twenty-five years is a four percent chance flood.

Floods have very low frequency or probability of occurrence are large, catastrophic floods whereas smaller floods will occur more often. The larger recurrence interval is the less of a chance there is for experiencing that flood in a particular year. However, the probability of occurrence cannot be zero, a very large, uncommon flood could occur every few years.

This recurrence interval gives a means of expressing the likelihood of in specified number of years, a certain magnitude of flood will be exceeded and hence is vital in flood control, emergency planning, land-use regulation and insurance considerations.

To avoid destruction, dams need to have sufficient spillway capacity and protection, bridges must have a waterway opening, flood walls and embankments that are high enough so they will not be overtopped and reservoirs must have the required capacity. The maximum flood these structures can safely accommodate for is called the design flood. The most direct way of estimating this design flood is to use a process called flood frequency analysis.

Historical records of past river heights and flows are generally used in the estimation of the largest floods that could occur in a given time period. These historical records only cover a very short amount of time and according to Malamud and Turcotte (2006), there is no general basis for extrapolation.

There is a wide range of statistical distributions used to analysis the frequency of floods and in current practise, the standard approach is to use the maximum annual flood for each year of flood data and obtain the best fit for the chosen statistical distribution. A Power Law relationship has been suggested by several authors as a better estimation of the flood hazard and this project will investigate the Power Law relationship for flood frequency analysis.

1.3 Project Aim

The main purpose of flood frequency analysis is the prediction of the frequency and magnitude of extreme events in a given catchment. Conventional analysis methods have been criticized for its questionable theoretical basis and failure in prediction of extreme flood events.

The aim of this research is to investigate the suitability of the Power Law statistical model as a preferred model to estimate flood frequency. To accurately confirm the suitability of the Power Law model, an analysis will be undertaken comparing the Power Law model against more conventional methods currently in use. Since no theoretical distribution can be considered to adequately fit the stream flow data of all the streams, selected streams will be considered individually for their suitability with the Power Law model.

1.4 Project Objectives

The project aim was reviewed and split into a number of objectives for completion.

- To research existing literature on the limitations and drawbacks of the current flood frequency analysis methods, the power law frequency model as well as background knowledge necessary to compare the different methods.

- Construct the annual and partial flood series from the peak discharges data obtained from the Queensland Department of Natural Resources and Mine's Water Monitoring website for ten unregulated gauging stations.
- Apply a Log Pearson type 3 distribution to the annual flood data and a negative exponential distribution to the partial series data as identified as conventional flood frequency distributions by the Australian Rainfall and Runoff (1987).
- Apply the power law distribution model to the partial flood data from each station
- Compare the peak discharges at similar average recurrence intervals (ARIs) as well as the overall fit of each distribution model for each station location

1.5 Justification

It is important to the civic society for accurate and reliable flood magnitude predictions, especially since local and regional communities have to make independent judgments regarding the actions that are required to prevent and manage natural disasters. Middelmann et al. (2000) estimates that more than 80% of the buildings that are at risk of flooding in Australia are located in New South Wales and Queensland and Queensland has the highest average annual damages from floods. Due to the importance for Queensland to have adequate flood protection measures, it is desirable to obtain accurate estimates of flood recurrence.

If a Power Law model is the underlying model for the flood behaviour of the Queensland rivers, then the industry standard 100-year estimated flood discharge will require revision and planning, making the design decisions more accurate and probably more conservative.

1.6 Consequential Effects

The overall improvement of the flood frequency analysis procedures used in identifying the probability model of flood peaks in some of Queensland's catchments is the main effect of this project. This will be achieved through investigating the accuracy of the Power Law relationship in regards to floods and comparing it with conventional analysis techniques.

1.6.1 Safety

The Power Law model has been proposed as a more conservative model in terms of estimating the peak discharges of the more extreme floods. This project could therefore improve the industry standard one hundred year flood, which is the basis for many design and planning decisions. The floodplain map for each community can be review for the new peak discharges and in turn reduce the impact of flooding. Emergency service organisations will also have a better understanding of the scale of flood risk and the logistical and access problems that may exist. Therefore in determining the more accurate model in predicting the frequencies of peak discharges of floods will help to

lessen the flood hazard where it is economic and socially acceptable, and reduce the devastating impacts on the communities.

1.6.2 Cost

Having an accurate flood frequency analysis procedure will reduce the costs associated with the construction and maintenance of infrastructure in the long term. Reviewing the floodplain maps to identify the flood areas of the 100-year flood with the Power Law model will better help the communities to reduce the frequency of damage as residents know to protect themselves and the recovery costs associated with flooding.

Flooding can affect everyone, through direct water damage through to disruptions to transport services, communication and power. Studies by the National Flood Risk Advisory Group (2008) show a household believes they cannot recover from a financial shock of more than \$10,000 with their own capital (National Flood Risk Advisory Group 2008). A minor flood that is just over the floor is likely to cause damages greater than \$10,000 and making minor flood events financially devastating. The floodplain map created using the discharges from the Power Law model will help with the awareness and readiness of the community and hence reduce the cost of repairs to the infrastructure. There are specifications set out in the Building Code of Australia and in the relevant Standards that give guidance on use of materials to reduce flood damages in new developments and in renovations. People who know they are in the floodplains can use these codes to reduce the impact of flood debris and to maintain structural integrity after a flood event, therefore reducing the long term cost of the infrastructure (National Flood Risk Advisory Group 2008).

1.7 Overview of Dissertation

This project dissertation is organised as follows:

Chapter 2 Literature Review

This chapter investigates the relevant literature on the topic of flood frequency analysis to provide an informative briefing on the subject. It presents current findings and opinions of professionals as a result of their investigative work on the more conventional flood frequency as well as the Power Law model.

Chapter 3 Methodology

This chapter discusses the methodologies undertaken in order for project completion. In particular, the criteria used to select the stream gauge stations and the process involved in determining the distribution models for each of these stations.

Chapter 4 Results

This section provides the output data from the methodology for each individual stream gauge station. It also discusses the accuracy of each distribution in terms of goodness of fit to the historical stream flows for those stations.

Chapter 5 Discussion

Using the results for each station identified in the previous chapter, this chapter discusses prevalent trends between the distribution models and identifies the effectiveness of each flood frequency model tested. Limitations of this study have also been identified in this chapter.

Chapter 6 Conclusions

A summary of the conclusions of the project are presented in this chapter, along with further work and recommendations.

Chapter 2 – Literature Review

2.1 Chapter Overview

A literature review of the topics relevant to this project is presented here. While this study was largely a statistical analysis, the literature review was necessary to gain background knowledge on a number of key topics. The chapter examines the basic concepts of flood frequency analysis, in particular the types of time series and current probability distributions recommended by the Australian Rainfall and Runoff (1987) guide. It also looks at the work conducted in the past to applying the Power Law to flood frequency analysis and methods of assessing the fit of a probability distribution models to the historic flood data.

2.2 Types of Data Series

With starting with just a basic time step, a range of types of time series that can be found that can be used for flood frequency analysis (Meylan, Favre & Musy 2012).

These include:

- A complete series
- Annual flood duration series

- Partial flood duration series (or Peaks Over Threshold)

The Australian Rainfall and Runoff (1987) specifies which series is to be used when selecting floods that are to be used in the frequency analysis in Australian conditions. The annual series is preferred when the average recurrence interval of flood discharge is greater than or equal to 10 years, which is generally used in design, as the higher recurrence intervals are used for determining the design flood for infrastructure at a particular location. The partial series is preferred when all floods are less than 10-year floods and is used for flows of low recurrence intervals, particularly in urban storm water environments. However, in conventional flood frequency analysis, it is common practice to apply both methods in order to determine the difference that data choice decisions make for prediction (Australian Rainfall and Runoff 1987). Therefore this investigation will use the annual flood series and the partial flood series in the analysis.

2.2.1 Annual Flood Series

The annual flood series is the more commonly used time series and is created using the largest discharge in each water year. The discharges in Australia are highly seasonal, therefore a water year is preferred to be used over a calendar year. (Australian Rainfall and Runoff 2010) A water year commences when the average discharge is lowest during the year, which for Queensland is defined as a 12 month period from July 1, through to June 30, of the following year (Kollmorgen et al. 2007). The highest flow in each water year is selected for the flood series, with all other smaller floods ignored.

According to the review of the Australian Rainfall and Runoff (2006), there are three major advantages in using the annual series for flood frequency analysis, these were expressed as:

- The flood peaks are independent of each other, as the peaks are generally separated by significant time intervals
- The annual series can be easily and unambiguously taken from Government websites, that are widely available
- Frequency distributions generally follow available theoretical distributions

A major drawback is that it could exclude floods from each year that are significant, especially if several large floods have occurred during the same water year. According to Armstrong, Collins and Snyder (2012), the annual series can include some small annual floods, causing small floods to occur more often than indicated by the annual series (Armstrong, Collins & Snyder 2012).

Keast and Ellison (2013) identify during a study involving the Northern Tasmanian stream gauging station data found that the annual series estimates at an average recurrence interval of 1.1 years were one third of the magnitude provided by the partial series estimations, demonstrating the annual flood series significantly underestimates the magnitude of the low scale floods. Keast and Ellison (2013) recommend the annual series is not to be used for recurrence intervals that are smaller than five years and instead use the partial series (Keast & Ellison 2013).

2.2.2 Partial Series

It is possible for the second or third largest peak in a particular year be greater than the maximum flood for another year. In an annual series, such additional events are ignored since only the largest annual event is allowed. The partial series model (also denoted the Peak Over Threshold model) is created with all the floods that have a peak discharges which is above a selected value, irrespective of the number of other floods that have occurred during the year.

Linsley et al. (1982) states that since the partial series is arbitrarily selected, it cannot be expected to fit a standard distribution (Linsley, Kohler & Paulhaus 1982). The Australian Rainfall and Runoff (2006) suggests a graphical interpolation of the historical flood data is sufficiently accurate to determine the distribution of flood recurrence of less than ten years. (Australian Rainfall and Runoff 2010). Inferences made for events greater than 10 years should be fitted with a probability distribution. The Generalized Pareto distribution has also been commonly used in flood frequency studies (Rosbjerg, Madsen & Rasmussen 1992); however, this analysis on the Queensland stream gauges only explored the graphical interpolation method.

In creating a partial flood series a selected base discharge is chosen and the discharges exceeding this base value are classified as floods. The number of floods (K) is different to the number of years of record (N) and depends on the base discharge. Rustomji, Bennett and Chiew (2009) determined that having more floods in the partial series is advantageous. Whereas a study by Keast (2013) on stream gauges in Northern Tasmania found the chosen base discharge had little effect on the partial series estimates at these low recurrence intervals (Keast & Ellison 2013). However, the greater number of small floods included, the distribution tends to match with the annual flood series.

Small events are excluded in the partial series as the selected base value is sufficiently high enough to exclude them, whereas the annual series includes the non-floods from the dry years influencing the shape of the distribution. In Australia, the range of flows and the time between floods is greater in there than in countries like the United Kingdom or the United States (Grayson et al. 1996). It is expected that in Australia, the ratio of floods, K , to years of record, N , be lower than the United States or the United Kingdom.

As recommended by the American Society of Civil Engineers (1949), the base discharge is chosen so that the number of floods is greater than the number of years of record, however, there also needs to be no more than four floods used in a water year (Meylan, Favre & Musy 2012). The United States Geological Survey (Dalrymple, 1960) on the other hand, recommended there should be three times the number of floods as years of record ($K = 3N$).

The American Society of Civil Engineers (1949) suggests the number of flood each year and therefore the base discharge should depend on the distribution being used. For fitting the Log Pearson 3 distribution, studies conducted by McDermott and Pilgrim (1982) and also by Jayasuriya and Mein (1985) determined the distribution models fitted more effectively when the number of floods equalled the number of years of record ($K = N$).

A single flood may have multiple peaks and therefore a minimum amount of time is needed for the river's discharge to be considered a separate flood peak. The partial flood series therefore may have floods that are not independent events and instead be all one flood. Some floods can be short lived and only inundation properties, whereas other floods cause the destruction of annual crops and can last as long as a year (Baker 1994).

Flood damage is caused by the highest flood; the secondary peaks therefore should not be included in the flood frequency analysis. Therefore is no specific criterion in determining when flood peaks are independent of each other. Malamud et al. (1996) found that using a time interval between 7 and 60 days between successive peaks gave reasonably robust flood frequency estimation (Malamud, Turcotte & Barton 1996). Some of the criteria used in past studies have included:

- According to a Uniting Kingdom Flood Studies Report, flood peaks are to be separated by three times the time taken for the flow to peak and the flow have decreased to two-thirds of the original peak between the two peaks. (Natural Environment Research Council 1975)
- Studies conducted by Pilgrim and McDermott (1982) and McDermott and Pilgrim (1983) used the monthly maximum peaks for a small to medium sized catchment. Basing this assumption on the fact that only minor additional damage would be caused by floods that have occurring within the same month and closer together flood peaks could not be classed as independent in terms of their effects. (Australian Rainfall and Runoff 1987)
- Malamud and Turcotte (2006) specified that each peak must be separated by at least 30 days either side of the peak to maintain independence (Malamud & Turcotte 2006).
- Kundzewicz et al. (2005) used a time period to determine the independent flood peaks based on the size of the catchment:
 - 5 days for the size of catchments less than 45,000 km²
 - 10 days for the size of catchments between 45,000 - 100,000 km²
 - 20 days for the size of catchments greater than 100,000 km²

The time period needs to be as short as possible while only including ‘statistically independent’ floods. While the flood series created for this project is not strictly uncorrelated, for the purposes of this analysis they are sufficiently independent.

2.3 Probability Distributions

The most commonly used method in the determination of the frequency at which peak stream flow occurs is statistical flood frequency analysis. This method involves fitting historical flood records to an extreme value probability distribution function and therefore relies on having long stream flow records.

An article by Ott and Linsley (1972) concluded that for flood frequency analysis to be performed accurately there needs to be a sufficiently long stream flow record, otherwise there is a high probability of uncertainty of the fitting of the distribution to the historical data. “Only extremely long-term records divulge the true frequency characteristic of a watershed. Moreover, there is a high probability (as much as 80 percent) that the flood peaks will be over-estimated when using the stream flow record as short as 20 years.” (Ott & Linsley 1972). Determining the average frequency of a once in 100-year flood when there is only 50 years of records available, the probable error is very high and suggested by Wisler and Brater (1959) be several hundred percent error because the not even one whole sample or period of observation is available upon which to base the judgement on. The probable error of the distribution is dependent on the number of

independent samples of data that is available and no amount of juggling or manipulation to the data can reduce the error (Wisler & Brater 1959).

The commonly used probability distributions for flood frequency analysis can be divided into four distinct groups (Malamud, Turcotte & Barton 1996):

- Generalized Pareto distribution
- General Extreme Value (GEV) family
 - includes the GEV, Weibull, Gumbel and Log Gumbel
- Normal family
 - includes the Normal, Log Normal, Log Normal type 3
- Pearson type 3 family
 - includes Pearson type 3, Log Pearson type 3

The federally adopted methodology in the United States is to fit an annual flood series to a Log Pearson type 3 distribution (USGS 1982). For Australia, it is suggested using the Generalized Extreme Value (GEV) and the Log Pearson type 3 distributions (Australian Rainfall and Runoff 2010).

Constructing an empirical distribution function or the probability plot is important for flood frequency estimation, as the probability plot estimates the average exceedence probability and plots it in regards to the observed peak discharges. It is then easier to draw the probability distribution as a smooth curve to allow for visually checking of the effectiveness of the proposed fitted flood distribution.

For plotting purposes, a general formula (shown below) is used to estimate the annual exceedence probability of an observed flood.

$$P_{(m)} = \frac{m - \alpha}{n + 1 - 2\alpha}$$

where,

n = the number of years of flood data

m = the rank of the flood

α = selected constant

There has been put forward several choices by a range of researchers for the constant α , which is chosen to maintain desirable statistical properties in the plotting position.

Various values of the constant α are summarized in Table 2.3.1.

Table 2.3.1 Plotting Position Formulas (Shabri 2002)

Proponent	α	Plotting Position Formula	Parent Distribution
Weibull (1939)	0	$\frac{m}{n + 1}$	All distributions
Beard (1943)	0.3175	$\frac{m - 0.3175}{n + 0.365}$	All distributions
APL	0.35	$\frac{m - 0.35}{n}$	Used with Probability Weighted Moment Method (PWM)
Blom (1958)	0.375	$\frac{m - 3/8}{n + 1/4}$	Normal distributions
Cunnane (1977)	0.40	$\frac{m - 0.4}{n + 0.2}$	General Extreme Value and Pearson type 3 distributions
Gringorten (1963)	0.44	$\frac{m - 0.44}{n + 0.12}$	Exponential, Extreme Value and General Extreme Value distributions
Hazen (1914)	0.50	$\frac{m - 0.5}{n}$	Extreme Value distributions

The plotting position of top ranked flood events is sensitive to the choice of plotting formula. According to Shabri (2002), these plotting position formula are design to correct bias and return a systematic underestimation of the recurrence interval of the top ranked floods. To maintain consistency, the Australian Rainfall and Runoff guide (2006) recommends that the Cunnane plotting position be used as it produces plotting positions that yield unbiased quantiles.

(2.3-2)

$$P_{(m)} = \frac{m - 0.4}{n + 0.2}$$

The probability plot for the partial flood series is prepared similar to the annual flood series. However, it involves estimating of the average recurrence interval for each historical flood instead of the annual exceedence probability and plotting that against the observed historical discharges, with the average recurrence interval of each flood calculated by:

(2.3-3)

$$ARI_{(m)} = \frac{n + 0.2}{m - 0.4}$$

where,

n = Number of years of flood data

m = Rank of flood (sorted in descending order)

The flood frequency curve is completed when a suitable distribution is fitted to the historical flood data (Figure 2.3.1). The distribution's curve correlates the return period (as shown on the x axis), to a specified flood magnitude (as shown on the y axis). Using

this curve, the peak discharge of a 'design flood' in relation to the designed lifespan of the proposed infrastructure can be found.

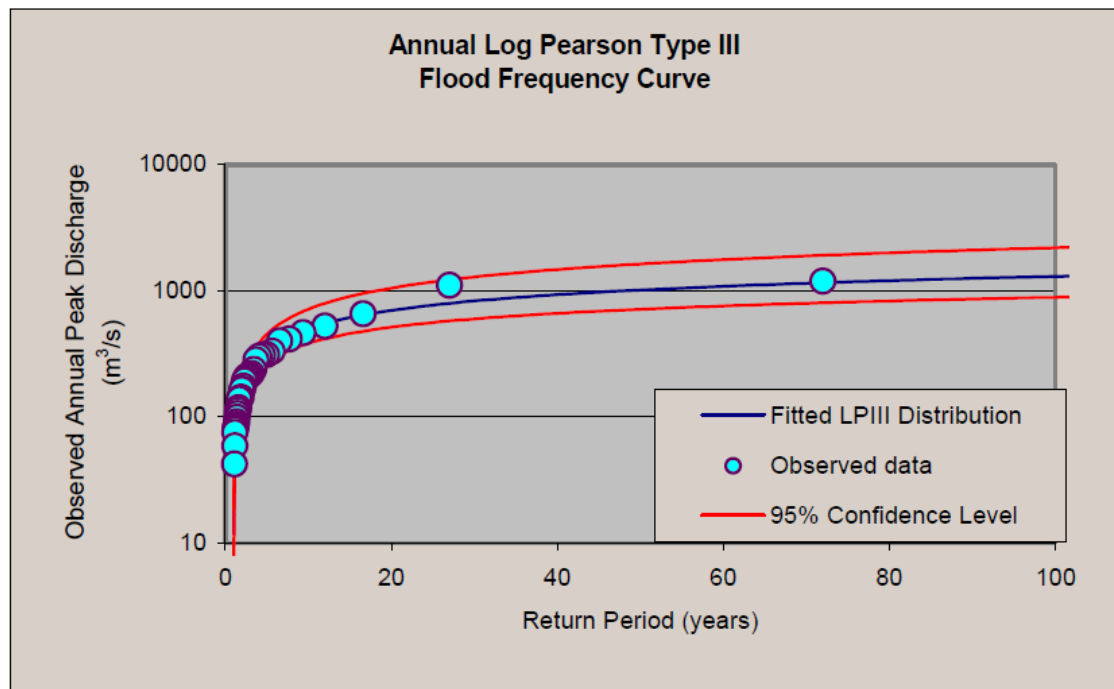


Figure 2.3.1 Typical Flood Frequency Curve (Sivandran 2002)

2.3.1 Log Pearson type 3 Distribution

The most common applied frequency distribution is the Log Pearson 3 distribution as it is recommended by the United States Water Resource Council (1982) for flood peak analysis (USGS 1982).

Studies conducted by Conway for New South Wales coastal streams and Kolittke et al. for Queensland streams established the Log Pearson 3 distribution as the most suitable distribution for their catchments (Rahman, Haddad & Rahman 2014). Based on findings from these studies, it was recommended in the Australian Rainfall and Runoff (1987) guide that flood frequency analysis in Australia should follow the United States of America and adopt the use of the Log Pearson type 3 distribution.

A study done by Boughton (1975) of the statistical frequency distribution of annual maximum flows in Queensland showed that of the distributions tested (Log Pearson type 3, Pearson 3, log-Normal, Gumbel and Potter), Pearson 3 and Log Pearson 3 distributions fitted the Queensland data most accurately due to their ability to change depending on the skew of the data (Boughton 1975).

This distribution is a three parameter distribution, meaning that it is flexible and can take on a many shapes, hence the reason for its wide use. Determining shape, scale and location of the Log Pearson 3 distribution requires calculating from the logarithms of the annual series, the skew, standard deviation and mean. These three statistical values determine a trend line and when it is plotted on a semi-log plot, it passed through the observed discharges. Some of these discharges are found to be outliers, not fitting with the general trend of the data. Since the data is ranked from the highest discharge to the smallest, these outliers occur at the low or high end of the distribution. Due to this, Cooper (2005) identified that the Log Pearson 3 distribution struggles to represent the outliers and the general trend of the discharges causing the distribution to significantly under or over-estimate the largest discharges.

The partial derivative function of the Log Pearson 3 distribution is given as (Ewemoje & Ewemooje 2011):

(2.3-4)

$$f(x) = \frac{\lambda^\beta (y-\varepsilon)^{\beta-1} e^{-\lambda(y-\varepsilon)}}{x\Gamma(\beta)}, \quad \log x \geq \varepsilon$$

where,

$$y = \log x$$

$$\lambda = \frac{S}{\sqrt{\beta}}$$

$$\Gamma(\beta) = \int_0^{\infty} u^{\beta-1} e^{-u} du, \quad \beta \geq 0$$

where,

x	=	Observed flow (m ³ /s)
y	=	Logarithm of observed flow (m ³ /s)
S	=	Standard deviation of floods
λ	=	Mean rate of recurrence
β	=	Shape parameter
$\Gamma(\beta)$	=	Gamma function
ε	=	Lower bound of gamma distribution

For Australian streams, excluding those in the arid zone, McMahon found that the annual peak flows had a mean standard deviation of the logarithms of 0.35 (range 0.12-1.3), while the world's non-arid zone average is 0.15 (range 0.06-0.36), hence indicating that the Australian streams are more than twice the world's average deviation (Hall 1984). Australian streams are considerably more variable than the rivers throughout the world, giving the Log Pearson 3 distribution an advantage as it can vary depending upon the standard deviation of the data.

However with a negative skew value, which is common for Australian flood data, there is an upper bound with the Log Pearson 3 distribution, causing difficulties in estimating floods of high recurrence intervals.

2.4 Power Law Flood Frequency Analysis

The conventional analysis methods often used to predict the return periods of extreme outlier events are highly unrealistic and have led many authors (Baker 1994) to be critical of these conventional techniques. The Power Law model offers a simpler alternative to the more complex probability models.

The Power Law relationship for flood frequency takes the form:

$$Q(T) = CT^\alpha \tag{2.4-1}$$

where C and α are regression coefficients in a log-log space.

The Power Law approach has found successful predictive and descriptive applications across a wide range of natural phenomenon. Turcotte (1997) presents a number of case studies using the Power Law relationship to model a variety of natural events including earthquakes and volcanic eruptions, while Scheidegger (1998) provides examples of its uses from landslides and Birkeland and Landry (2002) from avalanches (Kidson, Richards & Carling 2005). Malamud and Turcotte (1999) provided examples of the Power Law model with forest fires and Tzanis and Makropoulos (2002) provided more evidence of the use of the Power Law model for earthquakes.

A prominent application of the Power Law relationship to flood frequency was presented by Malamud et al., (1996), who demonstrated a close fit of the discharge to the recurrence interval of the extreme 1993 flood event on the Mississippi river with a Power Law model. Similar close fits were also demonstrated with historic data for the Colorado river.

Kidson et al. (2006) compared the Power Law model with conventional distributions (Log Pearson 3) for the prediction of the outlier flood events (large rainfall events). His study focussed on 50 United States rivers and 12 United Kingdom long term rainfall stations. It was demonstrated that the Power Law model produces a far more conservative return period estimates than the Log Pearson 3 at the higher discharge events.

The forecasted 100-year flood event using the Log Pearson 3 distribution has been found to be considerably smaller than forecasting using the Power Law approach. This difference was illustrated in the study by Malamud et al. (1996) on the 1993 Mississippi River flood. Where using the data at the Keukuk, Iowa gauging station found the flood to be a typical 100-year flood using the Power Law distribution and a 1000 to 10,000 year flood using the Log Pearson 3 distribution.

Malamud and Turcotte (2006) noted the Power Law relationship typically fitted better with the partial series flood record than the annual flood record, as the partial series is a better statistical sample. Kidson and Richards (2005) modelled a Power Law relationship to the partial duration series data which correlated with the extreme events better than the annual series data.

Studies conducted by both Kidson et al. (2006) and Malamud and Turcotte (2006) used the Weibull plotting position to calculate the recurrence intervals for the discharge values instead of Cunnane plotting position given in (2.3-3), recommended for use with the Log Pearson type 3 distribution. Kidson et al. (2006) states the Weibull plotting position formula was selected for their study due to its ability to provide simple unbiased exceedence probabilities independent of any distribution. This study has

however used the Cunnane plotting position for the Power Law model to maintain comparable results with the Log Pearson distribution.

According to an article by Malamud and Turotte (2006) a flood frequency factor can be found using the Power Law relationship. This factor is the ratio between the peak 10-year discharge and the peak 1-year discharge, which is the same as the ratio of the peak discharge of the 100-year event to the peak discharge of the 10-year event.

(2.4-2)

$$F = \frac{Q[10]}{Q[1]} = \frac{Q[100]}{Q[10]} = \text{constant}$$

According to Malamud and Turcotte (2006), this flood frequency factor is directly associated with the catchment's climate. Where it was found that for arid climates, the factor is relatively large. while for maritime climates the factor is relatively small (Malamud & Turcotte 2006).

2.5 Goodness of Fit

Cunnane (1985) describes how there is no analytical way of proving that a particular distribution is the correct distribution. For design purposes, the effectiveness of several types of probability distributions are able to be tested by identifying the fit of that the distribution has with the historical flood records.

The accurately of flood frequency analysis models are typically measured according to goodness of fit of the predicted values from the model against the observed data points.

These tests assist in determining the best fitting distribution to the given data and to describe differences between each distribution's expected values and the observed data values (Meylan, Favre & Musy 2012). However, goodness of fit tests can only conclude from the available data, whether to reject the hypothesis stating that a particular model is suitable and therefore these tests should not be used to pick the best distribution, rather to reject possible distributions.

The goodness of fit approach is flawed for a number of reasons. It assumes that the observed data taken from the stream gauges are precisely measured. Very high flows may be subjected to large absolute errors because of the difficulties of measurement and the lack of numerous confirming measurements.

Different goodness of fit tests will favour different models and not all models can be evaluated by the same test. Kidson and Richards (2005) demonstrated where different goodness of fit tests (absolute error and least squares error) favour alternative models. Therefore flood frequency analysis must be aware of the sensitivity to choice of goodness of fit tests and it is useful to employ a number of different tests to reduce the sensitivity. The goodness of fit test subjectivity can also be introduced through the choice of different plotting position formula for the distribution (Schertzer, Lovejoy & Lavallee 1993). These tests also tend to be insensitive in the prediction of extreme events.

The most commonly applied goodness of fit tests are the Chi-Squared test, the Kolmogorov-Smirnov test and the Anderson-Darling test (Meylan, Favre & Musy 2012). Due to the Power Law relationship not being a probability distribution, the Kolmogorov-Smirnov and Anderson-Darling tests were unable to be used and instead the more simple R-Squared test was used.

2.5.1 Chi-Squared Test, χ^2

The Chi-Squared test is a goodness of fit test as well as a conformity test, used to identify if the sample is from a specific distribution. According to D'Agostino and Stephens (1986) the Chi-Squared test is the most practical test of fit in many situations when the parameters are non-location-scale families or in uncommon distribution. If χ^2 equals zero, the distribution predicts the historical discharges, if χ^2 is greater than zero, it does not. The calculated chi squared value is compared with a critical value at selected significance level using degrees of freedom (Uregina n.d.). The chi square critical values are shown in table D.1 in Appendix D. A small chi squared value shows a close match between the observed values and the frequency distribution (Uregina n.d.). Therefore $\chi^2 < \chi^2_{crit}$, the fit is assumed to be satisfactory.

2.5.2 R-squared test, R^2

The R-squared test is a statistical measure of how close the observed data is to the fitted regression line.

The value of R^2 is a fraction that lies between 1 and 0. As demonstrated in Figure 2.5.1, when there is a R^2 value of 0, it means there is no relationship between the average recurrence interval and the historical stream discharges. When R^2 equals 1, the historical stream discharges all lie on the straight line provided by the probability distribution. (GraphPad Software n.d.)

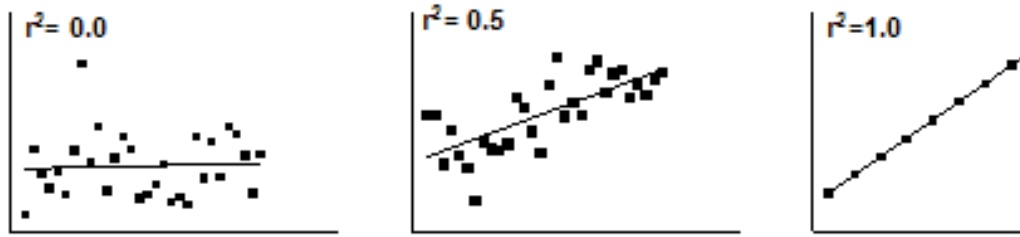


Figure 2.5.1 Examples of different values of R^2 (GraphPad Software n.d.)

The R-squared test is used in linear models. If the model is not linear, the R-squared test should not be used as the total sum-of-squares is will not be equal to the regression sum-of-squares and the residual sum-of-squares (Spiess & Neumeier 2010). Therefore the R-squared can easily be applied to the Power Law model. However, the majority of the statistical distributions including the Log Pearson 3 distribution are nonlinear and hence the R-squared test cannot be used. Ewemoje and Ewemooje (2011) in their study used the R-squared test to identify which plotting position fitted which statistical distribution best. To use this test Ewemoje and Ewemooje converted the predicted values from the probability distribution of the Normal distribution, Log Normal distributon and the Log Pearson 3 distribution into regression equations which were linear.

2.6 Poor Fits

A poor fit of the flood data to the probability distribution are characterised in two ways, according to the Australian Rainfall and Runoff (1987) guide. Firstly, by the presence of outliers in the distribution that is inconsistent with the overall trend of the remaining data and also by the discrepancies between the observed discharges and the fitted distribution.

There are a variety of reasons for poor fits of the probability distribution to plotting position and observed discharges (Australian Rainfall and Runoff 2010):

- 1) The smaller annual series maximums are not significant floods.
- 2) Rating curve extensions is biased causing over or under-estimate of the larger floods
- 3) Some of the observed floods could be unusually rare for the record length.
- 4) A change in the hydraulic control with discharge affecting the shape of the frequency curve
- 5) The flooding could be caused by multiple meteorological events (a storm and a tropical cyclone together), which is not responsive to most distributions.
- 6) Non-homogeneity of the flood record.

There are strategies are available to assist with a poor fit of the flood frequency distribution. A more flexible probability model could be fitted or the responsible data for the poor fit could be weighted less (Australian Rainfall and Runoff 2010).

2.7 Summary

This section has covered a large amount of literature regarding the conventional methods of flood frequency analysis used in Australia according to the Australian Rainfall and Runoff (1987) guide. It initially describes the types of data series used in flood analysis, in particular the annual flood series and the partial flood series.

The conventional methods of analysis using in Australia was identified as the Log Pearson type 3 probability distribution and this chapter also explains the limitations and complications of this distribution. The Power Law model was explained and previous studies of this model were discussed.

This chapter also describe a number of goodness of fit tests that will used to determine the accuracy of each model as compared to the historical peak discharges by the use of the Chi-Squared goodness of fit test and the R-Squared test.

Finally, research was conducted regarding poor fit of the flood data to the probability distributions. A variety of reasons for possible poor fits had been identified and along with strategies to deal with them.

Chapter 3 – Methodology

3.1 Chapter Overview

This chapter details the methods and procedures that were undertaken to complete this research project. In particular the selection criterion used for choosing the ten stream gauge stations from catchments across Queensland, as well as the specific approach taken to prepare both the annual flood series and the partial flood series. The approach used in creating the three frequency models to be analysed has also be stated. The chapter concludes with the methodology of applying the goodness of fit tests to the frequency models.

3.2 Resource Analysis

The number of resources that are required for this project is small. The data for the analysis will be obtain through two online databases, the Water Monitoring Data Portal produced by the Queensland Department of Natural Resources and Mines and the Hydrological Reference Station database produced by the Australian Government Bureau of Meteorology.

The spread sheet and data analysis software will be utilised in the form of Microsoft Office's Excel package and Mathwave's EasyFit, an Excel add-in which automates the process of fitting probability distributions to the data selection.

3.3 Station Selection

This project analyses peak flood discharges from generally unregulated gauging stations from catchments across Queensland. A strict criterion was set for the selection of stream gauges and analysis of stream gauge data. For this project, data sources were restricted to the Water Monitoring Data Portal, a database of stream gauging data produced by the Queensland Department of Natural Resources and Mines and the Hydrological Reference Station database, a database showing Australian stream flow trends by the Australian Government Bureau of Meteorology. This source produced an initial pool of approximately 410 potential stream gauges across Queensland. The stream gauge data was sorted according to the following criteria, which reduced the number of stations to a final list of ten stations; see Figure 3.3.1.

The initial criterion for gauge selection was that all data had to be concurrent for at least thirty water years from the 1 July 1984 to 30 June 2015. Data files were checked using the station summary reports to identify stations that suited this criterion. Gauging stations were not included if their data had a gap at a time when the stream flow had possibly peaked above the mean daily flow (or if the magnitude of a missing flow peak could not be estimated by linear interpolation). To check this, the flow records for the

nearby and/or upstream stations were examined to see if significant flow events had been recorded for the relevant period.

There was limited time was available for this project and in a more detailed study it should be possible to include more gauging stations because there would be more time to track down high quality stream gauging data.

Table 3.3.1 Summary of the final stream gauge stations used in this analysis

Station ID	Station Name	Basin	Catchment Area (km²)	Distance from Stream mouth (km)	Record Length (years)	Record Period (1 July to 30 June)
922101B	Coen River at Racecourse	Archer Basin	172	154.5	47	1968 – 2015
112002A	Fisher Creek at Nerada	Johnstone Basin	16	2.9	40	1975 – 2015
116006B	Herbert River at Abergowrie	Herbert Basin	7454	71.8	45	1970 – 2015
126003A	Carmila Creek at Carmila	Plane Basin	84	11.8	42	1973 – 2015
137201A	Isis River at Bruce Highway	Burrum Basin	446	22.7	49	1966 – 2015
138001A	Mary River at Miva	Mary Basin	4755	126	105	1910 – 2015
142001A	Caboolture River at Upper Caboolture	Pine Basin	94	31.4	49	1966 – 2015
143303A	Stanley River at Peachester	Brisbane Basin	104	89.2	88	1927 – 2015
146010A	Coomera River at Army Camp	South Coast Basin	88	45.2	52	1963 – 2015
422394A	Condamine River at Elbow Valley	Balonne-Condamine Basin	325	1136.8	42	1973 – 2015

The median start year for the stream gauge stations used was 1967, with a mean record length of 56 years. In total 559 years of flood data was analysed across the 10 Queensland stations. The location of these ten stream gauge stations can be seen on the map below in Figure 3.3.1.



Figure 3.3.1 The location of the ten stream gauge stations used in this analysis

3.4 Annual Series and Log Pearson 3 Distribution

The annual flood maxima for each site were ranked in descending order so that the floods range from largest to smallest. For the annual series, missing record periods is of no importance and is able to be included in analysis if it is determined that the largest discharge for that year did not occur during the missing record period (Australian Rainfall and Runoff 1987). The rainfall records and stream flows of nearby catchments were used in determine if a large flood occurred during the missing record period.

The average exceedence probability for each flood was estimated using the Cunnane plotting position recommended by the Australian Rainfall and Runoff guide for the Log Pearson 3 distribution. This plotting position formula is stated earlier as equation (2.3-2).

Engineers Australia recommends in the Australian Rainfall and Runoff guide (2006) that the fitting procedure of the method of moments is to be used for the Log Pearson distribution. This procedure is outlined here.

Firstly the logarithm of the observed flow data is needed.

(3.4-1)

$$y_i = \log x_i$$

where;

x_i = Observed historical flow data (m³/s)

y_i = Observed flow data logarithm

The key statistics of the observed dataset is then calculated from the logarithms, this includes the calculation of the mean, standard deviation and coefficient of skewness.

(3.4-2)

Mean $M = \frac{1}{N} \sum_{i=1}^N X_i$

(3.4-3)

Standard Deviation $S = \left[\frac{\sum (X_i - M)^2}{N-1} \right]^{0.5}$

(3.4-4)

Coefficient of Skewness $g = \frac{N \sum (X_i - M)^3}{(N-1)(N-2)S^3}$

where;

N = Number of years of record

X_i = Logarithm of observation i

If the coefficient of skewness is between -1 and +1, the value of the frequency factors for the distribution, K_T , can be determined using the equation (Chin 2006).

(3.4-5)

$$K_T = \frac{1}{3k} ([(z - k)k + 1]^3 - 1)$$

Where,

$$k = \frac{g}{6}$$

The standard variable z is computing by taking the inverse of the normal cumulative density function. (Abramowitz & Stegun 1965)

(3.4-6)

$$z = w - \frac{2.515517 + 0.80285w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}$$

Where,

$$w = \sqrt{\ln\left(\frac{1}{P^2}\right)}, \quad 0 < P \leq 0.5$$

With P equal to the observed return periods. The value of z is computed using (3.4-6) and given as a negative sign. Using equation (3.4-6) to estimate the frequency factor there is an error of less than 0.00045 (Chin 2006).

The values of K_T is more easily found using Table B.1 and Table B.2 shown in Appendix B from Water Resources Council (1967), these K_T values can also be found in the Australian Rainfall and Runoff (1987) guide.

The flood discharges at various return intervals is then determined using the expression:

(3.4-7)

$$\log Q_T = M + K_T S$$

Where,

M = Mean logarithm value of floods

S = Standard deviation of the logarithm of the peak floods

K_T = Frequency factor for the distribution corresponding to T years recurrence interval

$\log Q_T$ is the logarithm of a flood discharge having the same recurrence interval or percent chance. Find the antilog of $\log Q_T$ to get the flood discharge, Q_T .

3.4.1 Outliers

Outliers that are significantly different from the trend of the data can greatly affect the fitted distribution and the estimate of flood peaks from the distributions. The following formulas are used to identify the high and low outliers in the historical flood data.

The equation used to indicate high outliers is (Australian Rainfall and Runoff 1987):

(3.4-8)

$$X_H = M + \beta K_N S$$

Where,

X_H = High outlier threshold in log units

M = Mean logarithm value of floods

S = Standard deviation of the logarithm of the peak floods

K_N = Values from Table C. correlating to the number of years of data (N)

β = An adjustment factor depending on the number of years of data (N) and the skew of the logarithms of the flood data (g) given in Table D.2

If any of the values calculated used equation (3.4-1) are above the value of X_H from equation (3.4-8), then it is most likely an outlier, while values below the value of X_H have no statistical evidence of being an outlier. The omission or deletion of a value in the flood series, is regarded as an extreme step and only done when no other course of action is justifiable. If a value is omitted, the analysis should be completed as if the data omitted was not part of the series and the number of years of recorded data is reduced by one.

The equation used in determined low outliers is (Australian Rainfall and Runoff 1987):

$$X_L = M - \theta K_N S$$

Where,

X_L = Low outlier threshold in log units

M = Mean logarithm value of floods

S = Standard deviation of the logarithm of the peak floods

K_N = Values from Table C. correlating to the number of years of data (N)

θ = An adjustment factor depending on the number of years of data (N) and the skew of the logarithms of the flood data (g) given in Table D.3

If the logarithms of any peaks in the annual series are less than X_L , the data is considered as a low outlier and omitted from the data series.

3.5 Partial Series and Exponential Distribution

As stated in Chapter 2, the partial series requires firstly that a base discharge is chosen so that the series consists of all independent floods above the base value. To be classed as an independent flood peak, the maximum daily stream discharge needs to be separated by a certain number of days.

The construction of the partial flood series, the dataset consisting of daily stream discharge were obtained from the Department of Natural Resources and Mines and the Hydrological Reference Station database for each station and period of time considered.

The approach taken to maintain independent of the flood peaks was based on the approach taken by Malamud and Turcotte (2006) in their article, where the maximum daily discharge for the entire time period was found and all of the discharges found 30 days either side of this value were deleted, giving the largest ‘flood’ for that stream. Following on from that, the next maximum discharge of the values remaining was identified and all values that are within 30 days of it were deleted, to provide the second largest flood in the flood series. (Malamud & Turcotte 2006) The process was continued until there were K number of largest floods for the period considered. If possible, the base discharge for this approach should be selected so that the number of floods, K is two to three times N , which is the number of years of record, according to the Australian Runoff and Runoff (1987) guidelines. However, when the number of floods is low in regions due to low rainfall, it may be necessary to use a smaller value of K .

The base value for the peak discharges was chosen so that the number of floods (K) used for the partial series was approximately equal to 1.5 – 2 times the number of years of record (N). Due to the number of years of drought in the recent times in Queensland, the rainfall is low and hence a smaller ratio of K to N was used than the recommended by the Australian Runoff and Runoff (1987) guidelines.

There are several methods suggested by the Australian Rainfall and Runoff (2010) guide for the treatment of missing record periods in partial duration flood series. In this analysis where a nearby station record exists which covers the missing record period; the correlated data was used from the nearby station to predict the discharges in gap. Where there was no nearby station which covered the missing record period, the missing data was ignored in the analysis and the overall period of the record was left missing that record period.

A plotting position in the form of an average recurrence interval is needed for the partial series, this is estimated using equation (2.3-3). The design discharges for the partial series are calculated from the regression equation:

(3.5-1)

$$Q_T = a \log(ARI) + b$$

where, a and b are regression constants.

The partial series data should be plotted on a semi-log (or log-linear) scale where an exponential distribution plots as a straight line. Each observed flood peak is plotted on the linear scale and the average recurrence interval calculated using (2.3-3) is plotted on the logarithmic scale. A regression line can be drawn and the regression constants are identified.

3.6 Power Law Relationship

The Power Law relationship is constructed using for the most part the partial duration flood series, as described in section 3.5. However for the construction of this relationship the number of floods K used in the partial series is equal to the number of years of record N for each station, so that the Power Law relationship is not heavily influenced by low magnitude floods.

The plotting position for the Power Law model used the Cunnane plotting position. Hence the plotting position used is the same as (2.3-3) as the plotting position needs to be in the form of an average recurrence interval.

The design discharges for the Power Law distribution are calculated from the regression equation which was stated earlier in equation (2.4-1).

The partial series data should be plotted on a log-log scale, turning the Power Law relationship into a straight line when plotted. Where the each flood peak is plotted against the average recurrence interval calculated using (2.4-1). A regression line can be drawn and the regression constants are identified.

3.7 Applying Goodness of Fit tests

Goodness of fit tests were applied to each of the flood frequency models for each individual station. They involved comparing the results of each model to the historically observed flood records for each station to determine the effectiveness of the frequency model in predicted previous flood event's discharges.

3.7.1 Chi-Squared

This test is applied to binned data, so the value of the test statistic depends on how the data is binned. The following formula was used to determine the number of bins used in the calculation.

(3.7-1)

$$k = 1 + \log_2 N$$

where,

$k =$ The number of bins

$N =$ Number of years of flood

The peak discharges are then grouped into bins of equal widths and the frequency of discharges in each bin are calculated. The Chi-Squared goodness of fit test requires that there be at least five data points in each bin, so some adjacent bins may need to be joined together to make sure there is at least five data points in the bin or it may be necessary to group the data into large bins (Singh 1986).

The Chi-Square test statistic is then able to be calculated by taking an observed discharge frequency (O_i), subtracting the expected discharge frequency given from the probability distribution (E_i) and then squaring this difference. Squaring the differences makes the error positive and then can be divided by the expected discharge frequency. The standardised difference for each discharge is then summed.

The Chi-Squared test statistic (χ^2) given as:

(3.7-2)

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

where,

$O_i =$ the observed frequency for bin i

$E_i =$ the expected frequency for bin i

The test statistic is then compared to the critical value found in table D.1, where the degrees of freedom is one less than the number of discharges used to calculate the test statistic. The selected significance level for the analysis is 80 percent, so that a distribution model is rejected when the test statistic is greater than the critical test

statistic value at $\chi^2_{.80}$ in Appendix D. When comparing the three models for each station, the lower the test statistic value is the closer the match is between the probability distribution model and the historic observed values for the flood data.

For the Log Pearson type 3 distribution and the exponential distribution, EasyFit software produced by Mathwave was used to calculate the Chi-Squared value for the goodness of fit test, whereas the Chi-Squared value for the Power Law model was calculated in Microsoft Excel using the equation (3.7-2).

3.7.2 R-Squared

R-Squared values range from zero to one, with a value of one meaning that the historical flood discharges are completely explained by the frequency distribution. A high R-squared (usually 0.85 to 1) indicates a correlation between the historical data and the distribution; while a low R-squared (0.7 or less) indicates the model does not provide an effective estimate of the historical flood record (Invetopedia 2015).

To complete this test, the discharges were calculated from each of the distribution model for each of the recurrence intervals given by the plotting position for each station. That way the historical discharges are compared directly with the model's predicted discharge at the same recurrence interval.

The R-Squared value is calculated by (Khan Academy 2010):

(3.7-3)

$$R^2 = \sum_{i=0}^N \frac{(Q_{M,i} - Q_{H,i})^2}{(Q_{M,i} - \bar{Q}_M)^2}$$

where,

$Q_{H,i}$ = The historical discharge

$Q_{M,i}$ = The discharge provided by the model

$\overline{Q_M}$ = The mean of discharges provided by the model

N = Number of years of flood

This formula was applied to three data sets for each frequency model. The first dataset was the full number of year of flood data from the annual and partial series, the R-squared value found from this can be compared with the Chi-Squared test result. The next dataset uses the discharges which have a plotting position of less than ten years, this will be used to identify if the exponential distribution is the frequency model that is most effective in predicting the more common floods as suggested by the Australian Rainfall and Runoff (1987) guide. The final R-Squared value calculated is for the floods that have a plotted position that is greater than a ten year recurrence interval and used to identify the effectiveness of the Log Pearson 3 distribution and the Power Law model for the higher discharge flood.

3.7.3 Graphical Comparison

For each of the stream gauge stations, graphical comparison of the observed peaks discharges to the model predicted discharges was also completed to assist with the determination of the effectiveness of the three model types.

The values for the model predicted discharges were obtained with the plotting position formula from equation (2.3-3). The recurrence interval (in years) from each of the observed floods was used as the time period in each of the model's calculations to find

their respective discharges. The discharges then given by the models for each stream gauge can then be compared with the observed discharge from the historical data.

A one-to-one reference line was used in the graphical comparison to help identify where the observed discharges matches the model predicted discharges. The model which has their discharge predictions closest to the one-to-one reference line provides the most accurate prediction of the peak discharge according to the historical flood data. The average location of the ten year recurrence interval was also included in the graphs to help identify the effectiveness of the model with both the less than ten year floods and the greater than ten year floods.

This graphical comparison identifies the accuracy of each model to predicting the observed discharge using the Cunnane's plotting position formula as the average recurrence interval for each flood. Since this plotting position is an estimate of the return period of the historical floods, this graphical comparison has its limitations which are discussed in the discussion chapter.

3.8 Summary

This chapter has describes the processes used to generate the annual and partial flood duration series' and applying the probability distributions to them. The final part of this chapter covers goodness of fit testing procedures for each distribution type for each of the selected stream gauge stations and the critical values associated with these tests have also been identified.

Chapter 4 – Results

4.1 Chapter Overview

The previous chapter described methods of generating each frequency model for this project, along with the processes used in the analysis of each model for its effectiveness.

This chapter looks at the analysis that was performed on the historical stream gauge data with respect to the three distributions, Log-Pearson type 3, Exponential and Power Law. Each model has been identified for the ten stream gauge station and goodness of fit values for distribution's fit has also identified in this chapter both figuratively and graphically.

4.2 Stream Gauge Results

The section provides the individual results of each of the stream gauges used in this analysis so that the three flood frequency models can be discussed in in relation to each individual stream gauge station.

The ten stream gauges that have been used in this analysis and their results shown in this chapter are:

- 922101B – Coen River at Racecourse
- 112002A – Fisher Creek at Nerada
- 116006B – Herbert River at Abergowrie
- 126003A – Carmila Creek at Carmila
- 137201A – Isis River at the Bruce Highway
- 138001A – Mary River at Miva
- 142001A – Caboolture River at Upper Caboolture
- 143303A – Stanley River at Peachester
- 146010A – Coomera River at Army Camp
- 422394A – Condamine River at Elbow Valley

More details regarding each station can be found in Table 3.3.1 and Figure 3.3.1 in Chapter 3.

4.2.1 Coen River at Racecourse (922101B)

The Coen River stream gauging station at Racecourse is situated in the Archer basin in far north Queensland and has a catchment area of 172 km². The mean of the logarithms of the annual peak flows was found to be 2.388, while the standard deviation is 0.371 which is twice the world non-arid zone average deviation (Hall 1984). The skew of the logarithms of the flood data series is -0.268, which means the that mean of the flood peaks is less than the median value.

The flood series data was checked for high outliers. The high outlier threshold (X_H) was found to be 3.347, which correlates to a peak discharge of 2225 m³/s. The highest peak discharge for the Coen River stream gauge is 1100 m³/s, hence there are no high outliers

for this data series. The low outlier threshold (X_L) was found to be 1.233, which correlates to a peak discharge of 17.08 m³/s. The lowest peak discharge in the annual series is 38.09 m³/s, therefore there is no low outliers in the Coen River annual series data.

The Log Pearson type 3 distribution parameters for the Coen River at Racecourse found using MathWave Technologies' Easyfit software is $\alpha = 55.5, \beta = -0.115$ and $\gamma = 11.86$.

The exponential equation formed from the partial series data is, determining the peak discharge in cubic metres per second for a selected average recurrence interval in years is:

$$Q_{PEAK} = 534 \log(ARI) + 161$$

The Power Law model is formed using the equation: $Q_{PEAK} = 259.7(ARI)^{0.364}$, with a flood frequency factor of 2.31 between the associated 100-year flood to the 10-year flood discharge. Since the flood frequency factor is relatively small and the station under maritime climate.

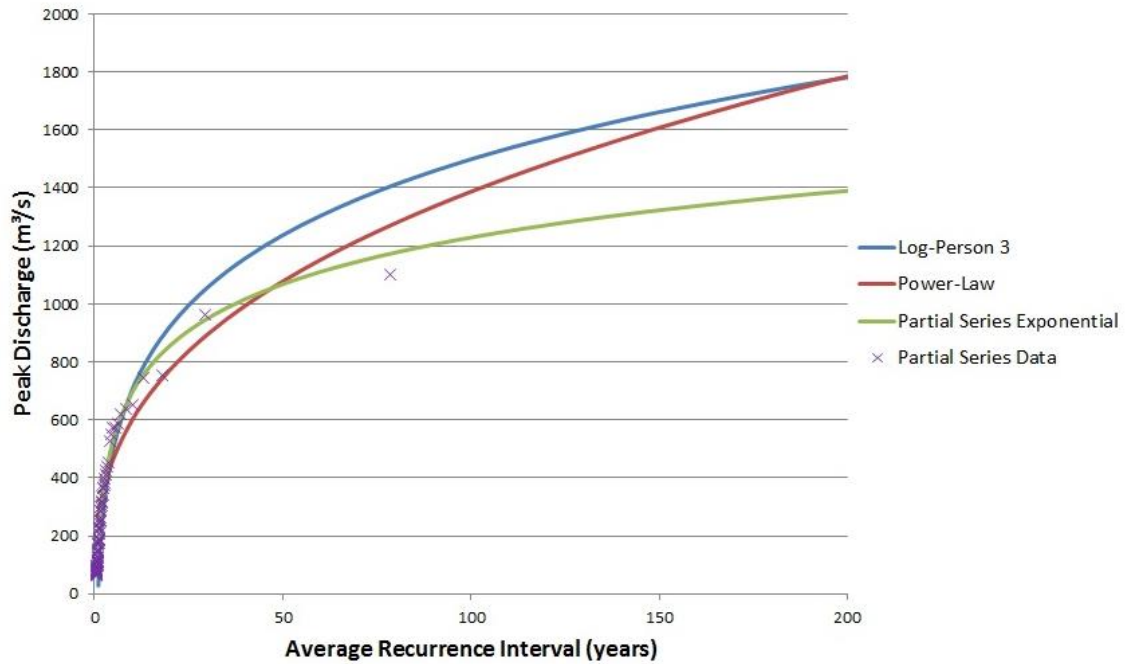


Figure 4.2.1 Coen River stream gauge model predicted peak discharges at Racecourse fitted to the average recurrence intervals

Figure 4.2.1 shows the results of each of the flood frequency models differs over a range of years of recurrence. It is noted that the Power Law model provides that highest predicted discharges for floods that are greater than the 200-year recurrence interval, while the exponential distribution provides the lowest predicted discharges.

The results of the goodness of fit tests are shown in Table 4.2.1 and Table 4.2.2. Table 4.2.1 gives the Chi-Squared critical test statistic value for the Coen River stream gauge at Racecourse, calculated using equation (3.7-2). It can clearly be seen that the model provided by the exponential distribution matches most closely with the historical data for this station as it has the lower test score. Both the Log Pearson 3 distribution and the Power Law provide high critical test statistic values and hence a probability that is under 80 percent, therefore according to the Chi-Squared test these models should be rejected as a poor fit for the Coen River.

Table 4.2.1 Coen River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	4.3025	0.367
Exponential (partial series)	2.2823	0.809
Power Law (partial series)	3.3968	0.639

Table 4.2.2 Coen River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.968	0.904	0.937
Exponential (partial series)	0.989	0.987	0.988
Power Law (partial series)	0.796	0.964	0.889

The R-Squared values for the Coen River are shown in Table 4.2.2. In the column labelled 'overall' in Table 4.2.2, gives the values of the R-Squared test using the entire flood data series. It can be identified that the goodness of fit from the R-Squared test of each of the distribution models matches the results of the Chi-Squared test in which the exponential distribution has the highest probability. However, the R-Squared test has the Log Pearson distribution a closer match to the historical data according to the R-Squared test than the Power Law, while the Chi-Squared test has the Power Law as the more accurate model. Since there the Log Pearson matches more closely with the smaller floods and there are more of them, the overall R-Squared test is biased towards the smaller floods.

The R-Squared values when the time period is less than the 10-year average recurrence interval shows the exponential distribution has the closest match to the historical flood data, reinforcing that the exponential distribution is sufficiently accurate when using the partial series where the recurrence interval is less than 10-years.

Lastly, Table 4.2.2 also shows the goodness of fit when the average recurrence interval plotted position is greater than 10-years. R-Squared values show that the exponential distribution of the partial series is again the closest fit to the historical flood data, however more interestingly is that the Power Law relationship was found by this goodness of fit test to be closer to the historical flood data than the Log Pearson type 3 distribution for floods great than the 10-year recurrence interval.

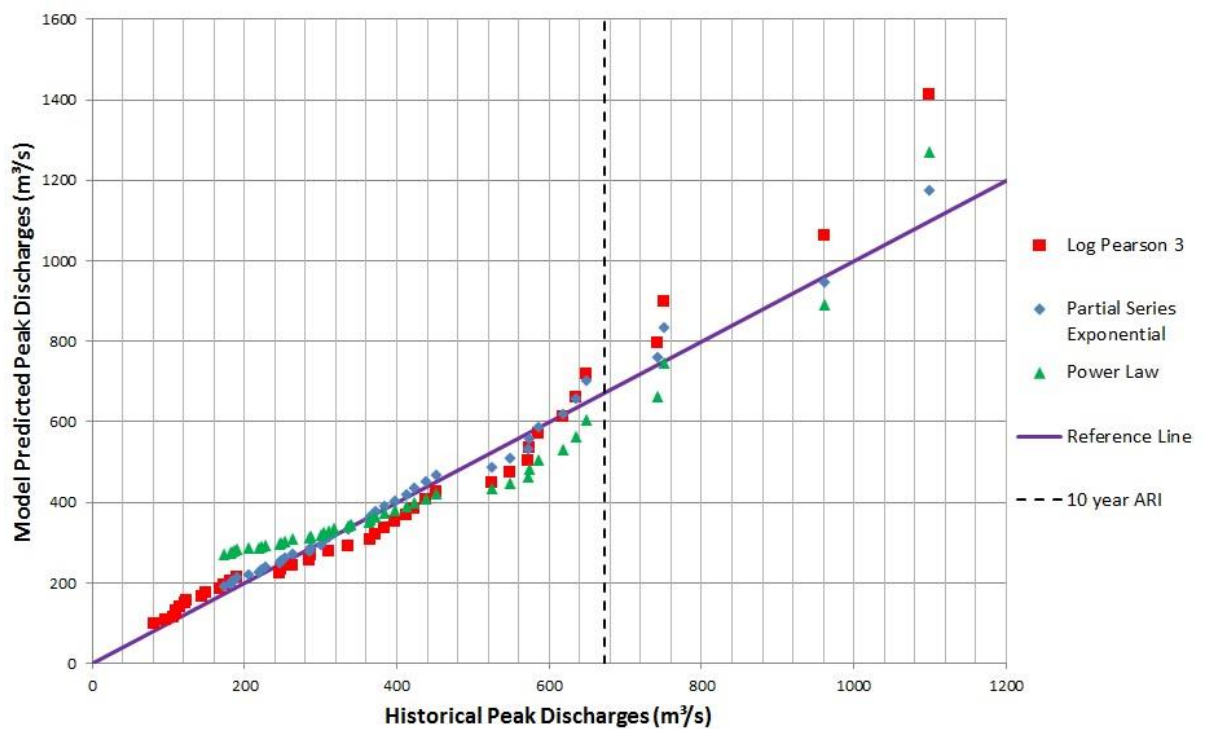


Figure 4.2.2 Coen River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data

The graphical demonstration shown in Figure 4.2.2, gives the predicted peak discharges of each model compared to the observed historical peak discharges for the Coen River stream gauge. The one-to-one line is given for reference and the approximate location of 10 percent average exceedence probability. It is identifiable that the exponential distribution predicts the peak discharges most accurately as the exponential plots closest to the one-to-one reference line, reinforcing the results obtained from the goodness of fit tests.

4.2.2 Fisher Creek at Nerada (112002A)

The stream gauge station on Fisher creek at Nerada is in the Johnstone basin in the far north region of Queensland, with a catchment area of 16 km². The mean of the logarithms of the annual peak flows was found to be 1.821, while the standard deviation is 0.386, reinforcing McMahon's article stating that the streams in Australia are more variable than the rivers throughout the world (Hall 1984). The skew of the flood data series is -0.098, which is similar to a normal distribution (skew = 0).

The flood series data was firstly checked for high outliers. The high outlier threshold (X_H) was found to be 2.884, which correlates to a peak discharge of 765 m³/s. The highest peak discharge for the Fisher Creek stream gauge is 416.5 m³/s; hence there are no high outliers for this data series. The low outlier threshold (X_L) was found to be 0.708, which correlates to a peak discharge of 5.10 m³/s. The lowest peak discharge in the annual series is 11.998 m³/s, therefore there is no low outliers in the Fisher Creek annual series data.

The Log Pearson type 3 distribution parameters for the Fisher Creek at Nerada found using MathWave Technologies' Easyfit software is $\alpha = 419.5$, $\beta = -0.043$ and $\gamma = 22.39$.

The exponential equation formed from the partial series data is:

$$Q_{PEAK} = 158 \log(ARI) + 45.3$$

While the Power Law model is found by: $Q_{PEAK} = 68.9(ARI)^{0.446}$. The estimated discharge associated with the 100-year flood is 2.79 times larger than the discharge associated with the 10-year flood, according to the frequency factor of the Stanley River stream gauge. Since the flood frequency factor is relatively small and the station under maritime climate.

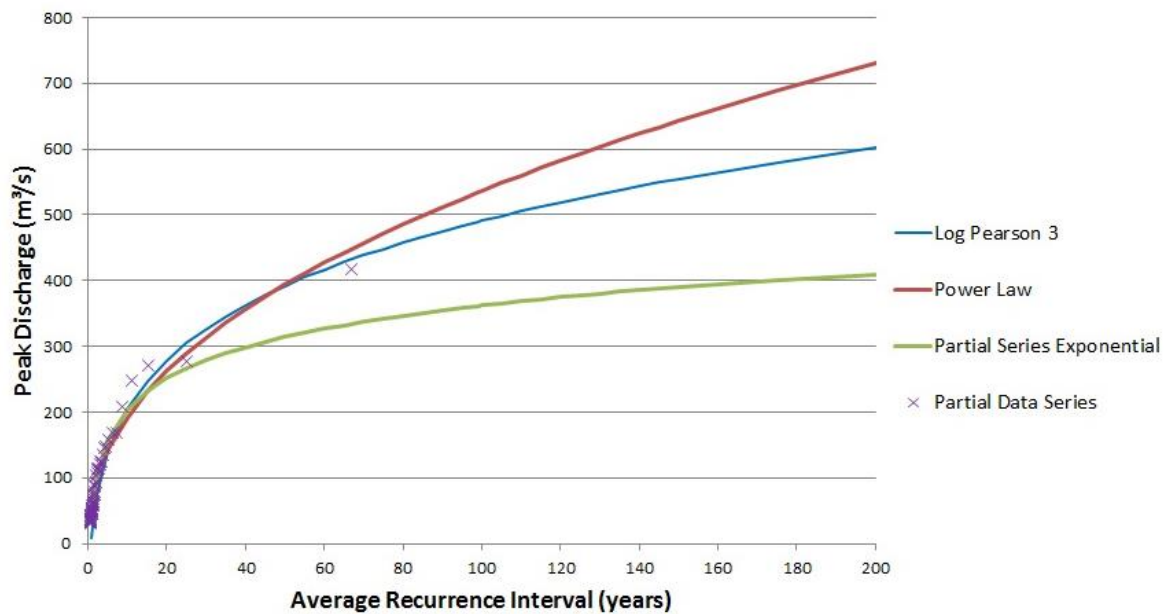


Figure 4.2.3 Fisher Creek stream gauge model predicted peak discharges at Nerada fitted to the average recurrence intervals

Figure 4.2.3 shows how each of the flood frequency models differs over the range of years of recurrence. It is noted that the Power Law model provides that highest

predicted discharges for floods that are greater than the 50-year recurrence interval, while the exponential distribution of the partial series data provides the lowest predicted discharges. From Figure 4.2.3, it can be seen that the Log Pearson 3 distribution seems to be closest to the partial series dataset, particularly for the floods of high recurrence intervals.

Table 4.2.3 gives the Chi-Squared critical test statistic values for the Fisher Creek stream gauge at Nerada, calculated using equation (3.7-2). The Log Pearson type 3 distribution of the annual series matches more closely to the historical data as it has the lower test statistic value, while the Power Law model has the highest Chi-Squared test value and has the least accurate fit with the historical flood partial series.

Table 4.2.3 Fisher Creek results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	1.035	0.960
Exponential (partial series)	2.2963	0.807
Power Law (partial series)	2.3096	0.805

Table 4.2.4 Fisher Creek analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.992	0.987	0.989
Exponential (partial series)	0.990	0.953	0.962
Power Law (partial series)	0.913	0.970	0.952

The R-Squared values for the Fisher Creek stream gauge models are shown in Table 4.2.4. From the values of the R-Squared test over the entire flood data series, it can be identified that the probabilities from the overall R-Squared test matches the probabilities of the Chi-Squared test with the Log Pearson distribution having the highest probability. The R-Squared values when the time period is less than 10 years shows the Log Pearson distribution has the closest match, closely followed by the exponential distribution.

Table 4.2.4 also shows the goodness of fit when the average recurrence interval plotted position is great than 10 years. The R-Squared values show the Log Pearson is again the closest fit to the historical flood data, while the Power Law relationship was found to be closer than the exponential distribution.

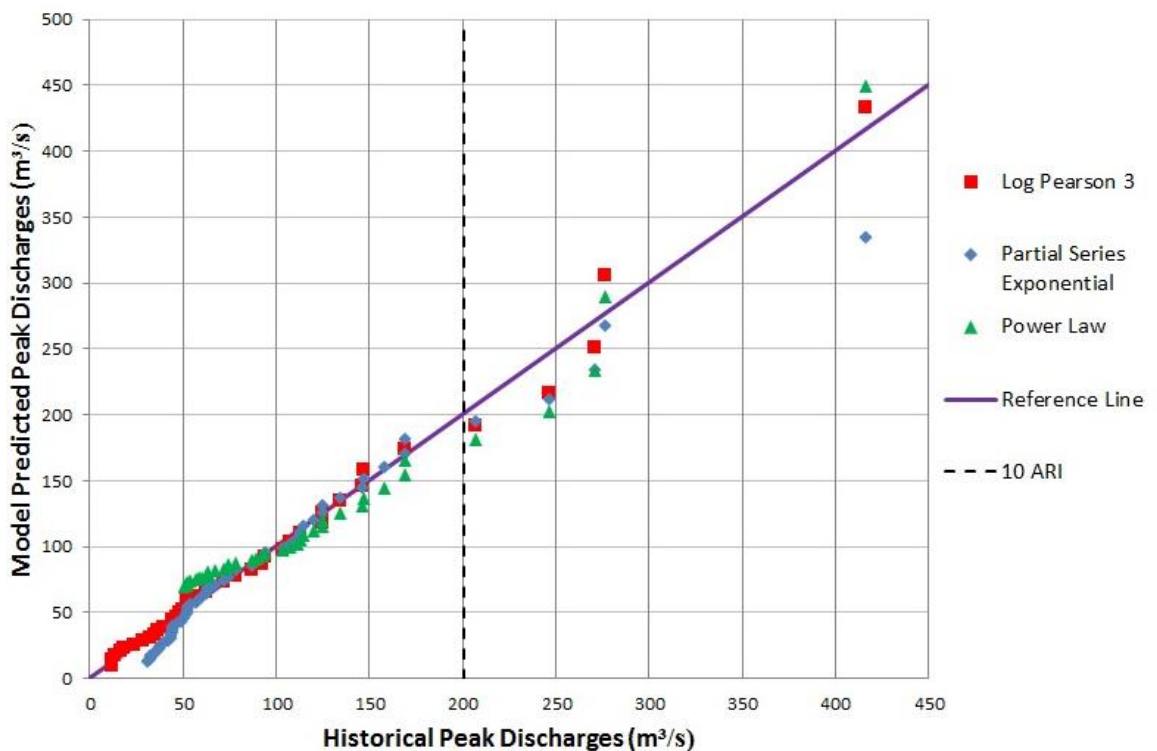


Figure 4.2.4 Fisher Creek stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data

The results discussed in Table 4.2.3 and Table 4.2.4 are shown graphically in Figure 4.2.4. This graphical demonstration shows the predicted peak discharges of each model compared to the observed historical peak discharges for the Fisher Creek stream gauge. It is identified that the Log Pearson distribution, given as the red squares are the closest data range to the purple reference line. This is appropriate for the discharges both less than and greater than the 10-year average recurrence interval line.

4.2.3 Herbert River at Abergowrie (116006B)

The Herbert River stream gauging station at Abergowrie is situated in the Herbert basin in the North Queensland region and has a catchment area of 7454 km². The mean of the logarithms of the annual peak flows was found to be 3.299, while the standard deviation is 0.412 which is greater than twice the world non-arid zone's average deviation (Hall 1984). The skew of the logarithms of the flood data series is -0.230, giving the mean of the flood peaks is less than the median value.

The flood series data was checked for high outliers. The high outlier threshold (X_H) was found to be 4.375, correlating to a peak discharge of 23,695 m³/s. The highest peak discharge for the Herbert River stream gauge is 9,458 m³/s, therefore there are no high outliers for this data series. The low outlier threshold (X_L) was found to be 2.036, which correlates to a peak discharge of 109 m³/s. The lowest peak discharge in the annual series is 241 m³/s, therefore there are no low outliers in the Herbert River annual series data either.

The Log Pearson type 3 distribution parameters for the Herbert River at Abergowrie found using MathWave Technologies' Easyfit software is $\alpha = 75.422$, $\beta = -0.10922$ and $\gamma = 15.834$.

The exponential equation formed from the partial series data is calculated to be:

$$Q_{PEAK} = 5863 \log(ARI) + 1111$$

The Power Law model is defined as: $Q_{PEAK} = 2234(ARI)^{0.392}$, with a flood frequency factor of 2.47. Hence the estimate for the discharge associated with the 100-year flood is 2.47 times larger than the discharge associated with the 10-year flood. Since the flood frequency factor is relatively small and the station under maritime climate.

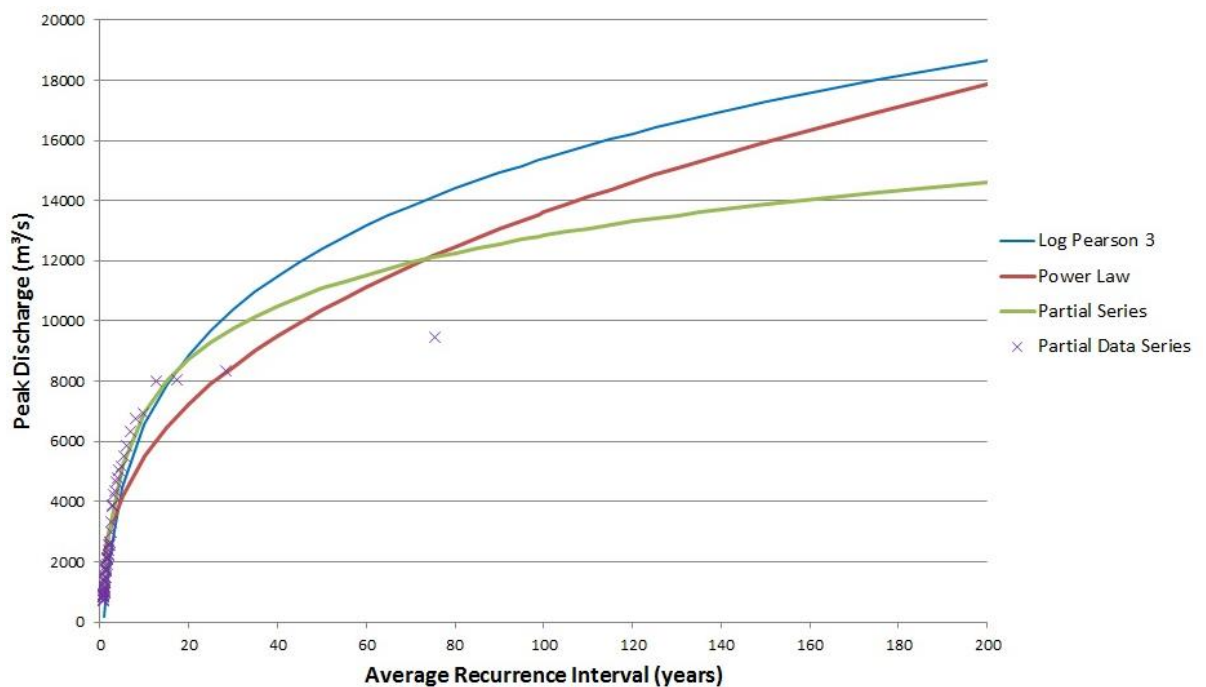


Figure 4.2.5 Herbert River stream gauge model predicted peak discharges at Abergowrie fitted to the average recurrence intervals

Figure 4.2.5 shows how each of the flood frequency models differs over the range of years of recurrence. It is noted that the Log Pearson distribution provides that highest

predicted discharges for floods; however, for the more extreme floods the Power Law model provides the large peak discharges, while the exponential distribution provides the lowest predicted peak discharges.

Table 4.2.5 gives the Chi-Squared critical test statistic value for the Herbert River stream gauge at Abergowrie, calculated using equation (3.7-2). The model provided by the exponential distribution has the probability closest to one and hence matches more closely with the historical flood data. The Power Law model on the other hand, has the highest Chi-Squared test value and therefore has the least accurate fit. Since the Power Law has a probability of less than the critical 80 percent, according to the Chi-Squared test show, the Power Law relationship should be rejected for the Herbert River.

Table 4.2.5 Herbert River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	1.1715	0.883
Exponential (partial series)	1.0425	0.903
Power Law (partial series)	2.696	0.610

Table 4.2.6 Herbert River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.952	0.791	0.882
Exponential (partial series)	0.980	0.911	0.949
Power Law (partial series)	0.742	0.874	0.801

The R-Squared values for the Herbert River are shown in Table 4.2.6 show that the goodness of fit from the overall R-Squared test matches the results of the Chi-Squared test as the exponential distribution of the partial series data fits the historical data most accurately. The R-Squared values when the time period is less than 10-year average recurrence interval gives similar results to the overall data results, the exponential distribution has the closest match reinforcing the idea that the exponential distribution is sufficiently accurate when the recurrence interval is less than 10 years.

The goodness of fit when the average recurrence interval plotted position is great than 10 years shows the exponential distribution is again the closest fit to the historical flood data, however more interestingly is that the Power Law relationship was found by this goodness of fit test to be closer to the historical flood data than the Log Pearson type 3 distribution for floods great than the 10-year recurrence interval.

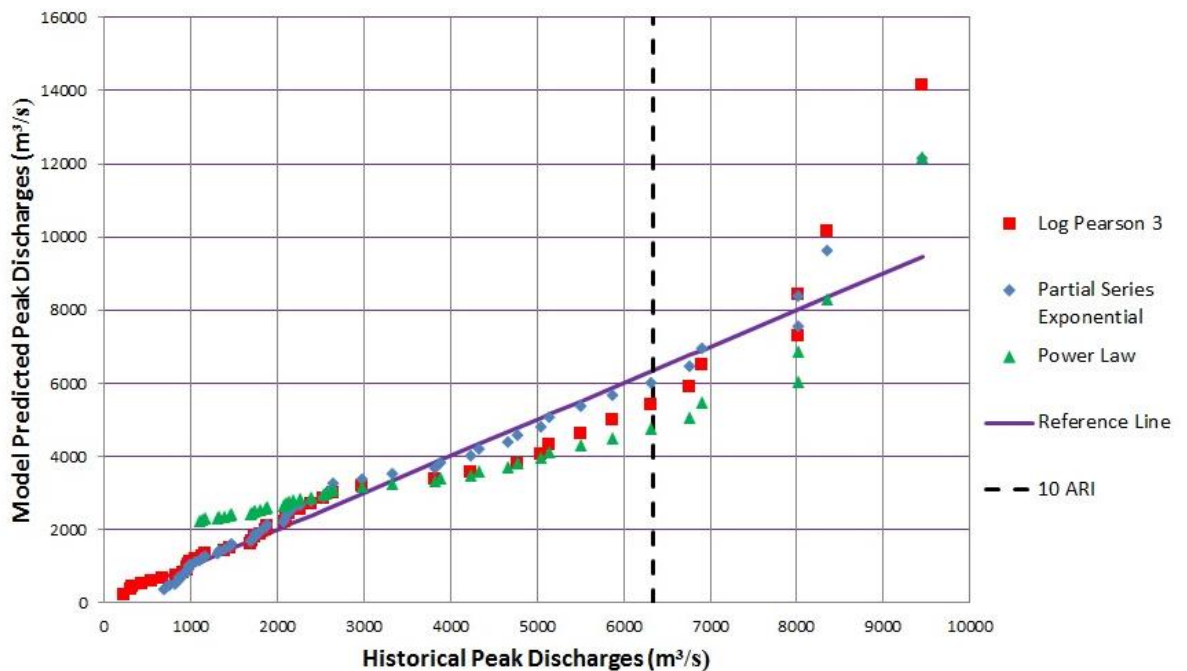


Figure 4.2.6 Herbert River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data

The graphical demonstration in Figure 4.2.6 shows the exponential distribution's predicted peak discharges provide a closer estimate to the historical data than both the Log Pearson and Power Law model, as the exponential distribution plots closer to the one-to-one reference line. Also seen from Figure 4.2.6, the discharge of largest flood that has occurred at the Herbert River stream gauge is most accurately predicted by the Power Law relationship.

4.2.4 Carmila Creek at Carmila (126003A)

The Carmila Creek stream gauging station at Carmila is situated in the Plane basin in the Central Queensland region and has a catchment area of 84 km². The mean of the logarithms of the annual peak flows was found to be 0.717, while the standard deviation is 2.247 which is significantly larger than the world non-arid zone's average deviation of 0.15 (Hall 1984). The skew of the logarithms of the annual flood data series for Carmila Creek is -2.16.

The flood series data was checked for low outliers first since the skew of the logarithms is less than -0.4. The low outlier threshold (X_L) was found to be -1.588, which correlates to a peak discharge of 0.026 m³/s. The lowest peak discharge in the annual series is 0.32 m³/s, therefore there is no low outliers for this data series. The high outlier threshold (X_H) was found to be 3.470, correlating to a peak discharge of 2,954 m³/s. The highest peak discharge for the Carmila Creek stream gauge is 1,304 m³/s, hence there are no high outliers in the Carmila Creek annual series data.

The Log Pearson type 3 distribution parameters for the Carmila Creek at Carmila found using MathWave Technologies' Easyfit software is $\alpha = 0.859, \beta = -1.781$ and $\gamma = 6.703$.

The exponential distribution equation was determined from the partial series to be:

$$Q_{PEAK} = 428.8 \log(ARI) + 211.5$$

While the Power Law model isn: $Q_{PEAK} = 261.8(ARI)^{0.375}$. The flood frequency factor of the Stanley River stream gauge is 2.37 and since this frequency factor is relatively small, the station under maritime climate.

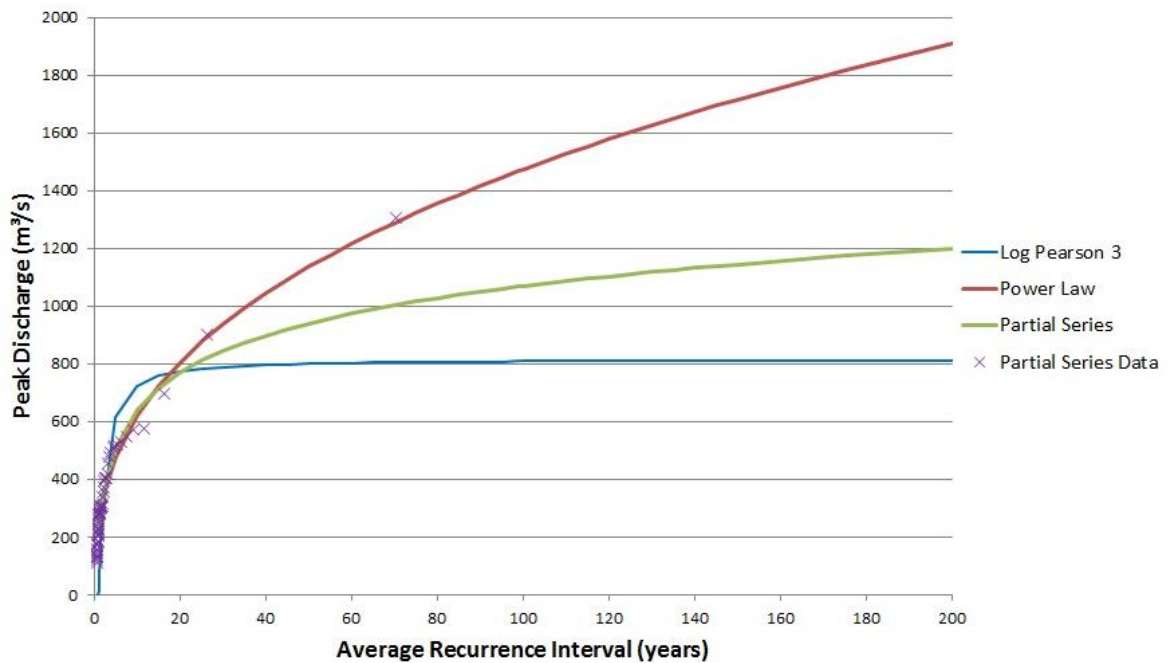


Figure 4.2.7 Carmila Creek stream gauge model predicted peak discharges at Carmila fitted to the average recurrence intervals

It is noted from Figure 4.2.7, that the Power Law model provides that highest predicted discharges for floods that are greater than the 20-year recurrence interval, while unlike the other stream gauge stations, the Log Pearson distribution provides the lowest

predicted discharges. The partial series data point can be seen to be closest to the Power Law model in Figure 4.2.7, especially for the discharges with higher recurrence intervals.

Chi-Squared critical test statistic results found in Table 4.2.7 show that the model provided by the exponential distribution matches more closely with the historical data for the Carmila Creek than the other models as the probability calculated is closest to one. The Log Pearson Chi-Squared test could not be calculated using the Easyfit software, as the standard deviation of the logarithms of the annual flood series is so high.

Table 4.2.7 Carmila Creek results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)		NA
Exponential (partial series)	0.31481	0.989
Power Law (partial series)	0.9465	0.9465

Table 4.2.8 Carmila Creek analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.904	0.802	0.851
Exponential (partial series)	0.961	0.908	0.923
Power Law (partial series)	0.951	0.992	0.981

The R-Squared test over the entire flood series determined the Power Law model most closely fitted with the historical data, as it has the highest value shown in Table 4.2.8. The R-Squared values when the recurrence interval is less than 10 years, show that according to the R-Squared test, the exponential distribution's predicted discharges are the closest match to the historical floods when the recurrence interval is less than 10 years. This supports the Australian Rainfall and Runoff (1987) suggestion that a graphical interpolation of the exponential distribution for floods of recurrence less than 10 years is adequate for making frequency estimates.

The goodness of fit as floods greater than 10 years shows the Power Law model is a closer fit with the historical flood peaks than the other two models.

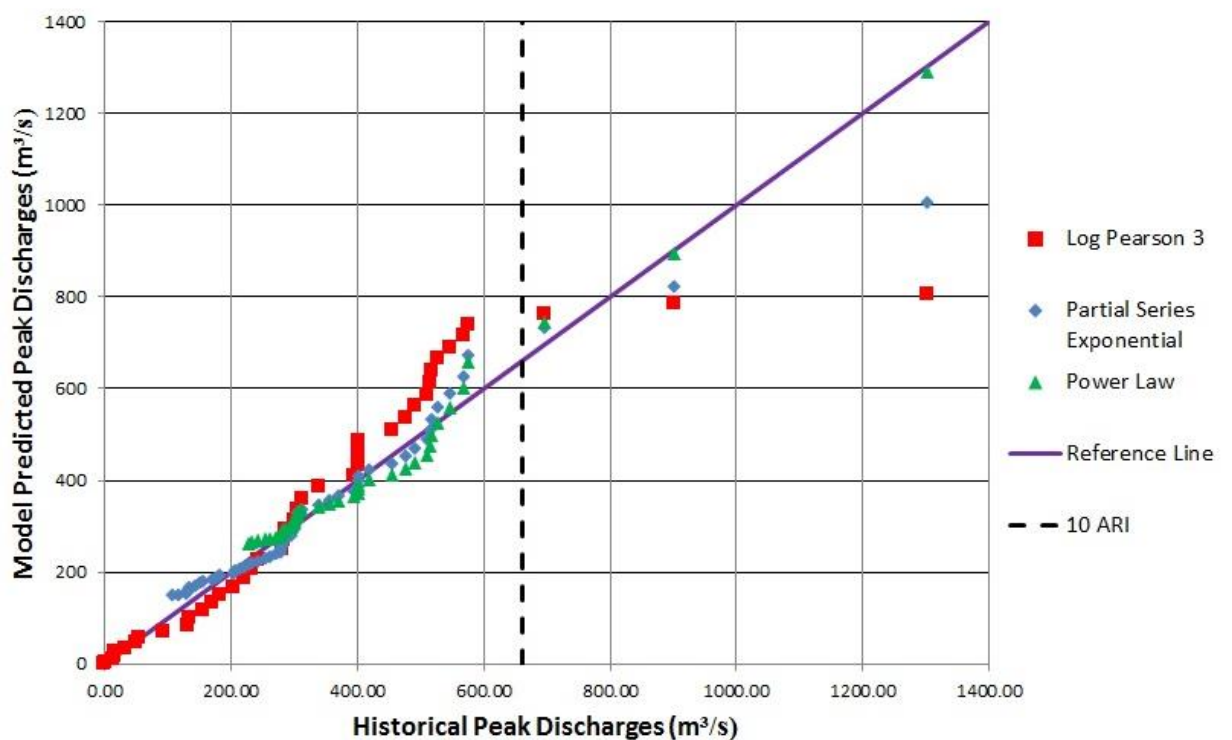


Figure 4.2.8 Carmila Creek stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data

It is identified from Figure 4.2.8, the Power Law relationship (green triangles) and the exponential distribution (blue diamonds) are the closest to the reference line for the discharges less than the 10-year recurrence interval line. While the Power Law model appears to be closer to the reference line for the discharges above the 10-year recurrence interval.

4.2.5 Isis River at Bruce Highway (137201A)

The Isis River stream gauging station at the Bruce Highway is situated in the Burrum basin in the Central Queensland region and has a catchment area of 446 km². The mean standard deviation of the logarithms of the annual peak flows was found to be 0.819 which is within the range of the annual peak flows found in Australia specified in McMahon's article (Hall 1984). The average value of the logarithms of the annual peak flows for the Isis River is 1.994, while the skew is -0.915, meaning of the flood peaks is less than the median value.

The flood series data was checked for low outliers first since the skew of the logarithms is less than -0.4. The low outlier threshold (X_L) was found to be -1.205, correlating to a peak discharge of 0.062 m³/s. The lowest peak discharge in the annual series is 0.364 m³/s, therefore there is no low outliers for this data series. The high outlier threshold (X_H) was found to be 3.754, which correlates to a peak discharge of 5,680 m³/s. The highest historical peak discharge for the Isis River stream gauge is 1,629 m³/s, hence there are no high outliers in the Isis River annual series data.

The Log Pearson type 3 distribution parameters found using MathWave Technologies' Easyfit software is $\alpha = 4.781$, $\beta = -0.862$ and $\gamma = 8.716$.

The exponential distribution expression formed from the partial series data was determined to be: $Q_{PEAK} = 657.8 \log(ARI) + 75.1$

The Power Law model is formed from: $Q_{PEAK} = 214.7(ARI)^{0.5.19}$. The flood frequency factor was calculated as 3.30, meaning there is a factor of 3.30 between the associated 100-year flood to the 10-year flood discharge. Since the flood frequency factor is relatively high and the station under arid climate.

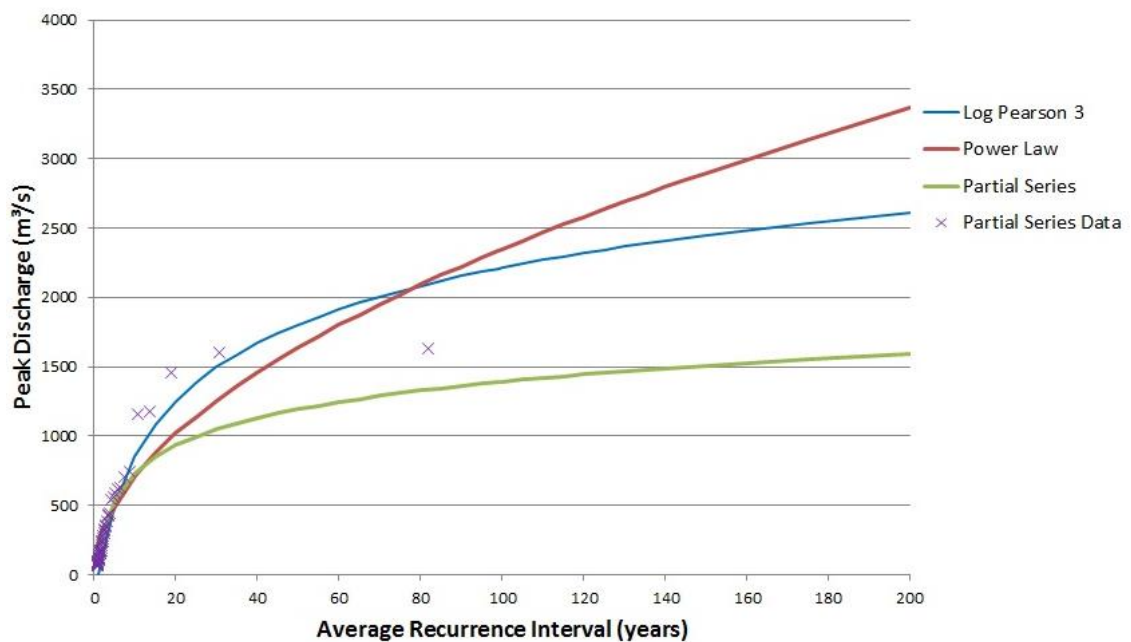


Figure 4.2.9 Isis River stream gauge model predicted peak discharges at the Bruce Highway fitted to the average recurrence intervals

Figure 4.2.9 shows how the three flood frequency models differ over a range of recurrence intervals for the Isis River stream gauge at the Bruce Highway. It should be noted that the Power Law model produces the highest predicted peak discharges for floods that have a recurrence interval greater than 80 years, therefore becoming the more conservative model for the rarer flood events. The exponential distribution of the partial series provides the lowest peak discharges for the same recurrence intervals.

Figure 4.2.9 also shows the data point of the partial series, it can be seen the discharges between 10 years and 40 years are historical higher than any of the model predictions.

Table 4.2.9 shows the results of the Chi-Squared test for the Isis River stream gauge at the Bruce Highway. The lowest critical test statistic value and hence best fitting model out of the three is the Log Pearson 3 distribution, followed by the Power Law model and lastly the exponential distribution. The probability of all three frequency models is below the critical probability of 80 percent and therefore according to the Chi-Squared test, all are poor fits for the Isis River.

Table 4.2.9 Isis River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	2.4792	0.780
Exponential (partial series)	3.2111	0.523
Power Law (partial series)	2.5481	0.636

Table 4.2.10 Isis River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.959	0.939	0.946
Exponential (partial series)	0.981	0.817	0.838
Power Law (partial series)	0.946	0.862	0.835

The R-Squared test results for the Isis River station are shown in Table 4.2.10. The overall R-Squared test which used all of the historical flood years in the analysis found that the Log Pearson type 3 distribution fits the most closely with the historical data, while the Power Law fits the least accurately. This is different to the Chi-Squared test, as according to the Chi-Squared test the Power Law model was more accurate than the exponential distribution. For the R-Squared goodness of fit test which the flood events that have a probability of recurrence in less than ten years, the exponential distribution of the partial series most closely fits with the historical flood peaks. The results also show that when the recurrence interval is greater than ten years, the Log Pearson distribution fits the best out of the three models.

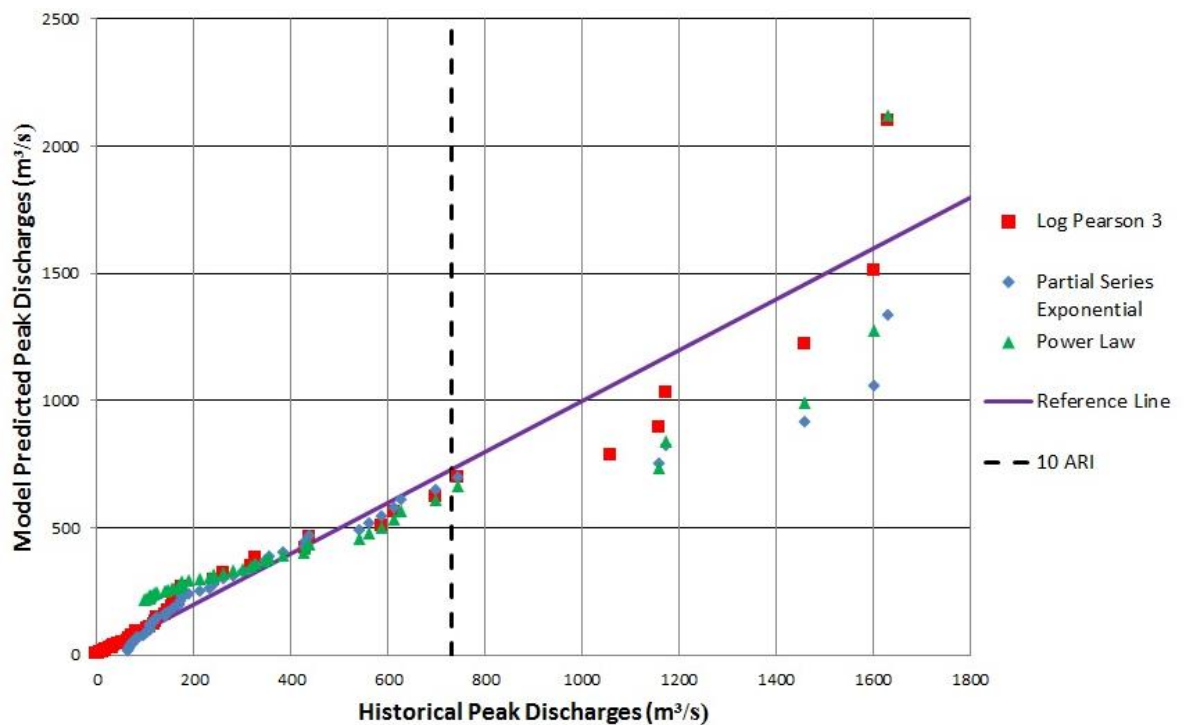


Figure 4.2.10 Isis River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data

Figure 4.2.10 shows the observed peak discharges for the Isis River stream gauge station against each model's predicted peak discharges at the plotting positions assigned in equation (2.3-3). It can be seen that the Log Pearson distribution is closest to the reference line for the peak discharges that are greater than the ten year recurrence line.

4.2.6 Mary River at Miva (138001A)

The Mary River stream gauging station at Miva is situated in the Mary basin in the Central Queensland region and has a catchment area of 4755 km². The Mary River stream gauge provided the largest number of years to analyse, as there is 106 years of flood peak data used from this station. The mean of the logarithms of the annual peak flows was found to be 2.943, while the standard deviation is 0.495 which reinforces McMahon's article stating that the streams in Australia are more variable than the rivers throughout the world (Hall 1984). The skew of the logarithms of the annual flood data series for Mary River stream gauge station at Miva is -0.487.

Since the skew of the logarithms of the annual flood series is less than -0.4, a test for low outliers was commuted first. The low outlier threshold (X_L) was found to be 1.128, which correlates to a peak discharge of 13.413 m³/s. The lowest peak discharge in the annual series for the Mary River at Miva is 12.358 m³/s, with the second lowest peak being 45.537 m³/s. Due the lowest peak being less than the lower threshold, the 12.358 m³/s flood peak was deleted and the frequency analysis was recomputed with 105 years of flood record. After taking the 2006-2007 flood peak out, the mean, standard deviation and skew statistics of the logarithms were recalculated to be 2.961, 0.464 and -0.094 respectively. The high outlier threshold (X_H) was then found to be 4.319, which correlates to a peak discharge of 20,857 m³/s. The highest peak discharge for the Mary

River stream gauge is 7,566 m³/s, hence there are no high outliers in the Mary River annual series data.

The Log Pearson type 3 distribution parameters found using MathWave Technologies' Easyfit software is $\alpha = 453.83$, $\beta = -0.050$ and $\gamma = 29.557$.

The exponential distribution expression formed from the partial series data is determined as: $Q_{PEAK} = 2837 \log(ARI) + 537$

While the Power Law model is formed using the equation: $Q_{PEAK} = 1176(ARI)^{0.435}$ and has a flood frequency factor of 2.72. Since the flood frequency factor is relatively small and the station under maritime climate.

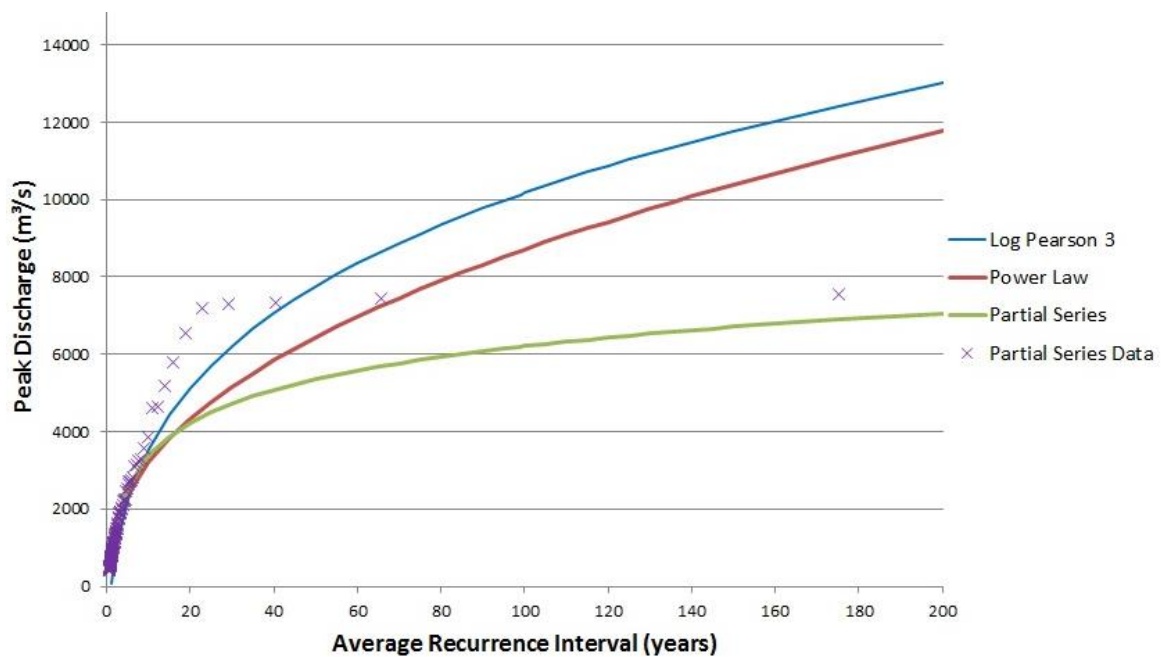


Figure 4.2.11 Mary River stream gauge model predicted peak discharges at Miva fitted to the average recurrence intervals

The three flood frequency models are shown in Figure 4.2.11 against a range of recurrence intervals for the Mary River stream gauge station at Miva. It can be seen in

Figure 4.2.11, the Log Pearson model returns the highest peak discharges and is therefore the most conservative. You can also see the data point of the partial series, it can be seen the discharges between 10 years and 40 years are historical higher than any of the model predictions.

The Chi-Squared goodness of fit test shown in Table 4.2.11 shows the Log Pearson type 3 distribution of the annual data series matches more closely with the historical data for the Mary River station. The exponential distribution of the partial series gave the least accurate fit to the historical flood data. The probabilities given for all three frequency models are below the critical probability of 80 percent and therefore according to the Chi-Squared test, are poor fits for modelling the Mary River flood data.

Table 4.2.11 Mary River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	2.4792	0.780
Exponential (partial series)	12.535	0.051
Power Law (partial series)	4.3104	0.635

Table 4.2.12 Mary River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.997	0.853	0.893
Exponential (partial series)	0.982	0.825	0.866
Power Law (partial series)	0.844	0.806	0.816

The overall R-Squared test shows that the Log Pearson distribution best fits with the historical flood data. It can also be seen that the Log Pearson fit best with the historical data when just looking at the floods that have an expectancy to return at less than 10-year intervals and also when the floods are expected to return at intervals greater than ten years. In all cases, the Power Law model, according to the results in Table 4.2.12, has the least accurate fit.

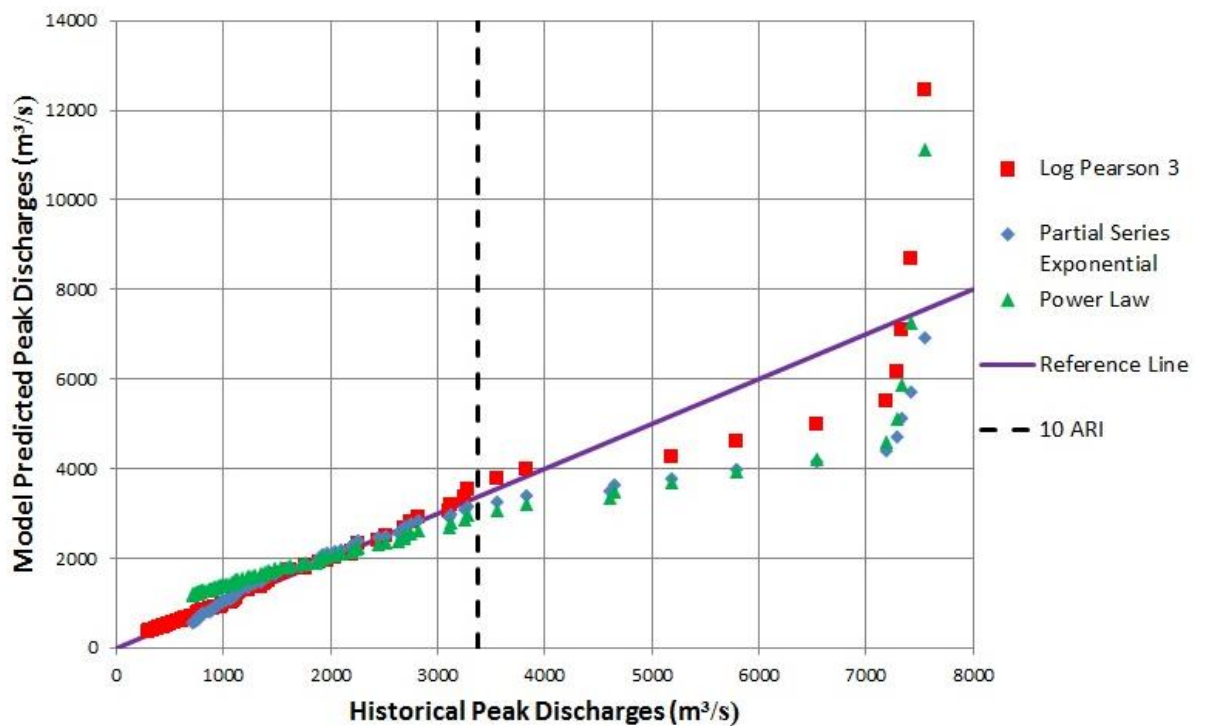


Figure 4.2.12 Mary River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data

The predicted peak discharge results of the three models at the plotting position assigned in equation (2.3-3) are shown in Figure 4.2.12, plotted against the historical peak discharges. When the model predicted peak discharge is equal to the historical discharge, the data point lies on the one-to-one reference line. It can be seen that the

Log Pearson 3 distribution has the data points the closest to the reference line, hence graphically reinforcing the results obtained from the goodness of fit tests.

4.2.7 Caboolture River at Upper Caboolture (142001A)

The Caboolture River stream gauging station at Upper Caboolture is situated in the Pine basin in the South Queensland region and has a catchment area of 94 km². The mean standard deviation of the logarithms of the annual peak flows was found to be 0.525 which is almost twice that of the world non-arid zone average deviation given in Hall (1984). The skew of the logarithms of the annual flood data series for Caboolture River is -1.06, which means the mean of the flood peaks is less than the median value.

The flood series data was checked for low outliers first since the skew of the logarithms is less than -0.4. The low outlier threshold (X_L) was found to be 0.038, correlating to a peak discharge of 1.092 m³/s. The lowest peak discharge in the annual series is 3.77 m³/s, therefore there is no low outliers for this data series. The high outlier threshold (X_H) was found to be 3.274, correlating to a peak discharge of 2,878 m³/s. The highest peak discharge for the Caboolture River stream gauge is 1,055 m³/s, hence there are no high outliers in the Caboolture River annual series data.

The Log Pearson type 3 distribution parameters for the Caboolture River at Upper Caboolture found using MathWave Technologies' Easyfit software is $\alpha = 3.560$, $\beta = -0.6402$ and $\gamma = 7.3093$.

The exponential equation created from the partial series data was calculated to be:

$$Q_{PEAK} = 430.5 \log(ARI) + 141.2$$

While the Power Law relationship is expressed by: $Q_{PEAK} = 204.8(ARI)^{0.396}$, with a flood frequency factor of 2.49. As the flood frequency factor is relatively small, the station under maritime climate conditions.

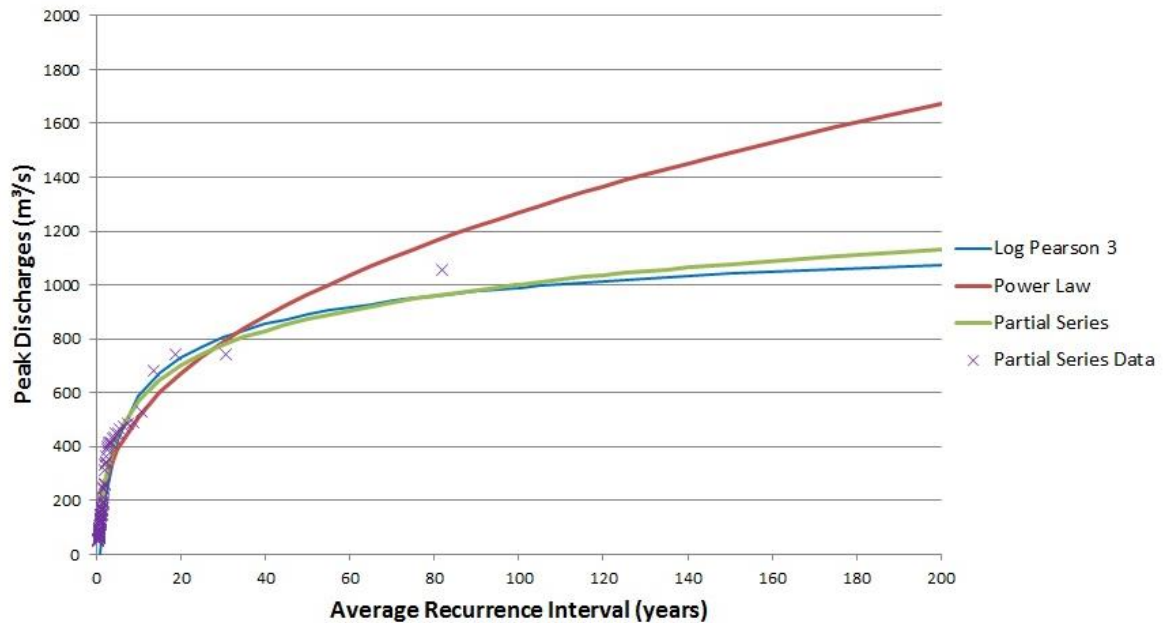


Figure 4.2.13 Caboolture River stream gauge model predicted peak discharges at Upper Caboolture fitted to the average recurrence intervals

Figure 4.2.13 shows how each of the flood frequency models differs over the range of years of recurrence. It is noted that the Power Law model provides that highest predicted discharges for floods that are greater than the 40-year recurrence interval, while the Log Pearson 3 and the exponential distribution provide very similar peak discharges.

Table 4.2.13 below displays the Chi-Squared critical test statistic values for the Caboolture River stream gauge at Upper Caboolture, calculated using equation (3.7-2). The Log Pearson type 3 distribution created with the annual flood series matches more

closely with the historical flood data. While the Power Law model has the highest Chi-Squared test value and therefore has the least accurate fit.

Table 4.2.13 Caboolture River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	1.4052	0.924
Exponential (partial series)	1.7472	0.883
Power Law (partial series)	1.9097	0.861

Table 4.2.14 Caboolture River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	overall
Log-Pearson type 3 (annual series)	0.980	0.986	0.983
Exponential (partial series)	0.930	0.983	0.960
Power Law (partial series)	0.766	0.967	0.880

The R-Squared values for the Caboolture River stream gauge models are shown in Table 4.14. It can be identified that the goodness of fit results from the overall R-Squared test matches the results of the Chi-Squared test, as the closer the R-Squared value is to one, the closer the model matches the historical flood data. Demonstrating that Log Pearson 3 distribution fits with the historical discharges better for the Caboolture river than the other models.

The R-Squared values when the recurrence interval is less than ten years, shows the Log Pearson distribution is the closest match with the smaller flood discharge predictions, closely followed by the exponential distribution. The goodness of fit for the flood greater than ten years also show the Log Pearson fits closest and the Power Law model having the least accurate fit.

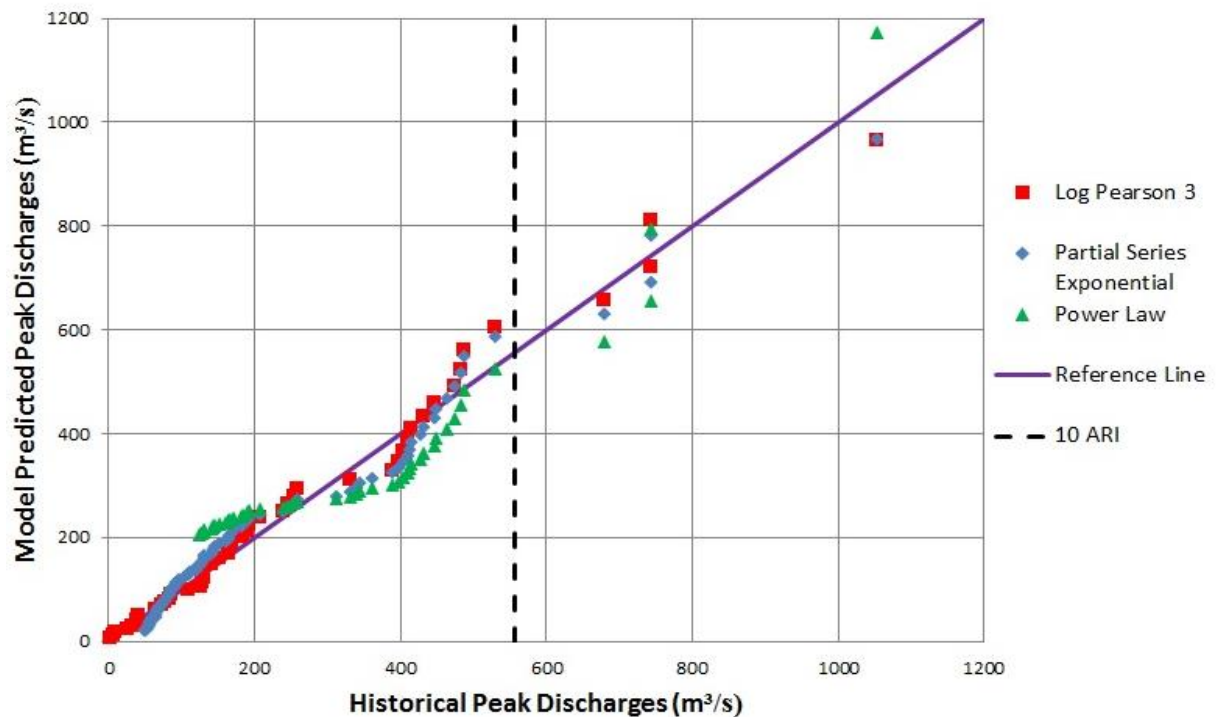


Figure 4.2.14 Caboolture River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data

The graphical demonstration in Figure 4.2.14, shows the predicted peak discharges of each model at the plotting position from equation (2.3-3) compared to the observed historical peak discharges for the Caboolture River stream gauge. It is identified that the Log Pearson distribution, given as the red square points are the closest to the purple reference line. Since the reference line displays the point where the historical discharges is equal to the model predicted peak discharges, the peak discharges predicted by the Log Pearson 3 distribution fits historical data most accurately.

4.2.8 Stanley River at Peachester (143303A)

The Stanley River stream gauging station at Peachester is situated in the Brisbane basin in the South Queensland region and has a catchment area of 104 km². The mean of the logarithms of the annual peak flows was found to be 2.163, while the standard deviation is 0.404 which is twice that of the world non-arid zone's average deviation. The skew of the logarithms of the annual flood data series for Caboolture River is -0.529.

Since the skew is identified as less than -0.4, the flood series data was checked for low outliers first. The low outlier threshold (X_L) was found to be 0.689, which correlates to a peak discharge of 4.88 m³/s. The lowest peak discharge in the annual series is 12.6 m³/s, therefore there is no low outliers for this data series. The high outlier threshold (X_H) was found to be 3.151, which correlates to a peak discharge of 1,417 m³/s. The highest peak discharge for the Stanley River stream gauge is 707 m³/s, hence no high outliers are in the Stanley River annual series data.

The Log Pearson type 3 distribution parameters for the Stanley River at Peachester from the MathWave Technologies' Easyfit software is $\alpha = 14.304$, $\beta = -0.246$ and $\gamma = 8.501$.

The exponential distribution equation is: $Q_{PEAK} = 354.9 \log(ARI) + 112$ and the Power Law model equation is: $Q_{PEAK} = 178.7(ARI)^{0.328}$. The flood frequency factor of the Stanley River stream gauge is 2.13, which is relatively small meaning the station under maritime climate conditions.

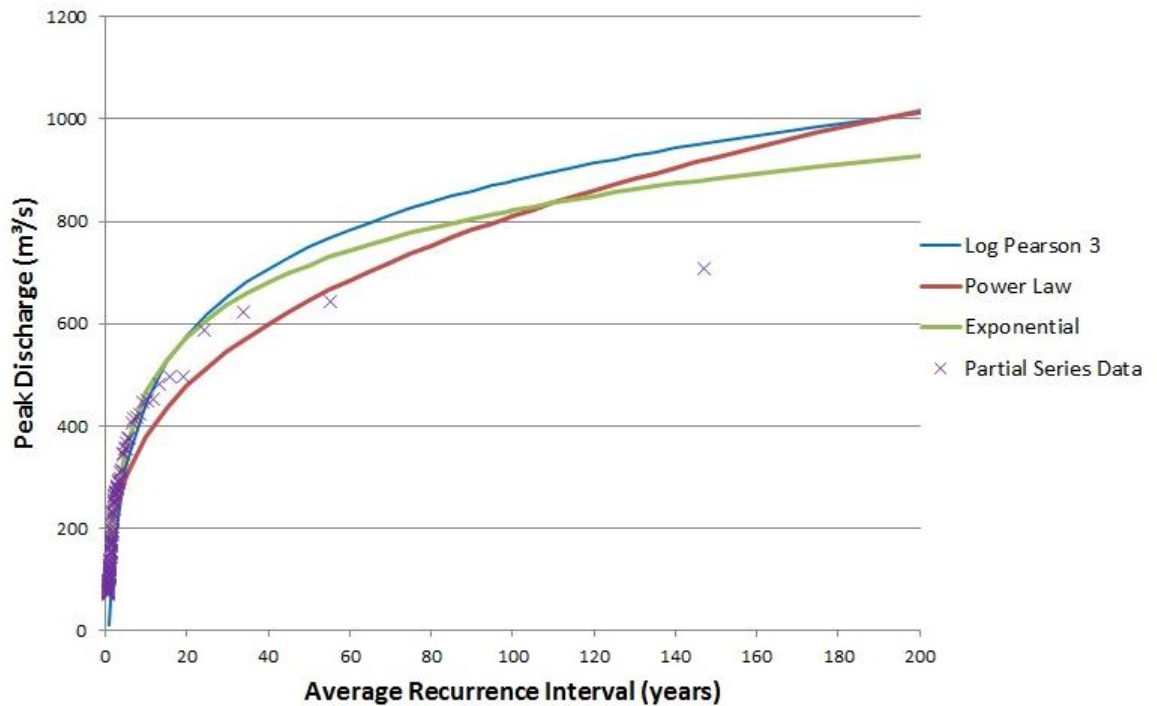


Figure 4.2.15 Stanley River stream gauge model predicted peak discharges at Peacheater fitted to the average recurrence intervals

The peak discharges from each flood frequency model is shown in Figure 4.2.15, over a range of recurrence years. It should be noted that the Power Law model provides the highest predicted peak discharges for the floods that have a recurrence interval greater than 200 years and as expected the exponential distribution provides the lowest predicted discharges for the larger flood events

Table 4.2.15 clearly shows the model provided by the Log Pearson 3 distribution has the lowest statistic value and hence the highest probability, indicating the Log Pearson 3 distribution most accurately fits with the historical flood data according to the Chi-Squared test. The Power Law has the least accurate fit.

The probability of all three frequency models is below the critical probability of 80 percent and therefore according to the Chi-Squared test, should be rejected as an effective fit for the Isis River.

Table 4.2.15 Stanley River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	5.4608	0.486
Exponential (partial series)	5.4523	0.487
Power Law (partial series)	10.096	0.121

Table 4.2.16 Stanley River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	Overall
Log-Pearson type 3 (annual series)	0.987	0.921	0.955
Exponential (partial series)	0.988	0.940	0.948
Power Law (partial series)	0.776	0.916	0.850

The R-Squared test using the entire flood series matches with the results of the Chi-Squared test, in which the Log Pearson 3 distribution most closely fits with the historical data for the Stanley River stream gauge. The R-Squared test completed of the floods with a recurrence interval of less than 10 years, shows the exponential distribution predicts the most accurate discharges, closely followed by the Log Pearson 3 distribution. The goodness of fit when the recurrence interval is greater than 10 years,

displays the Log Pearson 3 distribution as the best fitting model, but very closely followed by the Power Law model.

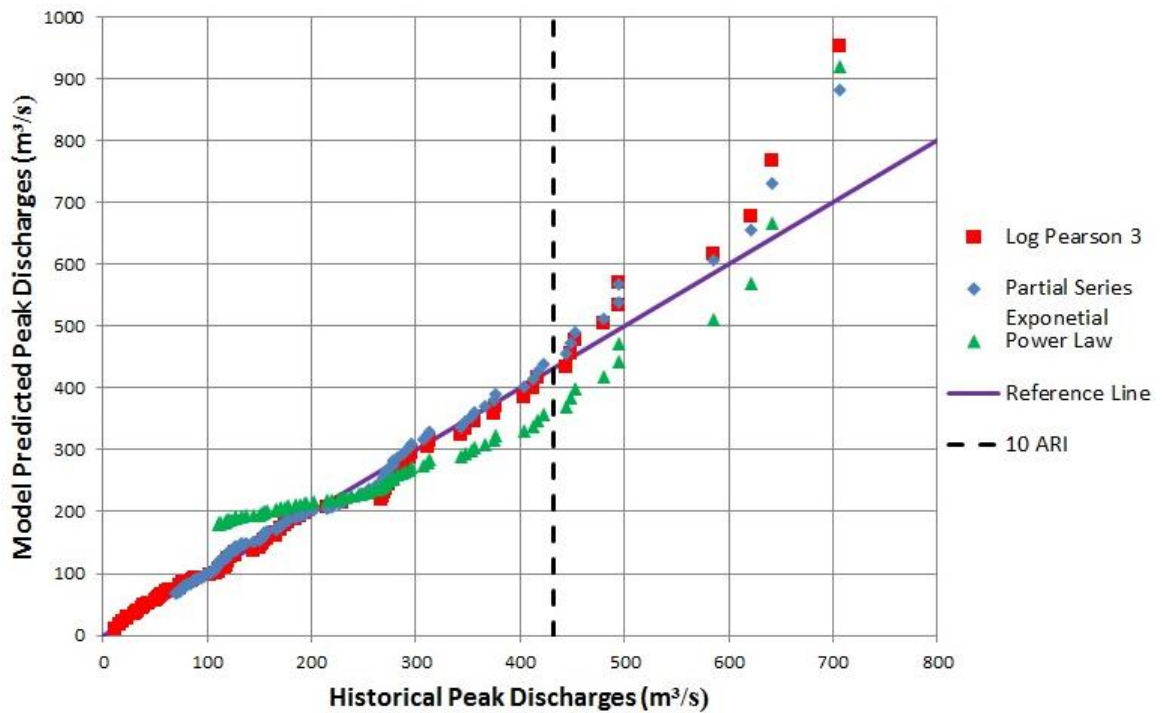


Figure 4.2.16 Stanley River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data

The results discussed in Table 4.2.15 and Table 4.2.16 are shown graphically in Figure 4.2.16. It is identified from Figure 4.2.16, the Log Pearson 3 distribution (red squares) is clearly the closest to the purple reference line, especially for the discharges less than the ten year recurrence interval line. For the discharges that are over the ten year recurrence line, the Power Law model generally predicts under-estimates the peak discharge of the flood, while the exponential distribution and the Log Pearson 3 distribution are more conservative and over-estimate the peak discharges.

4.2.9 Coomera River at Army Camp (146010A)

The Coomera River stream gauging station at Army Camp is situated in the South Coast basin in the South Queensland region and has a catchment area of 88 km². The mean of the logarithms of the annual peak flows was found to be 1.961, while the standard deviation is 0.684 which is much greater than the world non-arid zone's average deviation of 0.15 (Hall 1984). The skew of the logarithms of the annual flood data series for Caboolture River is -0.951.

The flood series data was checked for low outliers first since the skew of the logarithms is less than -0.4. The low outlier threshold (X_L) was found to be -0.769, correlating to a peak discharge of 0.17 m³/s. The lowest peak discharge in the annual series is 0.955 m³/s, therefore no low outliers are present in this data series. The high outlier threshold (X_H) was found to be 3.422, which correlates to a peak discharge of 2,642 m³/s. The highest peak discharge for the Coomera River stream gauge is 905 m³/s, hence there are no high outliers in the Coomera River annual series data.

The Log Pearson type 3 distribution parameters for the Coomera River at Army Camp found using MathWave Technologies' Easyfit software as: $\alpha = 4.42$, $\beta = -0.749$ and $\gamma = 7.83$.

The exponential distribution equation to determine the peak discharges formed from the partial series is: $Q_{PEAK} = 391 \log(ARI) + 69.1$

While the Power Law model equation is: $Q_{PEAK} = 137(ARI)^{0.470}$. The flood frequency factor is calculated as 2.95 and since the factor is relatively small, the station is under maritime climate conditions.

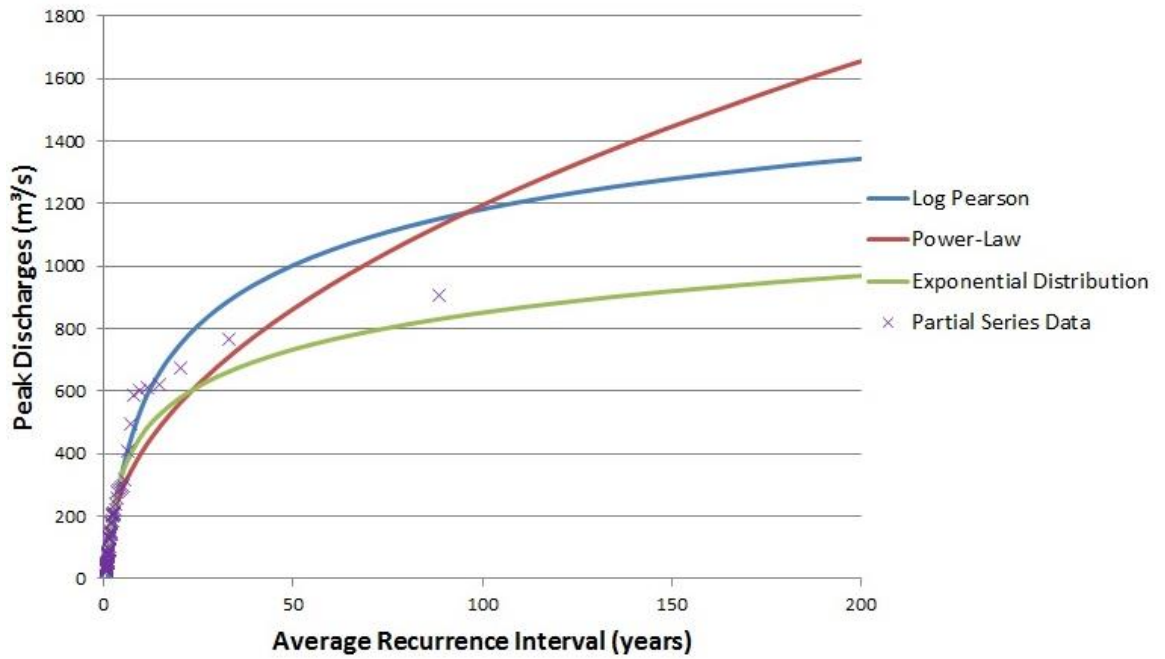


Figure 4.2.17 Coomera River stream gauge model predicted peak discharges at Army Camp fitted to the average recurrence intervals

It should be noted from Figure 4.2.17, the Power Law model provides the highest predicted discharges for floods that are greater than the 100-year recurrence interval, while the exponential distribution processes the lowest predicted discharges.

Table 4.2.17 gives the Chi-Squared critical test statistic values and can clearly identify the Log Pearson 3 distribution matches more closely with the historical flood data. The Power Law model on the other hand, is the model with the least accurate fit for the partial series.

Both the exponential distribution and the Power Law provide have a probability less than 80 percent, therefore according to the Chi-Squared test these models should be rejected as a poor fit for the Coomera River.

Table 4.2.17 Coomera River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	1.8752	0.866
Exponential (partial series)	4.4329	0.489
Power Law (partial series)	4.1341	0.388

Table 4.2.18 Coomera River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	Overall
Log-Pearson type 3 (annual series)	0.973	0.938	0.954
Exponential (partial series)	0.906	0.960	0.937
Power Law (partial series)	0.754	0.906	0.841

The R-Squared values for the Coomera River station are shown in Table 4.2.18, which found that the overall R-Squared was highest with the Log Pearson 3 distribution, same with the flood event with recurrence intervals greater than ten years. The exponential distribution had the better fit for the more common occurring flood events.

According to both the Chi-Squared test and all of the R-Squared tests, the Power Law relationship model for the Coomera River stream gauge has the least accurate fit to the historical flood data.

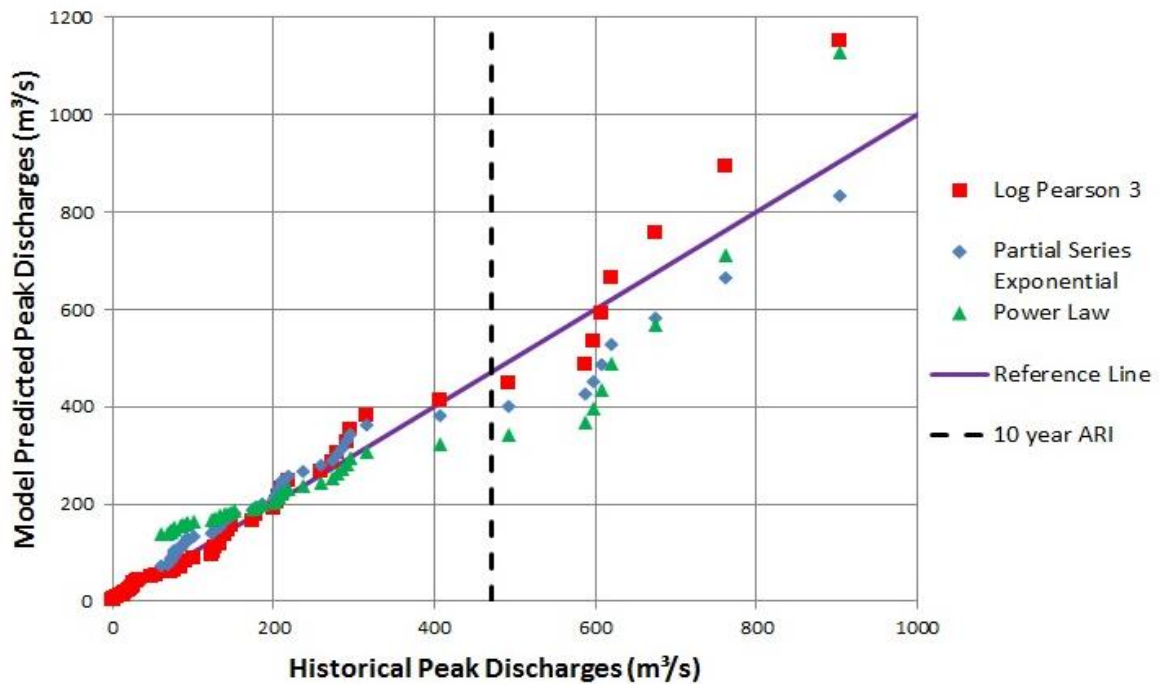


Figure 4.2.18 Coomera River stream gauge graphical comparison of the flood model’s peak predicted discharges compared to the observed historical data

It is identified from Figure 4.2.18, the Log Pearson distribution, given as the red squares, is the closest to the purple reference line for the floods less than the 10-year average recurrence interval line. For the discharges greater than the 10-year line, the model which is closest varies. The Log Pearson 3 distribution is generally the closer fit to the predicted values for the flood discharges between 500 m³/s and 700 m³/s, but as the discharge increases over 700 m³/s the Log Pearson 3 over-estimates the discharges.

4.2.10 Condamine River at Elbow Valley (422394A)

The Condamine River stream gauging station at Elbow Valley is situated in the Balonne-Condamine basin in the South Queensland region and has a catchment area of 325 km². The mean of the logarithms of the annual peak flows was found to be 1.561,

while the standard deviation is 0.663 which is much greater than the world non-arid zone's average deviation and the skew of the logarithms of the annual flood data series for Caboolture River is -0.166, which means the mean of the flood peaks is less than the median value.

The flood series data was firstly checked for high outliers, since the skew of the logarithms is between -0.04 and 0.4. The high outlier threshold (X_H) was found to be 3.318, which correlates to a peak discharge of 2077 m³/s. The highest peak discharge for the Condamine River stream gauge is 790 m³/s; hence there are no high outliers for this data series. The low outlier threshold (X_L) was found to be -0.400, which correlates to a peak discharge of 0.398 m³/s. The lowest peak discharge in the annual series is 0.553 m³/s, therefore there is no low outliers in the Condamine River annual series data.

The Log Pearson type 3 distribution parameters for the Condamine River at Elbow Valley found using MathWave Technologies' Easyfit software is $\alpha = 145.2$, $\beta = -0.127$ and $\gamma = 21.98$.

The exponential distribution equation formed from the partial series data was determined to be: $Q_{PEAK} = 137 \log(ARI) + 31.0$ and the Power Law model is formed from the equation: $Q_{PEAK} = 50.4(ARI)^{0.675}$. The flood frequency factor of the Stanley River stream gauge is 4.73 and being a relatively high frequency factor, the station is under arid climate conditions.

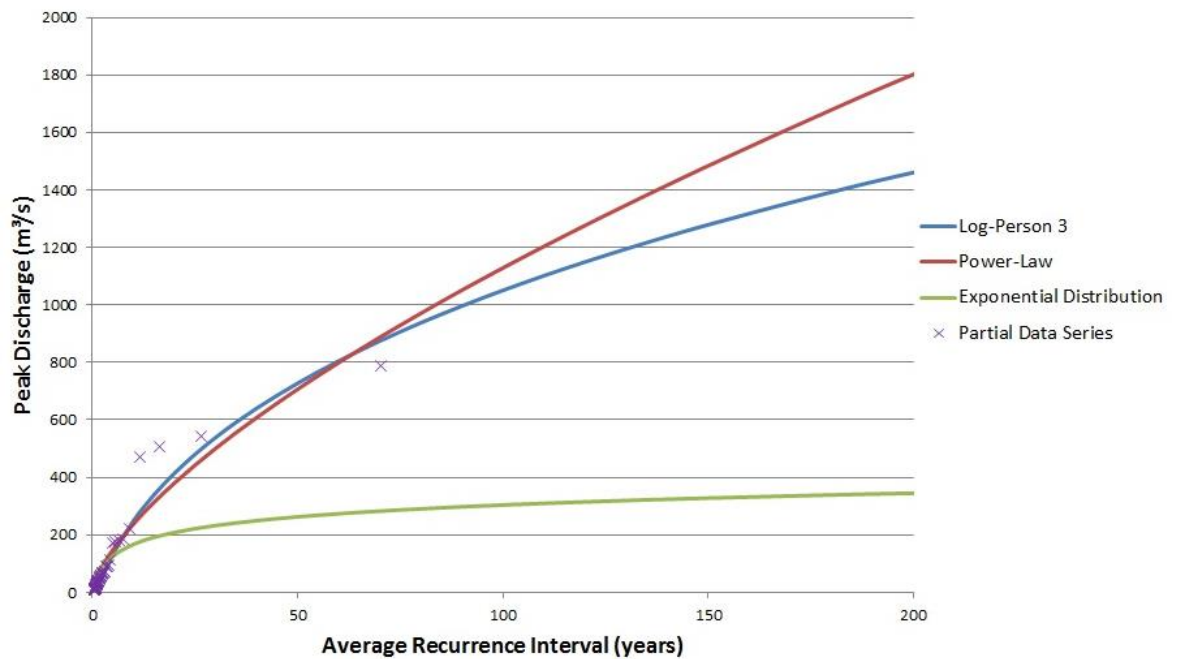


Figure 4.2.19 Condamine River stream gauge model predicted peak discharges at Elbow Valley fitted to the average recurrence intervals

The three flood frequency models have been shown in Figure 4.2.19, plotted against recurrence intervals for the Condamine River stream gauging station at Elbow Valley. It is clearly identifiable that the Power Law model returns the highest peak discharges for the recurrence intervals that are greater than 80 years and therefore making the Power Law model the more conservative model. It can also be seen that the exponential distribution produces the lowest peak discharges and therefore well under-estimates the

The Chi-Squared test results in Table 4.2.19, identified that the Log Pearson 3 distribution most accurately predicts the historical flood data for the river, while the exponential distribution provides the least accurate and with a probability of less than 80 percent.

Table 4.2.19 Condamine River results of Chi-Squared test

Model	Critical Test Statistic Value, χ^2	Probability
Log-Pearson type 3 (annual series)	1.3631	0.928
Exponential (partial series)	8.8596	0.012
Power Law (partial series)	1.3992	0.844

Table 4.2.20 Condamine River analysis of R-Squared values

Model	< 10-year ARI	> 10-year ARI	Overall
Log-Pearson type 3 (annual series)	0.989	0.931	0.942
Exponential (partial series)	0.909	0.402	0.488
Power Law (partial series)	0.880	0.900	0.897

The overall R-Squared test results in Table 4.2.20 agree with the results from the Chi-Squared test, hence determining that the Log Pearson 3 distribution most accurately fits with the historical data for the Condamine River. In fact the Log Pearson 3 distribution processes the highest R-Squared values for all three R-Squared tests, making it the most accurate model for the general trend of the Condamine River.

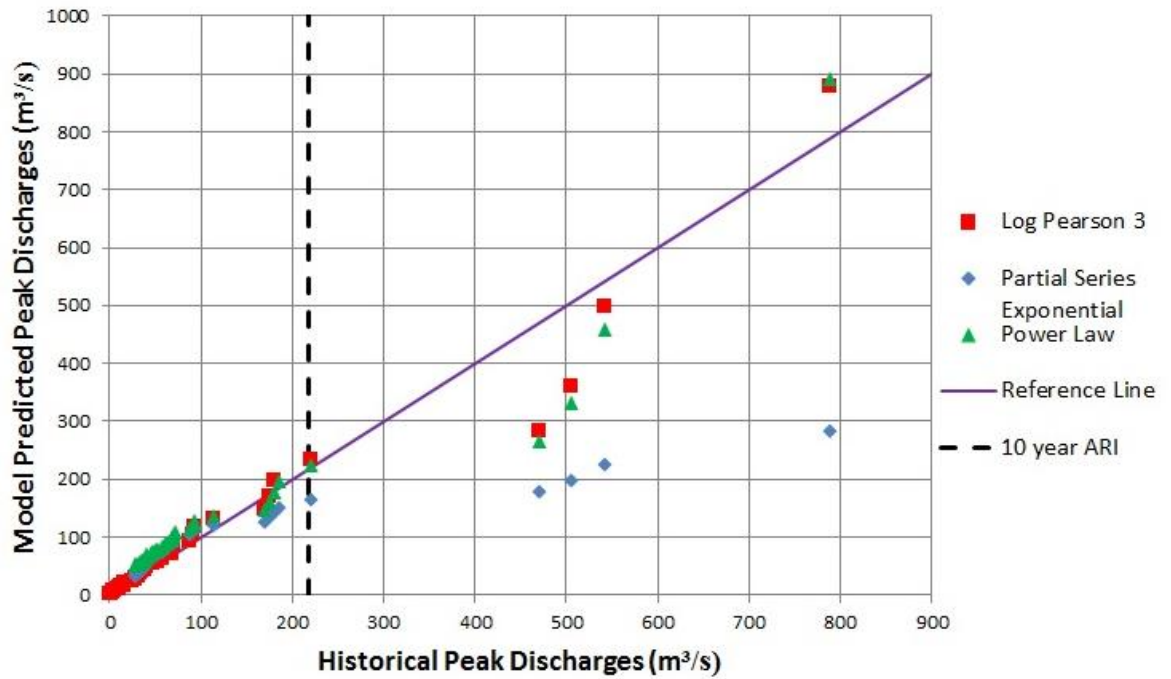


Figure 4.2.20 Condamine River stream gauge graphical comparison of the flood model's peak predicted discharges compared to the observed historical data

To reinforce the results obtained from the goodness of fit tests for the Condamine River stream gauge, Figure 4.2.20 shows a graphical representation of the difference between the peak discharges calculated by the models from the plotting positions assigned in equation (2.3-3) and the observed peak discharges. Clearly identifiable in Figure 4.2.20, is that the data point for the Log Pearson 3 distribution is plotted closer to the reference line than the exponential distribution and the Power Law model.

4.3 Summary

In this chapter, the results of each individual stream gauging station were discussed. For each of the ten stations analysed across Queensland, the three parameters (α , β and γ) were identified for the Log Pearson distribution as well as the regression coefficients for both the conventional exponential distributed partial series and the Power Law model. The first figures provided for each station gives each model plotted against recurrence intervals in years. As the recurrence interval increases, the Power Law becomes the most conservative model and predicts the highest peak discharges for the extreme events. Therefore supporting the findings from Kidson et. al (2006). The flood frequency factors were also identified from the Power Law model, relating the discharges associated with the 100-year flood to the discharge associated with the 10-year flood. These factors ranged from 2.13 to 4.73.

This chapter has also looked the goodness of fit of all three models to the historical peaks. Identifying which model fits best for each individual station by using the Chi-Square goodness of fit test and also the R-Squared test. In chapter 5, Table 5.2.1 has been provided that summarises the best fitting model for each station according to both goodness of fit tests.

Chapter 5 will also examine the results of each frequency model and provide a greater level of discussion and analysis. Each model will be analysed in terms of effectiveness of flood frequency analysis and the possible errors in analysis and limitations of the models and goodness of fit tests will be noted.

Chapter 5 – Discussion

5.1 Chapter Overview

Following on from the results detailed in Chapter 4, this chapter will provide a greater level of discussion and analysis of the overall effectiveness of each flood frequency model.

Firstly exponential distribution of the partial series is looked at, identifying its effectiveness for the ten Queensland stream gauge stations and in particular for the floods that have a recurrence interval of less than ten years. The Log Pearson 3 distribution will then be looked at in a similar approach, however with a greater focus on its effectiveness in predicting the larger floods.

The Power Law relationship models will then also be analysed for its effectiveness in predicting the frequency of floods compared to the current conventional methods used in Queensland.

Finally, the possible errors in the goodness of fit tests will be discussed and also the limitations of flood frequency analysis will be noted.

5.2 Summary of Results

The key outputs of the results from the previous chapter have been summarised in Table 5.2.1 for convenience. The table displays which of the three frequency models provided the highest test value for the Chi-Squared test and the multiple R-Squared tests. In the last column of Table 5.2.1 the model which most accurately predicted the maximum historical flood on record is displayed.

Table 5.2.1 Best fitting model of each stream gauge station found from the goodness of fit tests

Station	Chi-Squared Test	R-Squared Test			Maximum Observed Flood
		< 10-year ARI	> 10-year ARI	overall	
Coen River at Racecourse	Exponential	Exponential	Exponential	Exponential	Exponential
Fisher Creek at Nerada	Log-Pearson	Log-Pearson	Log-Pearson	Log-Pearson	Log-Pearson
Herbert River at Abergowrie	Exponential	Exponential	Exponential	Exponential	Power Law
Carmila Creek at Carmila	Exponential	Exponential	Power Law	Power Law	Power Law
Isis River at Bruce Highway	Log-Pearson	Exponential	Log-Pearson	Log-Pearson	Log-Pearson
Mary River at Miva	Log-Pearson	Log-Pearson	Log-Pearson	Log-Pearson	Power Law
Caboolture River at Upper Caboolture	Log-Pearson	Log-Pearson	Log-Pearson	Log-Pearson	Power Law
Stanley River at Peachester	Exponential	Exponential	Exponential	Log-Pearson	Log-Pearson
Coomera River at Army Camp	Log-Pearson	Log-Pearson	Exponential	Log-Pearson	Power Law
Condamine River at Elbow Valley	Log-Pearson	Log-Pearson	Log-Pearson	Log-Pearson	Log-Pearson

5.3 Exponential Distribution

It is noted in the Australian Rainfall and Runoff (1987) guide that a graphical interpolation for the exponential distribution of the partial series is a simple and safe method for determining the design floods with relatively low average recurrence intervals. Using the R-Squared test, the accuracy of this exponential distribution for the floods occurring every ten or less years was able to be tested. For all cases of these distribution, the R-Squared value for floods less than ten years was greater than 0.90, giving the exponential distribution at least a 90 percent accuracy in predicting these smaller peak discharges. According to the University Corporation for Atmospheric Research (2010), flood return periods for the ten year floods should have no more than 10 percent error. However, there should be at least 90 years of flood data and when there is less data available, the higher the percentage error is expected. Since the error according to the R-Squared test for the floods less than ten year is less than 10 percent and majority of these stream gauges have less than 90 years of data available, a graphical interpolation for the exponential distribution is suitable for the predicted of flood discharges of floods with a recurrence interval of less than ten years.

This R-Squared test found that five of the ten stream gauge stations had the exponential distribution providing the highest R-Squared value for the floods with recurrence intervals of less than ten years. These stations were Coen River at Racecourse, Herbert River at Abergowrie, Carmila Creek at Carmila, Isis River at the Bruce Highway and Stanley River at Peachester. The other five stations had the exponential as a close second. The design flood estimate of low recurrence intervals are rarely used in routine design and rarely cause large amount of damage to infrastructure, the peak discharge

predictions of these floods of recurrence intervals less than ten year is not as important in flood frequency analysis as the larger, less common floods. Hence the method of fitting an exponential distribution to the partial series data for the more common floods is effective and reliable if the peak of a flood of less than ten year recurrence is needed.

Both the Coen River and Herbert River were from to the closest fit to the historical data using both the full R-Squared test and the Chi-Squared test. When looking the graphical comparison for both these stream gauges, it can be seen the Log Pearson distribution and the Power Law model overestimate the discharges. Since the analysis for the Coen River is only considers 47 years of flood data and the Herbert River considers 45 years, the accuracy of the plotting positions of the historical floods would be improved with a greater number of years of historical flood data. Since there is always the possibility of a larger flood occurring in the future, greater than the current highest historical flood on record, using the Power Law and Log Pearson 3 models to determine the design flood for infrastructure design opposed to the exponential distribution would therefore provide a more conservative estimate of peak river discharge for this station.

5.4 Log Pearson 3 Distribution

The national adopted approach for flood frequency analysis uses the Log Pearson type 3 distribution to predict the peak discharge of future flood events. This study has used the Chi-Squared goodness of fit test and the R-Square test to identify the effectiveness of this probability model to the exponential distribution and the Power Law relationship

model. The results show that the Log Pearson 3 distribution appears to on the large part fit with the historical data, as six of the stations were found to have the highest possibility using the Chi-Squared test and seven stations using the R-Squared test with all the years of flood data. The Chi-Squared test found that five of these stations had an 80 percent or higher probability that the historical data is modelled by a Log Person 3 distribution. These results support the studies conducted by Kolittle et al. (Rahman, Haddad & Rahman 2014) and Boughton (1975) in stating the Log Pearson type 3 distribution is the most suitable for Queensland stream gauges.

The main objection to this approach, as identified in Section 2.3.1 is that the Log Pearson distribution cannot be fitted to predict both the outliers and the general trend of the peak discharges. As discussed, the Log Pearson 3 has most accurately predicted the peak discharges for majority of the chosen stream gauge stations. When looking at the R-Squared test to evaluate the goodness of fit of the Log Pearson 3 distribution to predict the flood event with a recurrence interval greater than ten years, the R-Squared test found that five of the stations had the highest value with the Log Pearson 3 model when compared to the goodness test results of the Power Law and exponential models. However, these floods are still part of the general trend of the data; therefore the maximum flood for each station should be looked at as the outlier event. From the graphical comparison figures shown in the results chapter, from the ten stream gauge station, seven of the stations were found that the Log Pearson distribution predicted the highest historical discharge most poorly by overestimated the peak discharge. These stream gauge stations were: Coen River at Racecourse, Herbert River at Abergowrie, Carmila Creek at Carmila, Isis River at the Bruce Highway, Mary River at Miva, Caboolture River at Upper Caboolture and Stanley River at Peachester. The last four of these stations also found R-Squared values for the floods greater than ten years, have

the Log Pearson distribution as more accurate than the Power Law, yet the Power Law model predicts the largest historical flood for the station more accurately. This reinforces Cooper's study (2005) in stating that the Log Pearson 3 distribution most commonly cannot fit with both the outlier events and the general trend, hence typically producing poor estimates of the extreme flood events.

5.5 Power Law Model

The return period of flood events has been said to follow a simple Power Law relationship and have been suggestion in recent literature to be the effective model in estimating event floods.

From the figures in the results chapter showing the frequency models distribution against the average recurrence interval, it can be seen that generally the Power Law model provided the higher predicted discharges for the flood that are less frequent to occur (i.e. 100-year flood event). This means that the design flood used in infrastructure design will be more conservative if the Power Law model is used for the flood frequency study. A conservative estimate for the design flows also recognises that the climate in the future may change to conditions that are not like current conditions with larger floods occurring more and more frequently.

The critical difference between the Power Law method and the conventional distributions is that the Power Law assumes a straight line rather than the Log Pearson's concave down scaling in log-log space. Since the plotted data of the flood discharges

with their recurrence intervals given by the plotting position generally concave downwards, the Log Pearson 3 distribution generally has a better fit for the historical floods for the majority of the flood series.

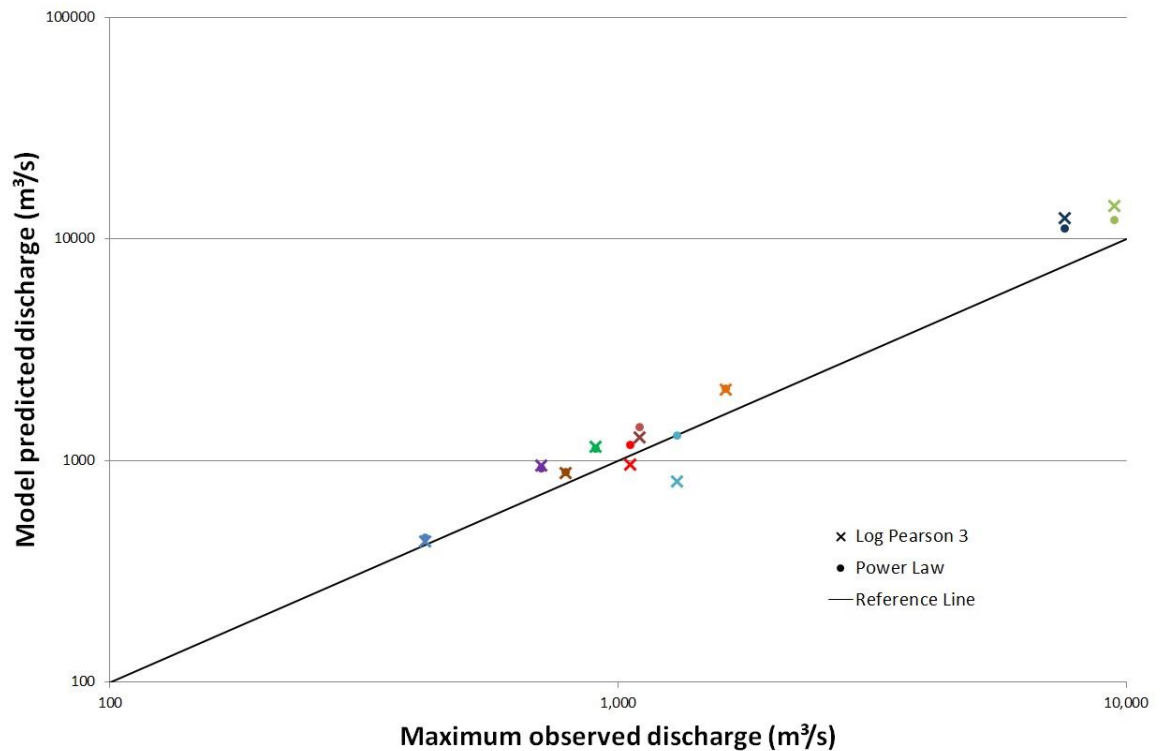


Figure 5.5.1 Graphical comparison of the Power Law model and the Log Pearson 3 distribution discharge predictions for the largest event compared to the observed largest event, for the ten stream gauge stations over Queensland.

Figure 5.5.1 provides the highest discharges for each of the stream gauging stations with their plotting positions and compares it to the model predicted discharges at the same recurrence interval as specified by the plotting position. The crosses display the results using the Log Pearson distribution, while the circles show the Power Law model. Each stream gauge station is also represented by different colours. The Mary River and Herbert River stream gauge station are the points nearest the 10,000 m³/s discharges. From these two stations, it can be seen that the Power Law model provides a closer estimate of historical peak discharges than the Log Pearson 3 distribution as the data

points are closer to the one-to-one reference. The rest of the discharges from the other eight stations (around the 1,000 m³/s discharge) are more random as to which of the two frequency models better predicted the historical data. Only the light blue data points, which are from the Carmila stream gauge station has a noticeable difference between the two model predicted values. Figure 5.5.1 suggests that the Power Law model is more accurate in predicting the extreme flood event's discharges when the average flow of the river is high, as both the Mary River and Herbert River have their average river flow higher than the other eight stations in this analysis. The repeatability of this result across a larger number of stations could be investigated to determine if this generalized pattern is accurate or a unique to individual stream gauge stations.

The partial flood series and the annual flood series were determined to be similar for the larger floods, but significantly different for the smaller floods. As the Power Law model predicts the larger floods, the Power Law model could be used with the partial series or the annual series as both flood series' are similar for the larger events.

5.6 Limitations

In this study, the flood series data are both plotted using the unbiased plotting position formula provided by Cunnane. These plotting positions serve as an estimate of the probability of exceedence and allow a visual examination of fit provided by the flood frequency models. According to articles by Shabri (2002), Ewemoje and Ewemooje (2011) and Adeboye and Alatise (2007), distributions show a best fit to specific plotting positions. The choice of a plotting position formula is the same as choosing an

underlying probability distribution. The goodness of fit tests compares the outputs of the distribution models to the historical peak flood and their associated plotting position. Had the Weibull plotting position been used instead of the Cunnane plotting position, the position of the higher flood may have been 75 years instead of the 80 years and therefore each of the flood frequency models may have fitted better with the Weibull plotting position. While these goodness of fit tests looks to determine the best fitting frequency model for each station with the Cunnane plotting position as its bases flood distribution, there may be a better fitting distribution using a difference plotting position for the station.

The overall R-Squared test is heavily weighted on the smaller floods that occur frequently as there is more of them in the flood series. Since the Power Law model is used to predict the larger and less frequent floods (Kidson, Richards & Carling 2006), the overall R-Squared value is not as accurate as the Chi-Squared test in identifying the model's effectiveness for the entire flood series.

The results must also be considered for measurement errors in discharge estimates, particularly for outlier events, which is often estimated by extrapolation from a rating curve. Kidson, Richards and Carling (2006) suggests that a 20 percent error in measurement is quite common and since a record length of 50 years only provides five plotting positions with a recurrence interval of greater than ten years which to base a reliable regression, there is a diagnostic error in determining the effectiveness of the Power Law model. Predictive models for determine the magnitude verse the frequency are less reliable when the record period is short, which is a truism frequently mentioned in the literature.

For each of the stream gauge stations, the best fitting distributions according to the Chi-Squared test were found using the Easyfit software. This results of this analysis are shown in Table 5.6.1. The results clearly show there is no one distribution that matches all the stream gauges or majority of them. Instead, the frequency of each river's flood events should be analysed for each station to verify the best fitting distribution. This clearly shows the limitations of the flood frequency distribution models as there is no correct answer for the best model for the Queensland stream gauges.

Table 5.6.1 The best fitting distribution for the annual and partial flood data series according to the Chi-Squared test performed using Mathwave's Easyfit software

Stream Gauge Station	Annual Series		Partial Series	
	Name	Probability	Name	Probability
Coen River at Racecourse	Nakagami distribution	0.961	Johnson SB distribution	0.999
Fisher Creek at Nerada	Frechet distribution	0.984	Pearson type 6 distribution	0.941
Herbert River at Abergowrie	Pearson type 6 distribution	0.975	Fatigue Life distribution	0.951
Carmila Creek at Carmila	Cauchy distribution	0.996	Exponential distribution	0.989
Isis River at Bruce Highway	Pareto 2 distribution	0.956	Fatigue Life distribution	0.983
Mary River at Miva	Phased Bi-Exponential distribution	0.953	Frechet distribution	0.987
Caboolture River at Upper Caboolture	Pearson type 6 distribution	0.977	Nakagami distribution	0.995
Stanley River at Peachester	Generalized Extreme Value distribution	0.892	Gamma distribution	0.705
Coomera River at Army Camp	Generalized Gamma distribution	0.938	Fatigue Life distribution	0.977
Condamine River at Elbow Valley	Generalized Logistic distribution	0.967.	General Extreme Value distribution	0.859

While Table 5.6.1 shows the best fitting distribution using the Chi-Squared goodness of fit test, if the Kolmogorov-Smirnov or Anderson-Darling test was used instead, the results would be completely different.

5.7 Summary

This chapter has delivered a greater insight into the use of the exponential distribution, Log-Pearson type 3 distribution and the Power Law relationship on the ten selected stream gauge stations within Queensland.

It has noted that the exponential distribution of the partial series data for the station is effective and reliable for the prediction of peak discharges of flood events that are estimated to occur in time periods of less than ten years. While the Log-Pearson type 3 distribution, has the highest probability of predicting the peak discharges for the less commonly occurring flood events.

It was also found that this analysis has supported suggestions made by Cooper (2005) in 'Estimation of Peak Discharges for Rural, Unregulated Streams in Western Oregon', in stating that the Log Pearson 3 distribution produces poor estimates of the high outlier events and general trend of the peak discharges. As it was found that the Power Law relationship model produced a more accurate prediction of the largest historical flood on record for a number of the stream gauge stations, particularly for the rivers that have an average flow that is high.

The limitations of this study were then included, taking note of the effects the choice of plotting position formula and goodness of fit tests have on the analysis results.

Finally the best fitting distributions were specified according to Mathwave's Easyfit software analysis for the ten stream gauge stations, taking note that there no single distribution that describes the relationship between the peak discharge of a flood event and its return period.

Chapter 6 – Conclusion

6.1 Chapter Overview

This chapter outline the final result of this project, as well as recommendations for further improvement and research beyond the scope of this project.

6.2 Project Conclusions

The entire project objectives have been executed in this project. The project objects involved many steps for the successful completion of the aim of this research to investigate the suitability of the Power Law model to estimate flood frequency.

The first objective was the exploration of relevant literature relating to the current conventional flood frequency methods and the Power Law frequency model. This was conducted and reported within Chapter 2 of this report. The literature was found through access to professional databases relevant to the project. This allowed for an appropriate analysis of any past work, techniques and relevant processes used in flood frequency analysis. It was found that the current conventional methods of predicting floods in Australia, according to the Australian Rainfall and Runoff (1987) guide is by probability distributions, more specifically the Log Pearson type 3 distribution for the larger floods and an exponential distribution for the smaller, more common events. Research has shown that the Log Pearson 3 distribution struggles to demonstrate the general trend of the peak discharge and the outlier and tends to significantly over or

under-estimate the largest peak discharge for the water way. Due to this, the Power Law model was investigated in other countries as a more effective method to estimate the largest peak discharge.

Following on from Chapter 2, an appropriate methodology was implemented in order to address the project specifications and complete the project objectives. This methodology included the criteria used in choosing the location of the ten stream gauges across Queensland to maintain a large number of years of flood records and also the process of using these flood records to provide the three frequency models; the Log Pearson type 3 distribution, the exponential distribution and the Power Law model. The analysis method based on the relevant literature was implemented in Chapter 3, which displays the results of each of the frequency models and their goodness of fit to the historical flood data for each of the individual stream gauges selected in Chapter 2. The three frequency model results for each station were displayed in Chapter 4 along with a discussion in terms of the results for each individual site. The general trends and conclusion made regarding each distribution type was discussed in Chapter 5.

The determination from the Australian Rainfall and Runoff (1987) that the exponential distribution of the partial series is accurate enough for the flood events of recurrence less than ten years (small peak discharges) was verified through this analysis, as the ten stations analysed found that in majority of cases, the exponential distribution had the highest goodness of fit values for the smaller flood events. Hence generally the exponential distribution of the partial flood series most accurately predicts the peak discharge for the floods of less than ten year intervals, more so than the Log Pearson or Power Law models.

The general trend of the historical flood data was most effectively modelled with the Log Pearson type 3 distribution, especially for the floods that are expected to return in intervals that are greater than ten years. However, a large number of the stream gauges analysed were found to have the Power Law model predict the largest historical discharge better than the Log Pearson distribution, supporting the recent literature in regards to the suitability of the Power Law relationship for the extreme flood events.

It was found that the main consequence of the Power Law model over the conventional models is that the Power Law model is a far more conservative estimate of the return period of large event. This has particular significance for managing the extreme flood events, as the current largest flood is expected to be exceeded over the course of time and the design of infrastructure should be designed conservatively.

6.3 Further Work

The ten stream gauge locations chosen in this study were found to have no high outlier discharges in the historical flood series data according to the high outlier test explain in Chapter 3.4.1. To accurately test the suitability of the Power Law relationship to the extremely outlier, stream gauges need to be used that contain these high outliers. More stream gauge station data should be used from other locations and tested for high outliers, so that the Power Law can be analysed with reference to them. Care needs to be taken to confirm that these high outliers are not just errors in the data and also that the data is homogeneous. These errors could be changes in the catchment conditions when

the extreme event occurred, unusual phenomenon causing the flood event and problems with the recording equipment.

The estimation of flood frequencies for the Power Law model could be further refined by research investigating the relationship of the various separate factors that influence the Power Law model gradient including the catchment size, slope, shape and climate and also by regionalisation studies as in conventional flood frequency analysis methods.

More recently, it has been suggested that the extreme flood events could be better predicted by the Pareto distribution, which has a 'power like' relationship and hence it would be worthwhile comparing to the Power Law model to determine the more effective method in determining the discharges of the flood with low probability of occurring.

Whilst research of effective flood frequency analysis methods can be complex, tedious and computationally expensive, the consequent results are important over the engineering industry. Accurate and reliable predictions of the peak discharges of large flood events are essential in the design of infrastructure and the safety of human life and it is envisioned that the findings of this project will assist in the continued research into an effective frequency model for predicting the discharges for the design storm.

List of References

Abramowitz, M & Stegun, IA 1965, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, US.

Adeboye, OB & Alatise, MO 2007, 'Performance of Probability Distributions and Plotting Positions in Estimating the Flood of River Osun at Apoje Sub-basin, Nigeria', *Agricultural Engineering International: the CIGR Ejournal*, vol IX, no. 7, pp. 1-21.

Armstrong, WH, Collins, MJ & Snyder, NP 2012, 'Increased Frequency of Low-Magnitude Flood in New England', *Journal of the American Water Resources Association*, vol 48, no. 2, pp. 307-320.

Australian Rainfall and Runoff 1987, 'A Guide to Flood Estimation', in DH Pilgrim (ed.), *Estimation of Design Peak Discharges*, Institution of Engineers, Australia, Barton, ACT.

Australian Rainfall and Runoff 2010, 'Australian Rainfall & Runoff', in *Peak Discharge Estimation*, The Institution of Engineers, Australia, Barton, Victoria.

Baker, VR 1994, 'Geomorphological understanding of floods', *Geomorphology*, vol 10, no. 1-4, pp. 139-156.

Birkeland, KW & Landry, CC 2002, 'Power-laws and snow avalanches', *Geophysical Research Letters*, vol 29, no. 11, pp. 49.1-49.3.

Boughton, WC 1975, 'A Study of Queensland Floods', *The Frequency of Floods in Queensland Symposium*, The Institution of Engineers, Australia, Brisbane.

Bureau of Meteorology 2015, *Australian Water Information Dictionary*, viewed 24 June 2015, <<http://www.bom.gov.au/water/awid/id-644.shtml>>.

Chin, DA 2006, *Water-Resources Engineering*, 2nd edn, Pearson Prentice Hall, Upper Saddle River, New Jersey.

Chow, VT, Maidment, DR & Mays, LW 1988, *Applied Hydrology*, McGraw Hill, New York.

Cooper, RM 2005, 'Estimation of Peak Discharges for Rural, Unregulated Streams in Western Oregon', Scientific Investigations Report 2005-5116, Oregon Water Resources Department, U.S. Geological Survey, 134 p.

Cunnane, C 1985, 'Factors affecting choice of distribution for flood series', *Hydrological Sciences Journal*, vol 30, no. 1, pp. 25-36.

D'Agostino, RB & Stephens, MA 1986, *Goodness-of-Fit Techniques*, Marcel Dekker, Inc., New York.

Ewemoje, TA & Ewemooje, OS 2011, 'Best Distribution and Plotting Positions of Daily Maximum Flood Estimation at Ona River in Ogun-Oshun River Basin, Nigeria', *Agricultural Engineering International: CIGR Journal*, vol 13, no. 3.

Filliben, JJ 2012, *Critical Values of the Chi-Square Distribution*, viewed 2 June 2015, <<http://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm>>.

GraphPad Software n.d., *r², a measure of goodness-of-fit of linear regression*, viewed 2 June 2015, <http://www.graphpad.com/guides/prism/6/curve-fitting/index.htm?r2_ameasureofgoodness_of_fitoflinearregression.htm>.

Grayson, RB, Argent, RM, Nathan, RJ, McMahon, TA & Mein, RG 1996, *Hydrologic Recipes: Estimation Techniques in Australian Hydrology*, CRC For Catchment Hydrology, Clayton, Victoria.

Hall, AJ 1984, 'Hydrology in Tropical Australia and Papua New Guinea', *Hydrological Sciences Journal*, vol 29, no. 4, pp. 399-423.

Hazen, A 1930, *Flood Flows*, John Wiley & Sons, Inc., New York.

Investopedia 2015, *R-Squared*, viewed 15 September 2015, <<http://www.investopedia.com/terms/r/r-squared.asp>>.

Keast, D & Ellison, J 2013, 'Magnitude Frequency Analysis of Small Floods Using the Annual and Partial Series', *Water*, no. 5, pp. 1816-1829.

Khan Academy 2010, *Calculating R-squared*, viewed 5 July 2015, <<https://www.khanacademy.org/math/probability/regression/regression-correlation/v/calculating-r-squared>>.

- Kidson, R & Richards, KS 2005, 'Flood Frequency Analysis: Assumptions and Alternatives', *Progress in Physical Geography*, vol 29, no. 3, pp. 392-410.
- Kidson, R, Richards, KS & Carling, PA 2005, 'Reconstructing the ca. 100-year flood in Northern Thailand', *Geomorphology*, vol 70, no. 3-4, pp. 279-295.
- Kidson, R, Richards, KS & Carling, PA 2006, 'Power-law extreme flood frequency', in G Cello, BD Malamud (eds.), *Fractal Analysis for Natural Hazards*, Geological Society, London.
- Kite, GW 1977, *Frequency and Risk Analysis in Hydrology*, Water Resources Publications.
- Kollmorgen, A, Little, P, Hostetler, S & Griffith, H 2007, 'A baseline assessment of water resources for the National Water Initiative', Australian Water Resources, Australian Government, ISBN 978-1-921107-38-2 , National Water Commission, Canberra.
- Kundzewicz, ZW, Graczyk, D, Maurer, T, Pińskwar, I, Radziejewski, M, Svensson, C & Szwed, M 2005, 'Trend detection in river flow series: 1. Annual maximum flow', *Hydrological Sciences Journal*, vol 50, no. 5, pp. 797-810.
- Linsley, RK, Kohler, MA & Paulhaus, JL 1982, *Hydrology for Engineers*, 3rd edn, McGraw-Hill, New York.
- Malamud, BD & Turcotte, DL 1999, 'Self-Organized Criticality Applied to Natural Hazards', *Natural Hazards*, vol 20, no. 2-3, pp. 93-116.
- Malamud, B & Turcotte, D 2006, 'The Applicability of Power-Law Frequency Statistics to Floods', *Journal of Hydrology*, no. 322, pp. 168-180.
- Malamud, BD, Turcotte, DL & Barton, CC 1996, 'The 1993 Mississippi River Flood: A One Hundred or a One Thousand Year Event?', *Environmental & Engineering Geoscience*, vol 2, no. 4, pp. 479-486.
- Meylan, P, Favre, A-C & Musy, A 2012, *Predictive Hydrology; A Frequency Analysis Approach*, 1st edn, Science Publishers, St. Helier, British Channel Islands.

- Middelmann, M, Granger, K, Harper, B, Baddiley, P & Malone, T 2000, 'Community Risk on Mackay, A multi-hazard risk assessment', Department of Industry, Science and Resources, Australian Geological Survey Organisation, ISBN: 0 642 398623, Canberra.
- National Flood Risk Advisory Group 2008, 'Flood risk management in Australia', *The Australian Journal of Emergency Management*, vol 23, no. 4, pp. 21-27.
- Natural Environment Research Council 1975, 'Flood Studies Report', 1, Hydrological Studies, London.
- Ott, RF & Linsley, RK 1972, 'Streamflow Frequency Using Stochastically Generated Hourly Rainfall', *International Symposium on Uncertainties in Hydrologic and Water Resource Systems*.
- Rahman, AS, Haddad, & Rahman, A 2014, 'Identification of Outliers in Flood Frequency Analysis: Comparison of Original and Multiple Grubbs-Beck Test', *International Journal of Environmental, Ecological, Geological and Geophysical Engineering*, vol 8, no. 12, pp. 765-773.
- Rahman, AS, Rahman, A, Zaman, MA, Haddad, K, Ahsan, A & Imteaz, M 2013, 'A study on selection of probability distributions for at-site flood frequency analysis in Australia', *Natural Hazards*, vol 69, no. 3, pp. 1803-1813.
- Rosbjerg, D, Madsen, H & Rasmussen, PF 1992, 'Prediction in partial duration series with generalized pareto-distributed exceedances', *Water Resources Research*, vol 28, no. 11, pp. 30001-3010.
- Rustomji, P, Bennett, N & Chiew, F 2009, 'Flood variability east of Australia's Great Dividing Range', *Journal of Hydrology*, vol 374, no. 3-4, pp. 196-208.
- Scheidegger, AE 1998, 'Complexity theory of natural diasters; boundaries of self-structured domains', *Earthquake and Atmospheric Hazards*, vol 16, pp. 103-112.
- Schertzer, D, Lovejoy, S & Lavallee, D 1993, 'Generic multifractal phase transitions and self organised criticality', *Colloquium on Cellular Automata*, pp. 217-227.
- Shabri, A 2002, 'A Comparison of Plotting Formulas for the Pearson Type 3 Distribution', *Jurnal Teknologi*, vol 18, no. 36, pp. 61-74.

Singh, VP 1986, *Hydrologic Frequency Modeling*, D. Reidel Publishing Company, Baton Rouge.

Sivandran, G 2002, 'Effects of Rising Water Tables and Climate Change on Annual and Monthly Flood Frequencies', Honours Thesis, Department of Environmental Engineering, University of Western Australia.

Spiess, A-N & Neumeier, 2010, 'Research article An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach', *BMC Pharmacology*, vol 10, no. 6.

Turcotte, DL 1997, *Fractals and Chaos in Geology and Geophysics*, 2nd edn, Cambridge University Press, Cambridge.

Tzanis, A & Makropoulos, K 2002, 'Did the 7/9/1999 M5.9 Athens Earthquake Come with a Warning?', *Natural Hazards*, vol 27, no. 1-2, pp. 85-103.

University Corporation for Atmospheric Research 2010, *Flood Frequency Analysis*, viewed 26 August 2015,

<http://stream2.cma.gov.cn/pub/comet/HydrologyFlooding/FloodFrequencyAnalysisInternationalEdition/comet/hydro/basic_int/flood_frequency/print.htm>.

Uregina, *The Chi Squared Distribution*, viewed 5 July 2015,

<<http://uregina.ca/~gingrich/appchi.pdf>>.

USGS 1982, *Guidelines for Determining Flood Flow Frequency*, U.S. Geological Survey, Reston, VA.

Wisler, CO & Brater, EF 1959, *Hydrology*, 2nd edn, John Wiley & Sons, London.

List of Appendices

Appendix A – Project Specification

Appendix B – Frequency Factors K_T for Log Pearson 3 Distribution

Appendix C – Tables used for identifying high and low outliers

Appendix D – Chi-Square Distribution Table

Appendix A Project Specification

University of Southern Queensland
FACULTY OF HEALTH, ENGINEERING AND SCIENCES

ENG 4111/4112 RESEARCH PROJECT

PROJECT SPECIFICATION

FOR: **Denika MOES**

TOPIC: **POWER-LAW FLOOD FREQUENCY ANALYSIS OF
SELECTED QUEENSLAND STREAMGAUGES**

SUPERVISOR: **Dr Ian Brodie**

PROJECT AIM: This project will provide a power-law statistical model to show
the relationship between discharge data from selected sites with
Queensland's stream gauges network with the average recurrence
interval.

PROGRAMME: (Issue A, 18th March 2015)

1. Undertake a literature review relating to current flood frequency analysis methods, the power-law model and 'goodness-of-fit' tests
2. Obtain peak discharge data both annual and monthly from selected unregulated gauging stations (10 sites) from the Water Monitoring Portal provided by Queensland's Department of Natural Resources and Mines
3. Determine the flood frequency distribution for each site by fitting a Log Pearson III distribution to the annual flood data as conventional method identified by AR&R
4. Analysis for each site the partial series data using the negative exponential distribution as second conventional method identified by AR&R
5. Applied the power-law distribution model to the partial flood data from each site
6. Compare the results of each model for each site (peak discharge at similar ARIs)
7. Apply Goodness of Fit testing to determine more accurate model using EasyFit Distribution Fitting Software from MathWave Technologies

As time permits:

8. Undertake analysis with other distribution formula (Weibull, Gumbel)

Appendix B Frequency Factors K_T for Log Pearson 3 Distribution

Table B.1 Frequency Factors K_T for use with Log Pearson Type 3 Distribution (Positive Skew Coefficients)

Table 1. -- K values for positive skew coefficients

Skew Coefficient (g)	Recurrence Interval in Years											
	1.0101		1.0526		1.1111		1.2500		Percent Chance			
	99	95	90	80	50	20	10	10	4	2		
3.0	-0.667	-0.665	-0.660	-0.636	-0.396	0.420	1.180	1.180	2.278	3.152	4.051	4.970
2.9	-0.690	-0.688	-0.681	-0.651	-0.390	0.440	1.195	1.195	2.277	3.134	4.013	4.909
2.8	-0.714	-0.711	-0.702	-0.666	-0.384	0.460	1.210	1.210	2.275	3.114	3.973	4.847
2.7	-0.740	-0.736	-0.724	-0.681	-0.376	0.479	1.224	1.224	2.272	3.093	3.932	4.783
2.6	-0.769	-0.762	-0.747	-0.696	-0.368	0.499	1.238	1.238	2.267	3.071	3.889	4.718
2.5	-0.799	-0.790	-0.771	-0.711	-0.360	0.518	1.250	1.250	2.262	3.048	3.845	4.652
2.4	-0.832	-0.819	-0.795	-0.725	-0.351	0.537	1.262	1.262	2.256	3.023	3.800	4.584
2.3	-0.867	-0.850	-0.819	-0.739	-0.341	0.555	1.274	1.274	2.248	2.997	3.753	4.515
2.2	-0.905	-0.882	-0.844	-0.752	-0.330	0.574	1.284	1.284	2.240	2.970	3.705	4.444
2.1	-0.946	-0.914	-0.869	-0.765	-0.319	0.592	1.294	1.294	2.230	2.942	3.656	4.372
2.0	-0.990	-0.949	-0.895	-0.777	-0.307	0.609	1.302	1.302	2.219	2.912	3.605	4.298
1.9	-1.037	-0.984	-0.920	-0.788	-0.294	0.627	1.310	1.310	2.207	2.881	3.553	4.223
1.8	-1.087	-1.020	-0.945	-0.799	-0.282	0.643	1.318	1.318	2.193	2.848	3.499	4.147
1.7	-1.140	-1.056	-0.970	-0.808	-0.268	0.660	1.324	1.324	2.179	2.815	3.444	4.069
1.6	-1.197	-1.093	-0.994	-0.817	-0.254	0.675	1.329	1.329	2.163	2.780	3.388	3.990
1.5	-1.256	-1.131	-1.018	-0.825	-0.240	0.690	1.333	1.333	2.146	2.743	3.330	3.910
1.4	-1.318	-1.168	-1.041	-0.832	-0.225	0.705	1.337	1.337	2.128	2.706	3.271	3.828
1.3	-1.383	-1.206	-1.064	-0.838	-0.210	0.719	1.339	1.339	2.108	2.666	3.211	3.745
1.2	-1.449	-1.243	-1.086	-0.844	-0.195	0.732	1.340	1.340	2.087	2.626	3.149	3.661
1.1	-1.518	-1.280	-1.107	-0.848	-0.180	0.745	1.341	1.341	2.066	2.585	3.087	3.575
1.0	-1.588	-1.317	-1.128	-0.852	-0.164	0.758	1.340	1.340	2.043	2.542	3.022	3.489
.9	-1.660	-1.353	-1.147	-0.854	-0.148	0.769	1.339	1.339	2.018	2.498	2.957	3.401
.8	-1.733	-1.388	-1.166	-0.856	-0.132	0.780	1.336	1.336	1.993	2.453	2.891	3.312
.7	-1.806	-1.423	-1.183	-0.857	-0.116	0.790	1.333	1.333	1.967	2.407	2.824	3.223
.6	-1.880	-1.458	-1.200	-0.857	-0.099	0.800	1.328	1.328	1.939	2.359	2.755	3.132
.5	-1.955	-1.491	-1.216	-0.856	-0.083	0.808	1.323	1.323	1.910	2.311	2.686	3.041
.4	-2.029	-1.524	-1.231	-0.855	-0.066	0.816	1.317	1.317	1.880	2.261	2.615	2.949
.3	-2.104	-1.555	-1.245	-0.853	-0.050	0.824	1.309	1.309	1.849	2.211	2.544	2.856
.2	-2.178	-1.586	-1.258	-0.850	-0.033	0.830	1.301	1.301	1.818	2.159	2.472	2.763
.1	-2.252	-1.616	-1.270	-0.846	-0.017	0.836	1.292	1.292	1.785	2.107	2.400	2.670
0	-2.326	-1.645	-1.282	-0.842	0	0.842	1.282	1.282	1.751	2.054	2.326	2.576

Source: Water Resources Council (1967)

Table B.2 Frequency Factors K_T for use with Log Pearson Type 3 Distribution (Negative Skew Coefficients)

Table 2.--K values for negative skew coefficients

Skew Coefficient (g)	Recurrence Interval in Years																					
	1.0101		1.0526		1.1111		1.2500		2		5		10		25		50		100		200	
	99	95	90	80	50	20	10	4	2	1	0.5											
0	-2.326	-1.645	-1.282	-0.842	0	0.842	1.282	1.751	2.054	2.326	2.576											
-.1	-2.400	-1.673	-1.292	-0.836	0.017	0.846	1.270	1.716	2.000	2.252	2.482											
-.2	-2.472	-1.700	-1.301	-0.830	0.033	0.850	1.258	1.680	1.945	2.178	2.388											
-.3	-2.544	-1.726	-1.309	-0.824	0.050	0.853	1.245	1.643	1.890	2.104	2.294											
-.4	-2.615	-1.750	-1.317	-0.816	0.066	0.855	1.231	1.606	1.834	2.029	2.201											
-.5	-2.686	-1.774	-1.323	-0.808	0.083	0.856	1.216	1.567	1.777	1.955	2.108											
-.6	-2.755	-1.797	-1.328	-0.800	0.099	0.857	1.200	1.528	1.720	1.880	2.016											
-.7	-2.824	-1.819	-1.333	-0.790	0.116	0.857	1.183	1.488	1.663	1.806	1.926											
-.8	-2.891	-1.839	-1.336	-0.780	0.132	0.856	1.166	1.448	1.606	1.733	1.837											
-.9	-2.957	-1.858	-1.339	-0.769	0.148	0.854	1.147	1.407	1.549	1.660	1.749											
-1.0	-3.022	-1.877	-1.340	-0.758	0.164	0.852	1.128	1.366	1.492	1.588	1.664											
-1.1	-3.087	-1.894	-1.341	-0.745	0.180	0.848	1.107	1.324	1.435	1.518	1.581											
-1.2	-3.149	-1.910	-1.340	-0.732	0.195	0.844	1.086	1.282	1.379	1.449	1.501											
-1.3	-3.211	-1.925	-1.339	-0.719	0.210	0.838	1.064	1.240	1.324	1.383	1.424											
-1.4	-3.271	-1.938	-1.337	-0.705	0.225	0.832	1.041	1.198	1.270	1.318	1.351											
-1.5	-3.330	-1.951	-1.333	-0.690	0.240	0.825	1.018	1.157	1.217	1.256	1.282											
-1.6	-3.388	-1.962	-1.329	-0.675	0.254	0.817	0.994	1.116	1.166	1.197	1.216											
-1.7	-3.444	-1.972	-1.324	-0.660	0.268	0.808	0.970	1.075	1.116	1.140	1.155											
-1.8	-3.499	-1.981	-1.318	-0.643	0.282	0.799	0.945	1.035	1.069	1.087	1.097											
-1.9	-3.553	-1.989	-1.310	-0.627	0.294	0.788	0.920	0.996	1.023	1.037	1.044											
-2.0	-3.605	-1.996	-1.302	-0.609	0.307	0.777	0.895	0.959	0.980	0.990	0.995											
-2.1	-3.656	-2.001	-1.294	-0.592	0.319	0.765	0.869	0.923	0.939	0.946	0.949											
-2.2	-3.705	-2.006	-1.284	-0.574	0.330	0.752	0.844	0.888	0.900	0.905	0.907											
-2.3	-3.753	-2.009	-1.274	-0.555	0.341	0.739	0.819	0.855	0.864	0.867	0.869											
-2.4	-3.800	-2.011	-1.262	-0.537	0.351	0.725	0.795	0.823	0.830	0.832	0.833											
-2.5	-3.845	-2.012	-1.250	-0.518	0.360	0.711	0.771	0.793	0.798	0.799	0.800											
-2.6	-3.889	-2.013	-1.238	-0.499	0.368	0.696	0.747	0.764	0.768	0.769	0.769											
-2.7	-3.932	-2.012	-1.224	-0.479	0.376	0.681	0.724	0.738	0.740	0.740	0.741											
-2.8	-3.973	-2.010	-1.210	-0.460	0.384	0.666	0.702	0.712	0.714	0.714	0.714											
-2.9	-4.013	-2.007	-1.195	-0.440	0.390	0.651	0.681	0.683	0.689	0.690	0.690											
-3.0	-4.051	-2.003	-1.180	-0.420	0.396	0.636	0.660	0.666	0.666	0.667	0.667											

Source: Water Resources Council (1967)

Appendix C Tables used for identifying high and low outliers

Table C.1 Values of K_N for outlier tests, 5% significance level values

Sample Size N	K_N	Sample Size N	K_N	Sample Size N	K_N
		51	2.964	101	3.210
		52	2.971	102	3.214
3	1.153	53	2.978	103	3.217
4	1.463	54	2.986	104	3.220
5	1.672	55	2.992	105	3.224
6	1.822	56	3.000	106	3.227
7	1.938	57	3.006	107	3.230
8	2.032	58	3.013	108	3.233
9	2.110	59	3.019	109	3.236
10	2.176	60	3.025	110	3.239
11	2.234	61	3.032	111	3.242
12	2.285	62	3.037	112	3.245
13	2.331	63	3.044	113	3.248
14	2.371	64	3.049	114	3.251
15	2.409	65	3.055	115	3.254
16	2.443	66	3.061	116	3.257
17	2.475	67	3.066	117	3.259
18	2.504	68	3.071	118	3.262
19	2.532	69	3.076	119	3.265
20	2.557	70	3.082	120	3.267
21	2.580	71	3.087	121	3.270
22	2.603	72	3.092	122	3.274
23	2.624	73	3.098	123	3.276
24	2.644	74	3.102	124	3.279
25	2.663	75	3.107	125	3.281
26	2.681	76	3.111	126	3.284
27	2.698	77	3.117	127	3.286
28	2.714	78	3.121	128	3.289
29	2.730	79	3.125	129	3.291
30	2.745	80	3.130	130	3.294
31	2.759	81	3.134	131	3.296
32	2.773	82	3.139	132	3.298
33	2.786	83	3.143	133	3.302
34	2.799	84	3.147	134	3.304
35	2.811	85	3.151	135	3.306
36	2.823	86	3.155	136	3.309
37	2.835	87	3.160	137	3.311
38	2.846	88	3.163	138	3.313
39	2.857	89	3.167	139	3.315
40	2.866	90	3.171	140	3.318
41	2.877	91	3.174	141	3.320
42	2.887	92	3.179	142	3.322
43	2.896	93	3.182	143	3.324
44	2.905	94	3.186	144	3.326
45	2.914	95	3.189	145	3.328
46	2.923	96	3.193	146	3.331
47	2.931	97	3.196	147	3.334
48	2.940	98	3.201		
49	2.948	99	3.204		
50	2.956	100	3.207		

Source: Australian Rainfall and Runoff (1987)

Table C.2 Values of β given from values of skew (g) and years of flood data (n)

<i>N</i>	-3.9	-3.0	-2.0	-1.0	-0.5	0.0	0.5	1.0	2.0	3.0	3.9
10	.86	.96	.96	.87	.95	1.10	1.26	1.42	1.67	1.80	1.94
15	.67	.77	.85	.82	.90	1.05	1.22	1.40	1.71	1.90	1.99
20	.58	.67	.77	.79	.87	1.02	1.20	1.39	1.73	1.98	2.11
30	.47	.55	.68	.76	.84	.99	1.17	1.37	1.77	2.09	2.28
40	.42	.49	.62	.73	.82	.97	1.16	1.37	1.79	2.15	2.40
50	.38	.45	.58	.71	.81	.96	1.15	1.36	1.80	2.19	2.47
60	.35	.42	.55	.70	.80	.95	1.15	1.36	1.81	2.21	2.52
70	.33	.40	.53	.68	.79	.95	1.15	1.36	1.81	2.21	2.52
80	.32	.38	.51	.67	.78	.94	1.14	1.36	1.81	2.26	2.60
90	.30	.37	.50	.66	.78	.94	1.14	1.36	1.82	2.27	2.63
100	.29	.35	.48	.66	.77	.94	1.14	1.36	1.83	2.29	2.64

Source: Australian Rainfall and Runoff (1987)

Table C.3 Values of θ given from values of skew (g) and years of flood data (n)

<i>N</i>	Skew <i>g</i>										
	-3.9	-3.0	-2.0	-1.0	-0.5	0.0	0.5	1.0	2.0	3.0	3.9
10	1.94	1.80	1.67	1.42	1.26	1.10	.95	.87	.96	.96	.86
15	1.99	1.90	1.71	1.40	1.22	1.05	.90	.82	.85	.77	.67
20	2.11	1.98	1.73	1.39	1.20	1.02	.87	.79	.77	.67	.58
30	2.28	2.09	1.77	1.37	1.17	.99	.84	.76	.68	.55	.47
40	2.40	2.15	1.79	1.37	1.16	.97	.82	.73	.62	.49	.42
50	2.47	2.19	1.80	1.36	1.15	.96	.81	.71	.58	.45	.38
60	2.52	2.21	1.81	1.36	1.15	.95	.80	.70	.55	.42	.35
70	2.57	2.24	1.81	1.36	1.14	.95	.79	.68	.53	.40	.33
80	2.60	2.26	1.81	1.36	1.14	.94	.78	.67	.51	.38	.32
90	2.63	2.27	1.82	1.36	1.14	.94	.78	.66	.50	.37	.30
100	2.64	2.29	1.83	1.36	1.14	.94	.77	.66	.48	.35	.29

Source: Australian Rainfall and Runoff (1987)

Appendix D Chi-Square Distribution Table

Table D.1 Critical values of chi-square distribution with degrees of freedom (df)

<i>df</i>	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Source: Filliben (2012)