A High Resolution Point Rainfall Model Calibrated to Short Pluviograph or Daily Rainfall Data

Shane Anthony Jennings

December 2006

Ph.D. Thesis

Department of Civil and Environmental Engineering
Adelaide University
Australia

ABSTRACT

The design of hydraulic systems that have to cope with natural flows of flood magnitude is risk-based The estimation of flood risk relies on joint probability theory where the combination of stochastic inputs such as rainfall and a description of the hydrological/hydraulic runoff process determine the probability distribution of flooding events. To date both the design storm approach presented in Australian Rainfall and Runoff (Institution of Engineers Australia, 1987) and continuous simulation through a Monte Carlo approach have provided workable methods for deriving empirical flood probability distributions as an estimate of this flood risk. While the continuous simulation approach has long been viewed as the best way to evaluate the probabilistic behavior of surface water systems, the design storm approach has remained the preferred choice due to its simplicity and ease of use. However with the onset of powerful personal computers providing the ability for increasingly complex analysis within the required timeframes, the tendency towards using a continuous simulation approach will continue to grow.

The idea behind the Monte Carlo continuous simulation approach is that a long model simulation will eventually sample all possible joint probability interactions (i.e. all combinations of rainfall input and runoff model conditions etc) within a system. If this is the case, the derived flood distribution from these simulations can be viewed as an accurate inference of the true flood distribution and therefore can be used for engineering analysis and evaluation of flood risk. A drawback of the Monte Carlo approach is the required input of a long rainfall record. In the absence of a significant historical record, rainfall models can be used to provide the required data but are in turn reliant on adequate historical data for calibration. Accurate calibration of rainfall models is particularly important in Australia where the variability of rainfall at short and long term time scales is large.

Australia does have an extensive network of rainfall recording stations. These sites record rainfall data in various forms ranging from a daily time step down to six-minute resolution. While the size of historical daily records is often large, there are very few six-minute (Pluviograph) records available of significant length. Indeed,

analysis of Australia's pluviograph records indicates that the average length of the more than 900 pluviograph data sets available from the Bureau of Meteorology is approximately 15 years. Only a small number of sites have a record length exceeding 40 years and of these only 40 or so remain active. Even with the high quality of rainfall data in Australia, periods of missing or corrupt data are often present. Not only does this lack of significant short time scale data provide a major obstacle in the application of a Monte Carlo approach to risk estimation, it also inhibits the application of rainfall simulation models that use this data for direct calibration. This lack of data is particularly important if we consider the tails of the flood probability distribution where it is unlikely that a 15-year historical record can provide accurate estimates of a 100 year flood event. While the advent of numerous stochastic rainfall models provide methods for extending historical rainfall records, without adequate historical rainfall data available for calibration their accuracy is questionable.

This thesis describes the development of a new technique which significantly extends the applicability of stochastic point rainfall models that require historical data for calibration. The technique is demonstrated using a high-resolution point rainfall model based on wet-dry alternating storm events. The original model presented by Heneker *et al.* (2001) uses storm events which are defined by the observed event distributions of dry periods, storm event durations and storm intensity conditioned on storm duration and replicates this event structure during simulation.

Significant improvements to the original model are presented as the first part of this thesis. The parameterisation used to describe the event distributions has been simplified and the number of parameters reduced resulting in a model that is more robust and easier to calibrate. In addition, the Metropolis algorithm (Metropolis *et al.* (1953)) was incorporated into the model providing a description of the posterior distribution of model parameters and as a result enables a description of parameter uncertainty within the model structure. These improvements have produced a model that is well defined and can be vigorously compared against numerous observed statistics in a quantitative manner. Simulation results indicate that the model is able to replicate both calibrated and non-calibrated statistics at various time scales.

The original model required the use of a long pluviograph record at the site of interest to ensure an accurate calibration of model parameters. To circumvent this restriction in the application of the model a new 'master'- 'target' scaling relationship has been developed and incorporated into the model. A model calibration is undertaken at a 'master' site with a long pluviograph record which is then updated and scaled to the 'target' site of interest using the information from either a short pluviograph or daily rainfall record. This structure has removed the need for significant pluviograph data at the 'target' site and enables the rainfall model to be applied at sites with short pluviograph or daily rainfall records.

The approach has been tested at numerous pairs of sites providing evidence of its success in generating accurate synthetic pluviograph data across the country and within various climatic regions. Model results are presented and compared for both the observed pluviograph data (for individual storm and sub-daily statistics) and daily data (for longer aggregated statistics) available at the target sites and compares well to Australian data. The rainfall model presented in this thesis can be used to provide accurate synthetic rainfall data at sites with minimal historical rainfall data providing a powerful tool for application in hydrological risk analysis across Australia.

STATEMENT OF ORIGINALITY

This work contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed Date: 22 JUNE 2007

Shane Anthony Jennings

ACKNOWLEDGEMENTS

This research was funded by an Australian Postgraduate Award (APA) and an Australian Research Council (ARC) grant.

I would like to thank the following people and organisations whose assistance was invaluable in completing the work presented in this thesis:

- Dr. Martin Lambert from the Department of Civil and Environmental Engineering at Adelaide University for his supervision of this project and for providing inspiration, ideas and guidance.
- Associate Professor George Kuczera from the Department of Civil, Surveying and Environmental Engineering at the University of Newcastle for many useful and essential ideas and discussions.
- Dr. Andrew Metcalfe from the Department of Mathematics at Adelaide University for his thoughts and ideas.

Andrew Frost (now Dr Andrew Frost) for his friendship, the numerous invaluable discussions on both his and my research and for his parent's house in Fingal Bay which provided a fantastic study retreat during this research.

My mum and dad for their ongoing support in everything I set out to achieve.

Finally, I thank and dedicate this thesis to my wife Grace for her love, understanding and support throughout this research. Without her encouragement and assistance, particularly in the final stages of completion, the end may not have been reached.

LIST OF PUBLICATIONS

Several papers have been published during this study. They reflect formative ideas for work or the application of the work presented in this thesis.

- Jennings, S., Lambert, M., Frost, A. & Kuczera, G. "Regionalisation of a High Resolution Point Rainfall Model". Proceedings of the 27th Hydrology and Water Resources Symposium, Melbourne May 2002.
- Frost, A., Jennings, S., Thyer, M., Lambert, M. & Kuczera, G. "Droughts, Floods and Everything Else in Between". Proceedings of the 3rd International Hydrology and Water Resources Symposium of the Institution of Engineers Australia, Perth Nov 2000.
- Kuczera, G., Lambert, M., Jennings, S., Frost, A., Heneker, T. & Coombes, P. "Are Design Rainfalls a Thing of the Past". NZ Water Wastewater Association Conference, NZ Nov 2002.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The design and analysis of complex hydraulic engineering systems is typically risk-based. Underground urban pipe networks, bridges, culverts, channels and wetlands are all designed to cope with natural stormwater flows of a certain flood magnitude. The estimation of how often these systems will fail (or the probability of observing an event that exceeds an assumed design level – flood risk) is fundamental to the risk analysis process.

The estimation of flood risks relies on joint probability theory where the combination of inputs such as rainfall and a description of the hydrological/hydraulic runoff process determine the probability distribution of flooding events. Within Australia engineers often use the approach presented in Australian Rainfall and Runoff (referred to as ARR) (Institution of Engineers Australia, 1987) which is known as the design storm approach. The method for evaluating flood risk probabilities is based on a design rainfall storm for which "the intention is to derive a flood of selected probability of exceedance from a design rainfall of the same probability" [ARR, p6]. This approach relies on the assumption that median values of all other variables other than rainfall (such as losses, base flow, temporal patterns and hydrograph model parameters) can be used and still estimate an accurate runoff representation providing a flood of the same exceedance probability as the input design storm. Unfortunately, ARR does not demonstrate that this objective is achieved and indeed admits that "there is a need for research to test this approach".

The problem of estimating flood risk can also be solved empirically using a Monte Carlo continuous simulation model which requires simulating the flood response due to a long rainfall record and empirically deriving the resultant flood probability distributions. This technique works on the basis that a sufficiently long simulation will eventually sample almost all possible joint probability interactions (i.e. all combinations of rainfall input and possible runoff conditions etc). If this can be achieved successfully, the derived flood probability distribution can be viewed as an accurate inference of the true flood probability distribution.

Despite the theoretical superiority of the continuous simulation approach, designers across Australia continue to adopt the design rainfall method as the method of choice not only due to its simplicity, but also due to the problems associated with using continuous simulation models in the past. These models rely on a large number of Monte Carlo simulations, which in turn requires significant computational effort and storage space. Previously this could be seen as prohibitive, however with the continual increase in the power and availability of personal computers, this issue has become less relevant.

A more significant issue is the availability and length of historical rainfall records available for use in Monte Carlo applications. This is particularly important if we consider the tails of the flood probability distribution where it is unlikely that a 15-year historical record can provide accurate estimates of a 100 year flood event. While the advent of numerous rainfall models can provide a method for extending historical rainfall records, without adequate historical rainfall data available for calibration their accuracy is questionable. In particular, models which attempt to produce synthetic rainfall at the sub-hourly time frame are often susceptible to insufficient calibration data for the complex processes that these models are attempting to reproduce. This is particularly relevant for analysis of systems where initial catchment conditions or storage volume is important (i.e. flood analysis or when investigating Water Sensitive Urban Design components in new or existing stormwater systems). Without the local availability of significant high resolution data for calibration or a technique to use alternate additional data sources (i.e. daily), these models will continue to remain restricted in their application and usefulness as an engineering tool.

The Australian Government's Bureau of Meteorology is responsible for the majority of climate stations across the country recording various climatic variables including wind, solar radiation, cloud cover and rainfall. If we consider rainfall recording stations to understand the calibration data available for high resolution rainfall models, then these stations are defined as belonging to one of two types, either the observed rainfall depth is recorded over a given day (Daily gauge) or continuously on a chart (Dines pluviograph) while the tipping bucket rain gauges record the time of tips. The distribution and length of these rainfall stations across the country provides a snapshot into the potential availability of calibration rainfall data within Australia.

At the time of writing the Bureau of Meteorology administered approximately 942 pluviograph rainfall sites in total across the country (see Figure 1.1).

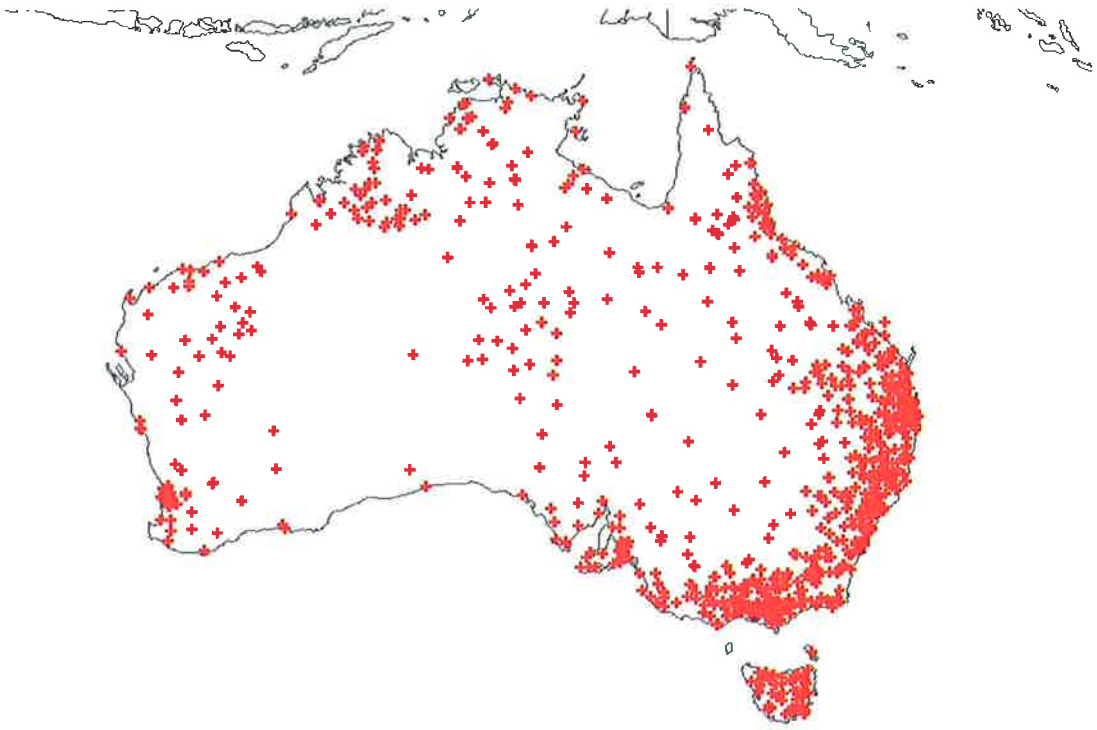

Figure 1.1: Australian Bureau of Meteorology: Pluviograph Recording Stations

A first glance gives the impression that this number of gauges could be considered an excellent basis for rainfall model calibration and provides a useful data tool for engineers in general. However, further analysis into the extent of this data reveals

significant inadequacies. Of the 942 pluviograph sites in Australia, the combined average length is only approximately 15 years. Even if this value is slightly biased by a number of sites that are relatively new (or were recorded for a specific purpose and contain only a few years of record), more alarming is the fact that of all sites that are still active, only 36 have a record length greater than 40 years.

Assuming 40 years of historical record is adequate for model calibration, Figure 1.2 displays the sparse nature of these pluviograph calibration sites available in Australia. Complicating issues further is that these historical records often contain sections of missing or erroneous data (faulty gauges, time aggregated rainfall totals), which present another obstacle (and a potential reduction in data length) in using this data to successfully calibrate high resolution rainfall models.

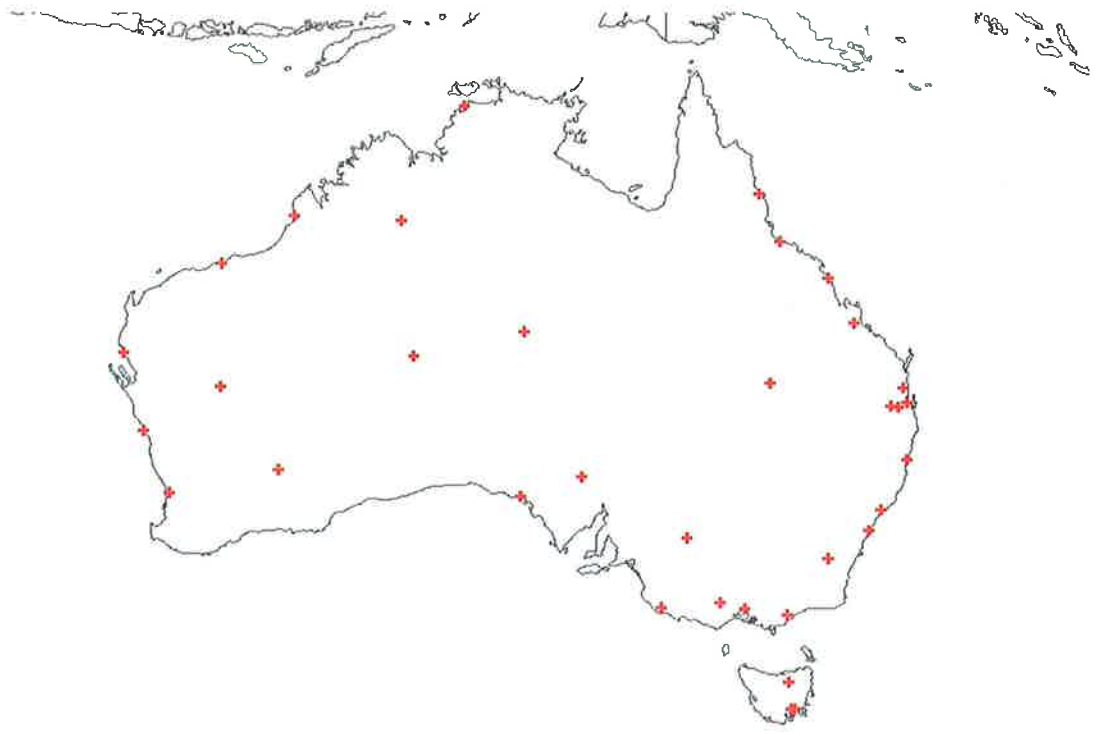

Figure 1.2: Australian Bureau of Meteorology: Pluviograph Recording Stations with a Historical Record Greater than 40 Years

Further analysis of Figure 1.2 shows that while sparse in number, the long term pluviograph records across Australia are located at the major Australian centres and are distributed throughout the major climatic regions. These climatic regions are

shown in Figure 1.3 (i.e. Temperate climate of Adelaide/Melbourne/Sydney, Sub-Tropics of Brisbane, the Desert of Alice Springs etc). The distribution of long term sites across major climatic regions enables these sites to be used as a basis for potential regionalisation work.

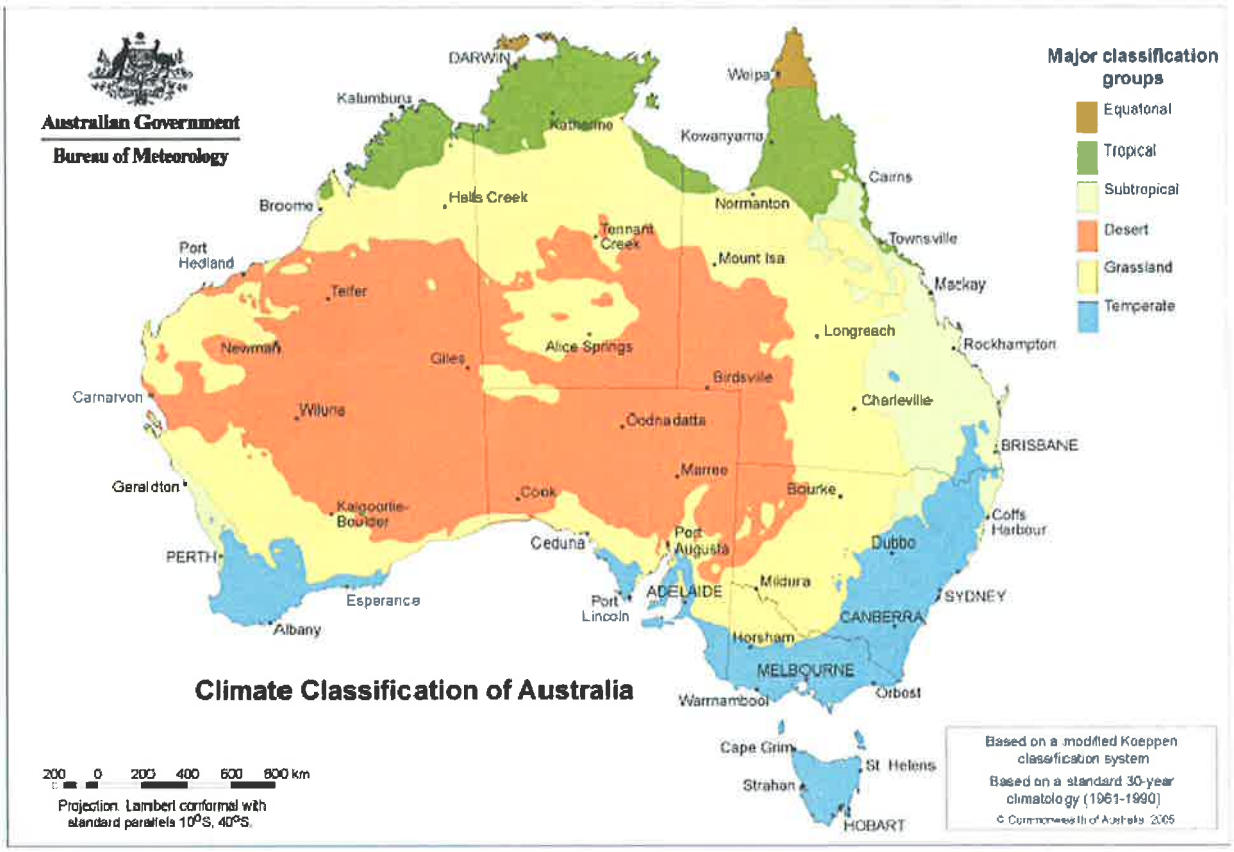

Figure 1.3: Climate Classification of Australia (Australian Bureau of Meteorology website)

In addition to the network of pluviograph stations, the Bureau of Meteorology administers approximately 18,000 daily recording stations (See Figure 1.4). These rainfall stations provide daily records not only across mainland Australia but also on islands off shore and even across Antarctica (not shown).

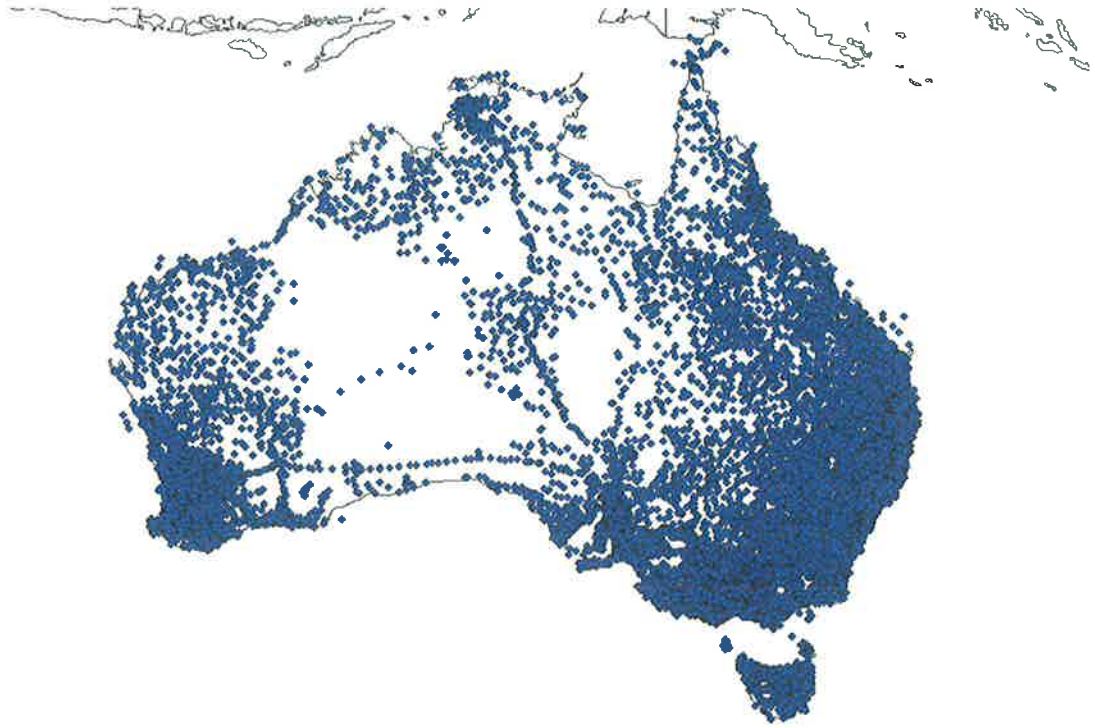

Figure 1.4: Australian Bureau of Meteorology: Daily Recording Stations

If the assumption of a 40 year calibration record is continued, then there are in excess of 7,300 daily sites which contain a sufficient data record and just over 4,400 of these are still active. For models that are able to utilise daily data for calibration, the Australian data set provides a comprehensive number of sites to choose from with a distribution across the nation that almost guarantees an adequate calibration site can be found within close proximity. In locations where no data is available, techniques also exist which provide interpolated daily data records at the site of interest based on neighbouring data sites. These in turn can be used for model calibration.

The comparison between potential pluviograph and daily calibration sites presents a stark contrast. Not only are there more active daily stations with long historical records (4,400) than the total number of pluviograph stations (932) within Australia, pluviograph stations in general have been focused on major centres along the coast. Figure 1.5 displays the extensive coverage of daily sites in direct comparison to their pluviograph counterparts when considering only those sites with more than 40 years of record.

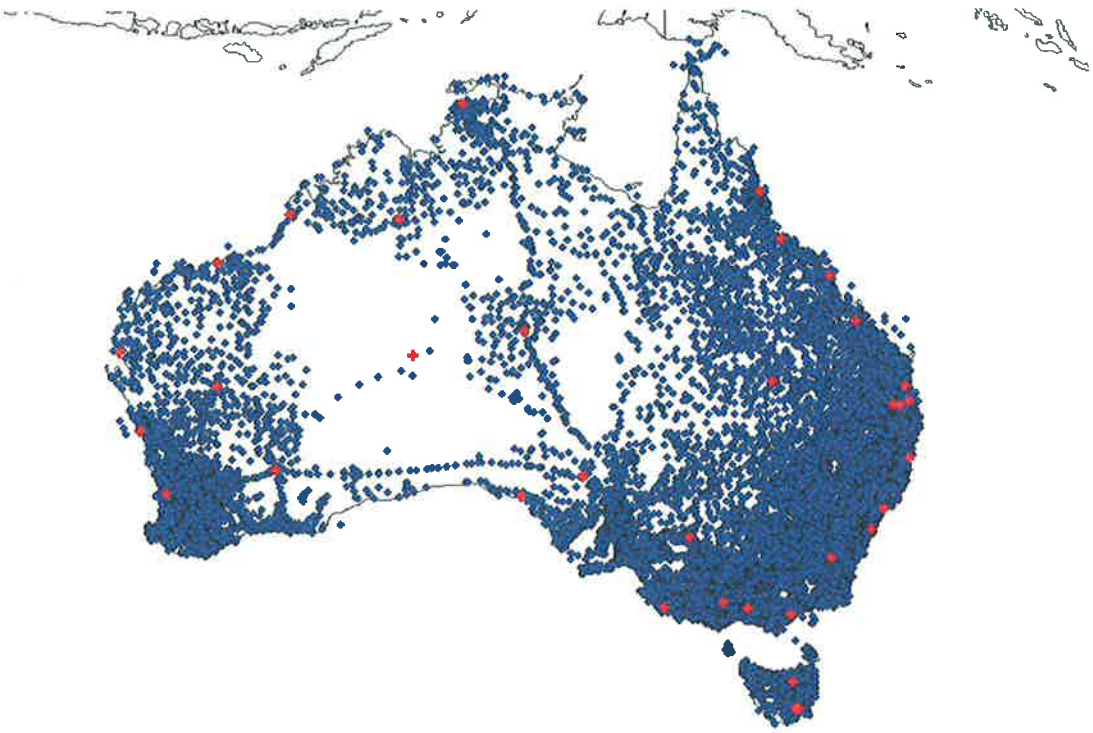

Figure 1.5: Australian Bureau of Meteorology: Pluviograph and Daily Recording Stations with a Historical Record Greater than 40 Years

It is clear that any model relying completely on significant historical pluviograph data for calibration is severely limited in its application in Australia in contrast to models capable of using daily data records. If a continuous rainfall simulation model was developed with adequate complexity to capture the sub-daily rainfall characteristics but also structured in a manner to utilise the limited information available from a daily calibration site, it is also clear that such a model would provide a valuable hydrological tool capable of wide spread application across Australia.

1.2 Aims

Continuous simulation models can provide significant advantages over the design storm approach to engineering analysis and design, however these models require a description of the stochastic rainfall input. In the absence of a significant historical record, rainfall models can supplement the historical information, however high resolution rainfall models in particular are limited in their application due to the sparse nature of calibration records and periods of time within these records which

are missing or have been influenced by malfunctions or errors with the recording apparatus.

The ultimate goal of this research was to provide a rainfall simulation model which could successfully simulate accurate synthetic pluviograph records at sites across Australia with minimal or no historical pluviograph data. To achieve this objective five aims were developed:

(1) To develop or select a rainfall model capable of simulating synthetic pluviograph data;

(2) To refine and improve the rainfall model by including uncertainty and a Monte Carlo simulation structure ensuring the calibration process is robust and comparison to observed data is accurate.

(3) To verify the accuracy of the rainfall model by analysing its performance and structure at sites with significant pluviograph records for calibration and comparison;

(4) To extend the application of the model to sites with minimal historical pluviograph data available for calibration and finally;

(5) To extend the application of the model to sites with only historical daily data and no pluviograph data available for calibration.

1.3 Research Outline

The outline of this research can be presented with a description of each chapter. Chapter 2 contains a review of current methods or approaches available for rainfall simulation and describes the rainfall model selected for re-development. After a review of point rainfall models (including Poisson, Cluster, Markov and Alternating Renewal Models), an alternating renewal model introduced by Heneker *et al.* (2001) was selected for further development and use. The selected point rainfall model

introduced by Heneker *et al.* (2001) calibrates to storm event data extracted from historical pluviograph records and as such is not vulnerable to missing or corrupt data periods in the historical record. The model was also able to reproduce both calibrated and non-calibrated statistics when applied to sites across various climatic regimes in Australia making it an ideal choice for further investigation. Chapter 3 considers the Heneker *et al.* (2001) model further and presents enhancements to the model structure and calibration process which improve the robustness and accuracy of the model calibration. This is particularly important if the model was to be calibrated successfully at sites with little or no pluviograph data. Chapter 4 considers parameter uncertainty and how including a description of uncertainty within the model can be used to improve statistical comparisons to observed data sets. Chapter 5 presents a validation of the improved model against numerous Australian data sites across a range of climatic conditions. This provides evidence of the models performance with adequate calibration data and its ability to replicate required observed statistics.

Chapter 6 develops a new calibration process which enables the rainfall model parameters to be calibrated at site locations with only a short historical pluviograph data record. In its original form, the adopted rainfall model (as is the case with most rainfall models that describe the rainfall process at a sub-hourly scale) required a significant length of historic data for accurate calibration. To provide a method for calibrating to a short historical record, a master-target relationship is introduced. This relationship is developed to exploit the similarities in the model parameters between two sites which reside in a similar region. This master – target framework uses the selection of a master site containing a long pluviograph record as a basis with which the model can be accurately calibrated initially. This initial calibration is then updated by the limited pluviograph information available at the target site of interest. This master-target calibration approach ensures the model is able to describe the rainfall process at the target site with minimal calibration data.

In addition to developing a process for shifting each model parameter, a technique was required to deal with the major issue of the different lengths and time periods over which the master and target data sets were observed. In order for the process to be successful, concurrent data periods were required for comparison at the master

and target site to ensure any calibration of the model when shifting from the master to the target site was a reflection of the requirements of the model parameters and not a reflection of the differences in data time periods. To circumvent this issue a pivotal intermediate calibration step is also introduced which enables the model to capture variations that exist between the two sites as a result of their different record lengths and non-concurrent data periods, which are then taken into account prior to a final comparison between the sites to shift the model parameters from the master to target sites. The overall approach is then tested by selecting target sites with sufficient data for comparison (but only using a sub-set of this data for calibration), providing evidence of its success in generating accurate synthetic pluviograph data at sites across the country and across various climatic conditions. Model results are presented and compared for both the observed pluviograph data (for individual storm and sub-daily statistics) and daily data (for longer aggregated statistics) available at the target sites.

Chapter 7 develops a process for calibrating the pluviograph rainfall model at sites with no historical pluviograph data. A master-target relationship similar in structure to Chapter 6 is developed with the master site consisting of a long pluviograph record while the target site contains only daily data. Initial calibration is completed at the master site, forming the basis for further calibration to the daily record at the target site. This approach ensures the model can successfully capture the underlying structure of the sub-daily rainfall (through the calibration at the master site) while having the ability to capture the required rainfall differences (through the use of daily data) that occur at the target site. A simulated likelihood fitting approach is developed to facilitate this process and enable direct comparison between aggregated master simulation data and the observed daily data at the target site. Model results are presented for the same pluviograph sites as used in Chapter 6 which enables a comparison of the accuracy of the two approaches (with and without pluviograph data) and ensures sub-daily statistics are successfully reproduced when calibrating with daily data at the target site.

Finally conclusions and recommendations of the research presented in this thesis are made in Chapter 8.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The modelling of hydrological engineering systems often requires the input of a long-term historical rainfall record. For example, volume based design of a stormwater detention or wetland system involves the interaction between existing system levels and incoming flow. Similarly the runoff response from the catchment itself which determines the incoming flow is dependent on the initial conditions of the catchment and whether it has been a relatively wet or dry period. These problems can only be accurately modelled through the use of a continuous water balance simulation. In the absence of a significant historical record or to provide further engineering system evaluation through continuous simulation, rainfall models can be used to provide synthetic records as inputs into such system simulations. A major issue which confronts the users of rainfall models, particularly those which attempt to reproduce rainfall at the sub daily time step, is the lack of quantity and quality within historical rainfall records available for model calibration. With this in mind this chapter reviews the developments of stochastic point rainfall models at various time scales and current regionalisation techniques. This review has identified a particular point rainfall model that is suitable for further development and the key features of a new regionalised framework that will enable the model to be applied at numerous sites across the country.

2.2 Physical Process of Precipitation

Precipitation is the primary source of the Earth's water supplies and is the blanket term used to describe rain, snow, hail and all other forms of moisture that fall from clouds in the atmosphere. The physical generation of precipitation can only proceed when four processes occur in sequence. Initially there must be a cooling of moist air to the dew point temperature. This is usually a product of warm moist air cooling as it rises through the atmosphere. Once this occurs, condensation follows forming cloud droplets or ice crystals. This is a complex process and relies on the presence of dust and aerosols in the atmosphere to provide a surface for condensation to occur. The efficiency of the condensation process is influenced by the size and number of these microscopic particles that are available (Burroughs, 1999). If these particles are not present in the atmosphere, the condensation process and therefore precipitation cannot begin. As time progresses, crystals formed during the condensation process continue to develop and grow forming raindrops. Finally a constant supply of water vapour provides the fuel required to ensure these processes continue to produce precipitation in one of its many forms (Gilman, 1964).

The three dimensional interaction between these processes and the surrounding atmosphere due to local and global circulation patterns ensures the precipitation process can be considered pseudo-chaotic. Due to the complexity involved and our lack of complete understanding, physical based models which attempt to mathematically describe these underlying physical processes can only have limited application both in the temporal and spatial scale (Cho 1985, Cho and Chan 1987). The difficulty in producing such deterministic physical models that adequately describe this complex and evolving process of rainfall have lead researchers to concentrate on modelling rainfall stochastically to reproduce certain statistical attributes of the observed data.

2.3 Stochastic Rainfall Modelling

The stochastic modelling of rainfall is an invaluable tool for describing and providing rainfall information to designers and engineers. Unlike deterministic models, which attempt to model and reproduce the exact physical processes involved during the precipitation process, stochastic rainfall models are a generalised mathematical model of the system used to reproduce observable quantities and statistics. In the stochastic model, the actual physical process of rainfall plays no part. Stochastic models with a good theoretical structure will have model parameters that are based on a physical characteristic (such as storm duration); however this is not a pre-requisite in models of this type. Yevjevich (1974) provided an interesting comparison between determinism and stochasticity in hydrology and noted that the best form of hydrological model is likely to be a combination of deterministic measurable parameters with a stochastic model structure. This would seem to provide a model that had reliable and understandable inputs but also enabled the model to capture the variability and randomness often observed in natural processes and systems. Raudkivi and Lawgun (1974) also discussed the use of deterministic and stochastic models and concluded that it is unlikely deterministic models could be applied in design situations due to the complexity and number of variables involved. Continuing in this vein, Cho (1985) argued that it is more practical and justifiable to assume a stochastic process for rainfall rather than try to develop a deterministic model.

Typically stochastic rainfall models can be classified into two types. Stochastic rainfall models that try to capture the characteristics of rainfall through time at a single point are commonly referred to as point rainfall models. Models that also include the development and decay of rainfall over space as well as time and are interested in how the rainfall fluxes across a catchment are referred to as spatial rainfall models. As the purpose of this study was to provide accurate representation of rainfall at a single point given limited calibration data, spatial rainfall models were not considered.

2.4 Point Rainfall Models

The history of stochastic point rainfall modelling extends back to the 1930's. A paper written by Shamir (1965) refers to earlier work by Slade (1936) where a probability distribution was fitted to annual rainfall data. Since this beginning, numerous techniques and methods of rainfall modelling have been developed. While the number of rainfall models is large, generally they can be classified as one of three main types. The first and possibly the best known rainfall models are those which are based on the Point Process. The general theory behind point processes is outlined by Cox and Isham (1980) while the use of these processes and the mathematical structure of various rainfall modelling approaches can be found in Waymire and Gupta (1981a, 1981b, 1981c). Poisson models belong to the group of Point Process models, as do cluster models which are an extension to the Poisson assumption. The second group of models are referred to as Markov models. Markov models employ a discrete time step and inherit their name through the use of a Markov conditional structure between subsequent time steps. As they are based in discrete time, generally they have not been applied to a sub-daily scale. The final group of rainfall models are based on the assumption of independent storm events which are typically defined as 'wet' or 'dry'. As the model generates a synthetic rainfall record through the alternate simulation of these storm events, the models are often referred to as 'Wet-Dry Spell' or Alternating Renewal models.

The development of rainfall models has been dominated by Poisson, Markov or Alternating Renewal models, allowing the selection of an appropriate model for this study to focus in these areas. The development and properties of each type of model is presented in further detail below. In selecting a model to extend through a new regionalisation technique, attention was paid to the ability of the model to reproduce statistics not used during its calibration process, which provide a valuable check of the validity of the model. Attention was also paid to statistics which are critical for engineering design purposes such as Intensity-Frequency-Duration curves. Finally if a model is to be used as an engineering tool, the model parameters should be identifiable and easy to calibrate at any site. The parameter structure was also taken into account when selecting the appropriate rainfall model.

2.4.1 Poisson Models

A number of well known rainfall models in existence today are based on the Poisson model structure which utilises the basic point process approach. The point process approach describes the occurrence (position) of independent events in the modelling space. For rainfall models the occurrence process defines the position of rainfall occurrences in time or the temporal position. Generally these models can assume a continuous time process where events can occur at any point on the time axis, or a discrete time process where events must occur at certain fixed intervals i.e. marks located at daily intervals.

The simplest continuous-time point process is the Poisson process in which the events can occur randomly at any point in continuous time. The Poisson process assumes that the time between events is independent and exponentially distributed. In addition, the number of events over a time interval is also independent and Poisson distributed.

In a Poisson model if a magnitude or rainfall intensity is attached to each occurrence, the process is known as a marked point process, i.e. a mark or magnitude of intensity for each rainfall event. If in turn these magnitudes do not have an associated duration, i.e. the entire magnitude/depth occurs instantaneously at each mark in the Poisson process then this can be referred to as a Marked Poisson or White Noise model. Early work on models of this type includes that of Todorovich and Yevjevich (1967) (see Figure 2.1) who considered the individual storm depths (U_i) associated with a Poisson arrival process to be gamma distributed. Todorovic (1968) and Todorovic and Yevjevich (1969) continued with this model structure but abandoned the gamma distribution for rainfall amounts in favour of an exponential distribution for seasonal precipitation.

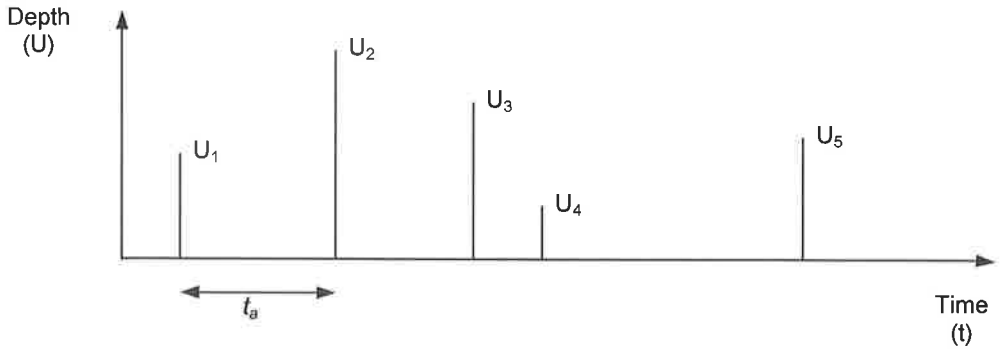

Figure 2.1: Marked Single Poisson Arrival Model / Poisson White Noise Model

Poisson White Noise models require the determination of 'independent' storm events to provide calibration storm data. In the latter work of Todorovic and Yevjevic (1969) they used the assumption that a consecutive sequence of rain observations was considered an event. For the case of hourly data, if an hour of rain was surrounded by two dry hours, then a storm of duration 1 hour is recorded. Similarly the daily record can be distinguished into independent storms by grouping consecutive runs of wet days located between dry days. In this way the number of storms and rainfall amounts over a certain time interval could be calculated and the model parameters calibrated.

Eagleson (1978) extended the basic Poisson arrival model by including rectangular rainfall pulses, rather than instantaneous rainfall bursts. He explicitly defined the time between storms or inter-event time (t_a) and the storm duration (t_d) as exponential distributions and used the previous two-parameter gamma distribution (Todorovic and Yevjevic, 1967) to model individual storm depths (see Figure 2.2). His model was calibrated using independent storm events from a 10-minute time resolution record and was successfully applied to estimate the distribution of annual rainfall given limited calibration data.

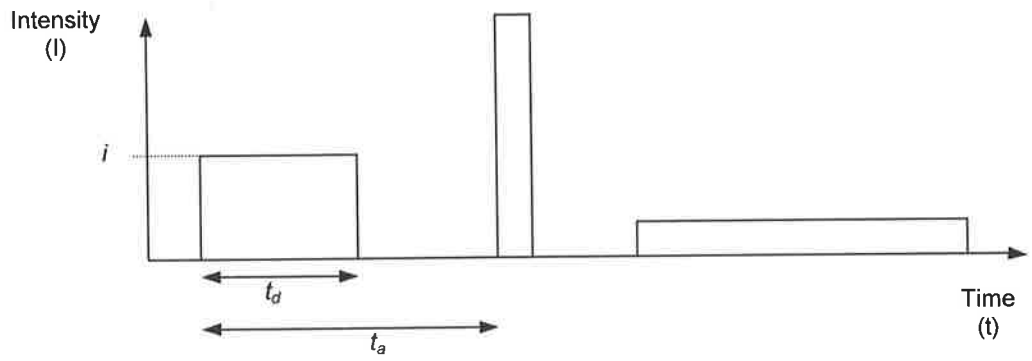

Figure 2.2: Independent Poisson Marks (IPM) Model / Rectangular Pulse Model (adapted from Eagleson, 1978b).

A further improvement to models of this type was proposed by Rodriguez-Iturbe *et al.* (1984). The Poisson Rectangular Pulse model developed by Rodriguez-Iturbe *et al.* (1984) was still characterised by an intensity (I_i) and associated duration (t_d) that were independent and identically distributed. The major advantage of their model over previous rectangular pulse models was the ability of rainfall pulses from different storms to overlap (see Figure 2.3).

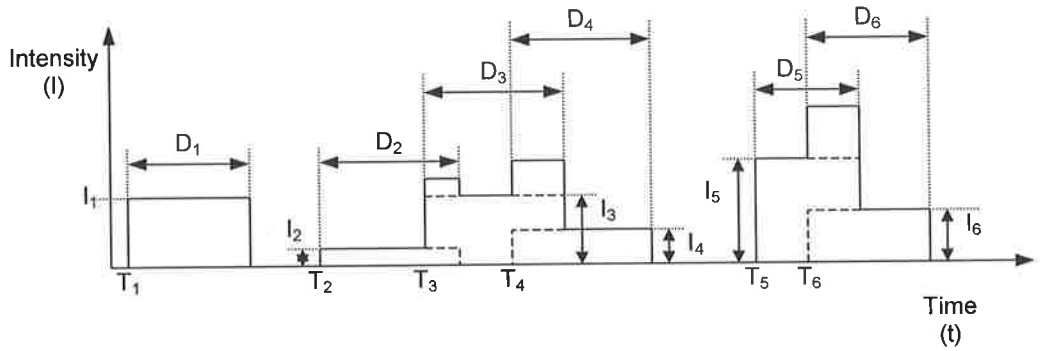

Figure 2.3 Poisson Rectangular Pulses Model (adapted from Rodriguez-Iturbe *et al.*, 1984).

This is in contrast to the earlier work of Eagleson (1978) where storm events were considered independent and did not overlap. The simulated intensity at a point in time is determined by the sum of intensities from each individual active storm event. The resultant simulation had a more realistic variation of storm intensity over time and ensured aggregated statistics over numerous time scales could be calculated. Rodriguez-Iturbe *et al.* (1984) exploited this result by deriving aggregated rainfall

moments at a given time scale and then used these parameters to fit to observed data. In a similar result to Todorovic and Yevjevich (1969), they noted that the nature of this process lends models of this type to be scale dependent given that parameter estimates and accuracy in results are different when using data at different scales (i.e. daily rather than hourly time scales). They also suggested Poisson White Noise models could not reproduce rainfall statistics for the hourly or daily scale due to the correlation present at these time scales and the underlying independence assumption used to form the Poisson model. Poisson White Noise models were also shown by Foufoula-Georgiou and Guttorp (1986) to be inadequate for representing short-time increment rainfall. Burlando and Rosso (1993) were also able to show that there was no significant improvement in extreme event estimation when shifting from a Poisson White Noise to a Poisson Rectangular Pulse model. They also suggested that these models did not provide an adequate description of the temporal variation of intensity evident in real rainfall events. In order to improve the reproduction of the important internal storm event structure, a further extension of Poisson models was developed which have become well known as Cluster models.

2.4.2 Cluster Models

Cluster models are an extension of the Poisson and therefore the Point Process approach and are generally a two-level process. At the primary level, rainfall generating mechanisms or storm origins occur according to a Poisson process. Each storm origin then gives rise to a group, or cluster of rain cells. Within a cluster, the distribution of rain cells is completely defined by the number and the distribution of their position with reference to the storm origin. The superposition of these rain cells provides the temporal definition of each storm event.

Two of the better-known cluster Poisson models are the Neyman-Scott and the Bartlett-Lewis models. Both models are able to take into account the apparent clustering of rainfall cells with respect to time. However, they differ in their treatment of the position of rain cells within the cluster structure. Neyman-Scott models explicitly define the distances from each cell to the primary storm origin and assign an appropriate distribution to this independent variable. In contrast Bartlett-Lewis models assume that the interval between subsequent cells is independently

distributed. This is the major difference between these two representations of the clustering nature of rainfall and can be seen in Figure 2.4.

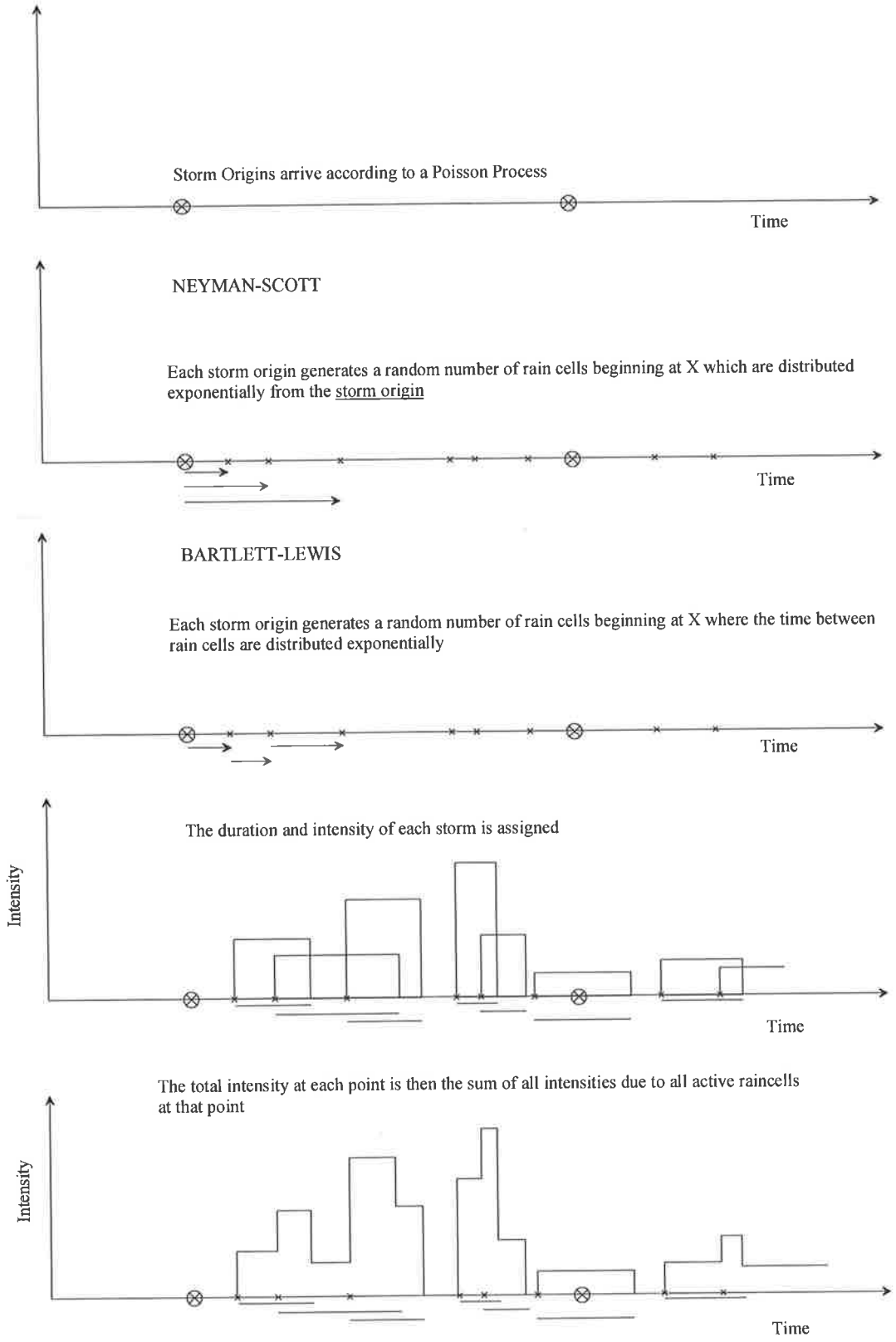

Figure 2.4: Schematic of the Neyman-Scott and Bartlett Lewis Models

The origins of the Neyman and Scott model trace back to the clustering structure being used to describe the spatial distribution of galaxies (Neyman and Scott 1958). Since then this structure has been successfully applied to model such things as the occurrence of earthquakes (Vere-Jones (1970), Lawrance (1972)) and then rainfall. Kavvas and Delleur (1981) first noted that the occurrence of daily rainfall in Indiana exhibited a clustering structure which could be modelled using the basic Neyman-Scott process. They used a geometric distribution for the number of rainfall events in a given cluster and an exponential distribution for the distances of events from their cluster centres. Their work concluded that the cluster model has the ability to preserve the dependence structure and marginal probabilities of the rainfall process, but the model form was homogeneous and could only be applied to stationary rainfall occurrences.

During their comparison of three rainfall models, Rodriguez-Iturbe *et al.* (1984) derived the second order properties of the accumulated rainfall amounts over different time scales for a Neyman-Scott White Noise Model. They then applied these results to fit a model to daily data in Denver, Colorado and Agua Fria, Venezuela. Rodriguez-Iturbe *et al.* (1984) observed that the Neyman-Scott model was superior in describing the rainfall process than the Poisson models for both the hourly and daily rainfall data they examined.

Valdes *et al.* (1985) re-examined the time scale dependency for the three models (Neyman-Scott White Noise, Poisson White Noise and Poisson Rectangular Pulse). This involved determining whether the parameters are consistent when estimated from data at different time scales. From their analysis, the Neyman-Scott process appeared the best option in terms of parameter stability. While their Neyman-Scott White Noise model outperformed other models over the various time scales that were analysed, their results indicated an inability of the model structure to preserve the extreme value distributions (storms at the extreme high end of the rainfall distribution).

Foufoula-Georgiou and Guttorp (1986) also studied the Neyman-Scott White Noise model in their analysis of event based data and concluded that the model cannot be time-scale invariant and that the application of the model should be restricted to the

time scale of the available data. They examined alternate fitting routines and noted the insufficiency of the second-order statistics in identifying the underlying continuous process, particularly when using daily data. Foufoula-Georgiou and Guttorp (1986) were also able to show that the choice of distribution to represent the cluster size was a great influence on the generated rainfall sequence and found a better fit through the application of a negative binomial rather than a geometric or Poisson distribution. While the Neyman-Scott White Noise model provided good agreement with observed rainfall, it was concluded that the model did not provide an adequate description of the underlying rainfall process and that model parameters were difficult to define due to a lack of physical meaning.

The inadequacies of previous models led Rodriguez *et al.* (1987a) to introduce the Neyman-Scott Rectangular Pulse and the Bartlett-Lewis Rectangular Pulse model. A Poisson and a geometric distribution for the number of cells within a cluster were considered for the Neyman-Scott model and a geometric distribution chosen for the Bartlett-Lewis. Both models assumed the distribution of storm cell duration to be exponential. Rodriguez-Iturbe *et al.* (1987a) derived the second order properties of the aggregated process for both the Neyman-Scott and Bartlett-Lewis models and the probability that an arbitrary interval is dry for the Bartlett-Lewis model. Cowpertwait (1991a, 1991b) extended this and provided an expression for the probability that an arbitrary interval is dry for the Neyman-Scott model.

In applying their new models and analysing data from Denver, Rodriguez-Iturbe *et al.* (1987b) concluded that they were capable of reproducing the observed statistics at various levels of aggregation but had problems reproducing the probability of zero rain and extreme rainfall values. They found that one set of parameters could effectively represent all levels of cumulative rainfall from 1 to 24 hours and that the parameters of the two models (Neyman-Scott and Bartlett-Lewis) were often identical. These results were reinforced by Cowpertwait (1991b) who applied the Neyman-Scott model to hourly data in England.

To improve the prediction of zero rain periods at each aggregation level, Rodriguez-Iturbe *et al.* (1988) proposed a modification to the Bartlett-Lewis model and introduced randomness into the mean rain cell duration parameter between events.

Application to hourly data in Denver and Boston showed an improvement in the reproduction of zero rain periods, however the simulation of extreme rainfall values remained a problem. Islam *et al.* (1990) also applied the modified Bartlett-Lewis model to hourly data in Italy and showed similar results.

Burlando and Rosso (1991) provided a comparison between the modified model structure introduced by Rodriguez-Iturbe *et al.* (1988) and Islam *et al.* (1990) and the original Bartlett-Lewis and Neyman Scott Rectangular Pulse models. Using the same hourly data from Italy, they were able to show that the modified model did not overcome the inadequacies of the original models and that the increase in complexity was not warranted. Burlando and Rosso (1991) also pointed out that while the Bartlett Lewis process allows for an easier mathematical framework and a larger number of relationships for calibration, it is regularly outperformed by its Neyman-Scott counterpart.

Entekhabi *et al.* (1989) followed a similar path to Rodriguez-Iturbe *et al.* (1988) and introduced modifications to the Neyman-Scott model in order to improve the reproduction of dry probabilities and extreme values. By applying a gamma distribution to randomly vary the mean rain cell duration, structural inter-storm variability was introduced. Similar to the results found by Rodriguez-Iturbe *et al.* (1988) with the modified Bartlett-Lewis model, there was a significant improvement in the reproduction of dry probabilities; however the reproduction of extreme values remained a problem.

Velghe *et al.* (1994) motivated by the increasing number of modified models in the literature provided a comparison between the original Bartlett-Lewis and Neyman-Scott models and the modified models introduced by Rodriguez-Iturbe *et al.* (1988) and Entekhabi *et al.* (1989). They confirmed an improvement on the original models in terms of the reproduction of zero depth probability and to a lesser extent the extreme values. However Velghe *et al.* concluded that the models had lost accuracy in reproducing the second order statistics due to the higher complexity of parameter estimation involved. They also found that the Bartlett-Lewis model was more sensitive to the moment equations used during calibration, and that the reported

improved results were the result of only one data set and were not sufficient to suggest that cluster models in general were suitable for modelling rainfall.

Cowpertwait *et al.* (1996a, 1996b) applied the Neyman-Scott Rectangular Pulse model in their work on sewer rehabilitation studies in the United Kingdom. The introduction of daily transition probabilities provided an improvement to the original poor reproduction of dry sequences and dry day proportions. Results indicated that the model was still not able to adequately reproduce extreme values particularly with a return period greater than 5 years. This was satisfactory for the purposes of a sewer rehabilitation study where events with return periods less than 5 years are of primary concern; however for the purposes of most engineering design practices, this is not adequate. An over-simplification in the parameterisation of the model in particular the averaged intensity of the model's rain cells, was given as a reason for the difficulty in reproducing extreme values.

Onof and Wheater (1993) provided a further extension to the Bartlett Lewis model by initially incorporating a gamma distribution to describe rain cell duration. Results indicated an overestimation of the auto-correlation statistics and mean inter-event times; however the model did improve the reproduction of dry interval proportions. The simulation of hourly data indicated additional problems with the overestimation of extreme values for return periods greater than two years and the average event duration. Further modification with the superposition of a jitter process on each rectangular pulse was incorporated by Onof and Wheater (1994). This was introduced to provide a more realistic representation of the rainfall process and to improve the auto-correlation results. While improvements in the reproduction of extreme events and auto-correlations were evident, difficulties were encountered during parameter estimation due to the large number of model parameters.

Gyasi-Agyei and Willgoose (1997) followed Onof and Wheater (1994) and developed a combination model. Based on the Bartlett-Lewis model, they incorporated an auto-regressive jitter process to fix the deficiencies in the modelled second-order properties. Using 15 minute data from central Queensland, they were able to show improved results when compared to the original Bartlett-Lewis model (Rodriguez-Iturbe *et al.* (1987)) and the modified randomised Bartlett-Lewis model (Rodriguez-

Iturbe *et al.* (1988)). Gyasi-Agyei and Willgoose (1999) further generalised their model by replacing the Bartlett-Lewis model with a binary chain to represent the rainfall process. The binary chain preserved the dry and wet sequences as well as the rainfall mean while the correlated jitter was again employed to improve deficiencies in the second order properties. Two possible binary chain models were considered (a non-randomised Bartlett-Lewis model and a Markov chain model) with the Bartlett-Lewis model being preferred due to a lesser number of parameters. While the authors show improved results in comparison to earlier models, no aggregation levels greater than 24 hours, Intensity-Frequency-Duration (IFD) curves or extreme values were shown.

Calenda and Napolitano (1999) provided an investigation into the estimation of parameters for the Neyman-Scott models. Their calculations showed that the estimation of model parameters by the method of moments is significantly affected by the choice of the aggregation scale of the data and if equations at different aggregation scales are used, the difference in these scales is also an issue. When the aggregation scales chosen are too close, the resulting objective function is very flat and ensures the optimal parameters are difficult to find.

Cameron *et al.* (2000) provided an evaluation of an exponential model adapted from Eagleson (1972), a data-based model from Cameron *et al.* (1999) and the random parameter Bartlett-Lewis Gamma Model of Onof and Wheater (1994). While they were able to again show the merits of using a cluster based model to represent rainfall and general statistics, the reproduction of extreme values was poor. To overcome this, Cameron *et al.* (2001) introduced a generalised Pareto distribution to represent depths of high intensity rain cells and improve the reproduction of extreme values. They concluded that the reproduction of previous aggregated statistics was reasonably consistent and that the modelling of the extremes was improved.

Koutsoyiannis and Onof (2001) introduced a disaggregation process for the generation of hourly data that aggregate up to the given daily data totals. Their work combined a cluster Bartlett Lewis process with a process to adjust the finer time scale data so as to obtain the required coarser data. They used data from the United

Kingdom and the United States to test the performance of their methodology and were able to preserve most of the statistical properties of the rainfall process.

Extending the single site disaggregation, Koutsoyayiannis, Onof and Wheater (2003) developed a methodology for the spatial-temporal disaggregation of rainfall. Using a hybrid model with temporal characteristics based on the Bartlett-Lewis Rectangular Pulse model, they combined univariate and multivariate rainfall models operating at different time scales. These models were used to derive spatially consistent hourly rainfall series in areas where only daily data is available. While providing encouraging results, the authors noted deficiencies in the results with over predicting very low values of dry periods and simulating low intensity during events when the intensity is actually zero.

All of the work on Poisson models discussed above rely on near perfect data sets for calibration. Their calibration process traditionally relies on aggregation statistics which can be corrupted by sections of erroneous or missing data. Cowpertwait (1991a) proposed that if only a few data points were missing in a month of data, these were to be taken as zero and if a significant number of points were missing that month was to be deleted from the record. If numerous months were deleted, then that year was discarded from the record. If a different month was missing from the same record, then the corresponding month in the deleted section is inserted. Cowpertwait had the advantage of dealing with rainfall records that contained very small amounts of corrupted or missing data (1% missing). Numerous data sets in Australia have a missing or corrupted percentage closer to 6% - 10% (Heneker *et al.* (2001)). Gyasi-Agyei (1999) in their application to Australian data discards entire months with any missing data, ensuring valuable information is lost. This may have an adverse impact on the ability of the model to replicate the observed statistics given the increase in missing data may affect the calibration statistics, particularly for short records.

While Poisson models and in particular cluster models are extremely popular in the literature, questions remain as to the ability of these models to represent zero depth probabilities and extreme values. The reproduction of extreme values is often critical

in a design situation and these models must be used with particular attention to this detail particularly at the small time scale.

2.4.3 Markov Models

An alternative approach which can be used to reproduce the rainfall occurrence mechanism and any associated dependence structure is the Markov Chain. Markov chains work on the basis that the present state for a given time interval is dependent upon the state(s) of the previous time interval(s). A first order Markov chain calibrated to simulate daily data would simulate day x based upon the state of day x-1. A second order model would take into account the states of day's x-1 and x-2 etc. This same Markovian structure can be applied to models that are required to work at smaller or larger time scales. The controlling factor in a Markov chain is the transition probabilities, which govern the chance of observing a particular new state, given the previous state(s) of the model. Markov models are completely defined by the initial probability distribution and their transition probabilities, which are usually in the form of a probability matrix. Waymire and Gupta (1981a) provide a good mathematical summary of the Markov process. Assuming the variable (Y_i) to be Markov dependent, let the transition matrix be denoted by

$$P = \begin{bmatrix} p_{DD} & p_{WW} \\ p_{DW} & p_{WD} \end{bmatrix}$$

Where p_{DD} is the probability that the ith interval is dry given the previous (i-1)th interval is dry and is described as:

$$p_{DD} = \text{prob}(\text{dry} - \text{dry}) = P(Y_i = 0 \mid Y_{i-1} = 0) \qquad i = 1,2,...$$

and

$$p_{WD} = 1 - p_{DD}$$

while p_{WW} is the probability that the ith interval is wet given the previous (i-1)th interval is wet and is described as:

$$p_{WW} = \text{prob}(\text{wet} - \text{wet}) = P(Y_i = 0 \mid Y_{i-1} = 1) \qquad i = 1,2,...$$

and

$$p_{WD} = 1 - p_{WW}$$

A graphical representation of a first order two state Markov chain can be seen in Figure 2.5.

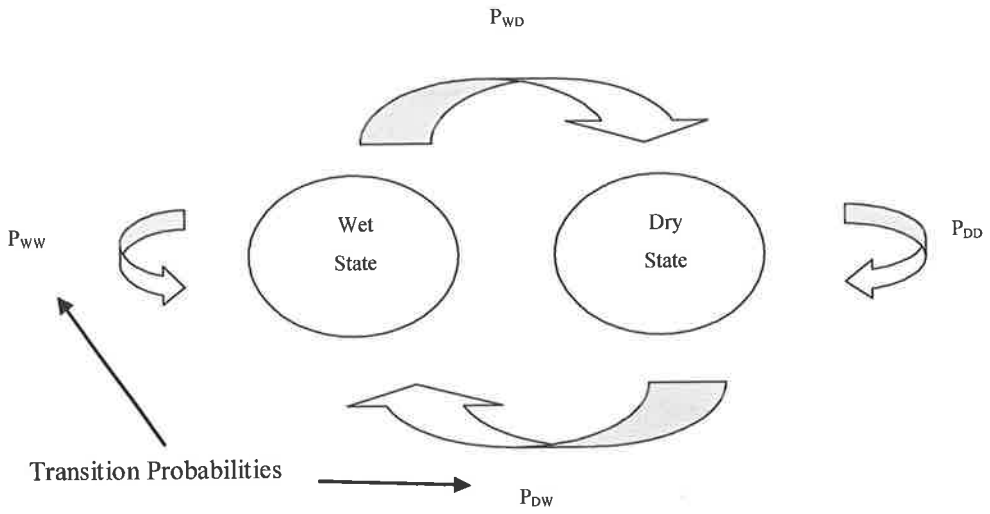

Figure 2.5: Two-State First Order Markov Model Structure

Markov chains have been a popular method for representing the daily occurrence process due to their explicit description of the dependence structure evident at this scale. Gabriel and Neumann (1957, 1962) were one of the first to apply a first order Markov chain to model the daily rainfall occurrence process with some success. Their Markov chain consisted of constant transition probabilities over the year (a homogeneous parameter set), and found that their model was adequate to describe data for Tel Aviv. Adapting this approach, Caskey (1963) and Weiss (1964) incorporated variation in their transition probabilities over the year (non-homogeneous) and applied this to several cities in the United States with varying success.

A first order Markov chain was applied by Hopkins and Robillard (1964) to data from Canada. Their model was unable to describe the daily rainfall process in months that historically contained very little rain with only a few rain days. Green (1964) showed that the first order Markov chain proposed by Gabriel and Neumann (1957, 1962) was outperformed by an alternating renewal model when comparing

certain conditional probabilities (See Section 2.4.4 for a description of the alternating renewal process). In further findings, Green (1965) concluded that the geometric memory of the first-order Markov chain is not adequate to describe long droughts or wet spells. This result is particularly important for models calibrated to Australian data where often long term dry spells are evident. Feyerherm and Bark (1967) also showed that a first order model was inadequate in describing the higher-order dependence of daily rainfall in Indiana and introduced a second order Markov chain for this purpose.

Smith and Schreiber (1973) suggested that the earlier work of Weiss (1964) and Gabriel and Neumann (1962) was based on what could be referred to as frontal rain storms, and extended this earlier work by analysing air mass thunderstorm data in North America. They compared a first order Markov chain to a simple independent event-based Bernoulli model. The Markov chain model was able to describe numerous statistical properties of the thunderstorm occurrence process. However, the authors conceded that there is no evidence that this occurrence process is a simple Markov structure or that a higher order model would not improve results. In order to increase their Markov chain model's accuracy from year to year, an additional annual variance was incorporated on top of the nonhomogeneous transition probabilities.

Initially the application of Markov chains was predominately restricted to replicating the occurrence process of daily rainfall and less attention was focused on incorporating a model for rainfall amounts. Haan *et al.* (1976) combined the simulation of both the rainfall occurrence process and depths via a first order Markov chain with a seven-state (one dry, six wet) transition probability matrix. Rainfall was divided into classes, where the class boundaries for the states of the Markov chain were found using a geometric progression. The distribution of rainfall amounts within a class were considered uniform with the exception of the last class in which a shifted exponential was used. While the seven-state model performed well, there was an over-estimation of simulated rainfall amounts in each class, which in turn produced a consistent error in annual rainfall totals.

Pattison (1965) modelled the hourly rainfall process by the combination of a first order and a sixth order Markov chain. During the simulation procedure, if the state of the previous hour was wet, the model used a first order dependence structure to determine the current state. If the previous hour was dry, the model used a sixth order structure to determine the state of the current hour. Pattison simulated rain amounts by incorporating 20 rainfall states and using transition probabilities. These transition probabilities were varied from month to month. While the model was capable of replicating the characteristics of the hourly rainfall process during storm periods, dry periods between storms were overestimated.

Srikanthan and McMahon (1985) developed multiple state models for the generation of daily, hourly and six-minute data and applied them to locations around Australia. They used a sliding scale technique where generated rainfall values were consecutively scaled using numerous transition probabilities from the original daily value to hourly and then to 6 minute interval rainfall events. The daily model consisted of seven wet states with associated transition probabilities calibrated monthly. The simulated daily depths were a result of the current state of the model. For the largest wet state, a normally distributed random variable was transformed via the Box-Cox technique to give large simulated daily depths. The remaining wet states in the model employed a similar procedure using a linear distribution. Results indicate a reasonable comparison to various aggregated statistics, but the authors acknowledged a less than satisfactory result in the reproduction of several maximum daily rainfalls, wet and dry day runs and mean rainfall depths for three types of wet days.

Their generation of hourly rainfall was developed as a two-stage process. Initially the daily model was used to determine whether a day was wet or dry and given that a day was wet, the type of wet day. Wet days were divided into two types based on the daily rainfall depth and a dividing rainfall threshold. For each type of wet day a time dependent second order Markov chain was used to determine whether each hour in the wet day was wet or dry. An hourly transition probability matrix was finally used to generate the hourly rainfall depths. The use of hourly transition probability matrices that vary from month to month produces a model with a large number of parameters. While this was reduced by an adjustment to the model that grouped

hours into blocks of four, the number of parameters is still large. A number of statistics were compared to the observed data with a concession from the authors that the results are less successful than the daily model.

Srikanthan and McMahon's (1985) 6-minute model was a combination of three sub-models, a daily, hourly and a 6-minute model. The 6-minute model followed a similar progression to the hourly model with the exception that hourly rainfall was now also divided into one of four types. A 6-minute transition probability matrix was incorporated to generate 6-minute rainfall values over a wet hour. If the hourly rainfall depth exceeds 5mm, then continuous rain is assumed over the hour. While showing satisfactory results, the increase in parameters to use this type of model at the 6-minute level limits its application.

Rajagopalan *et al.* (1996) presented a nonhomogeneous Markov model for daily precipitation. They assumed that the transition probabilities from state to state vary smoothly over the year and estimated the corresponding Fourier series using non-parametric techniques. This enabled the model to be fitted for the entire year rather than in homogeneous seasons or months, as is usually the case. While the model was able to reproduce numerous statistics, the application of the non-parametric fitting approach limits the extrapolation of daily rainfall values beyond the observed maximum. This is also an underlying problem of Markov models in general as the estimation of transition probabilities generally relies on the observation of transitions in the historical record. For short records, or records with large erroneous or missing sections of data this could introduce a bias to the calibrated probabilities.

Jimoh and Webster (1999) showed a first order Markov model was capable of describing the occurrence process of wet and dry days in Nigeria. Their focus was on techniques incorporating variation in the transition probabilities over the year. The use of a Fourier series, averaging techniques or a combinations of these were found to be equally as good at providing the required non-homogeneity of model parameters.

Although Markov chains appear to provide a simple mathematical model of rainfall, their dependence structure cannot describe the long-term persistence (droughts and

floods) evident in short time-increment rainfall records. The frequency and implication of these long dry periods in Australia ensures their accurate description is of particular concern. When considering short time increment rainfall records (high resolution; 5,6,10 minutes), small order Markov chains are also unable to describe the clustering effect present at such a fine time scale. Increasing the order degree of a Markov chain to successfully describe the complexities of the high resolution rainfall process would present a significant computational challenge and with alternate effective models available, this approach is unnecessary. Of additional concern is the reliance and restriction evident through estimating transition probabilities from short records, or records with large erroneous or missing sections of data which can introduce a bias to the calibrated probabilities. Disaggregation techniques can be employed to enable a daily Markov model to generate sub-daily time scales, however the extra model and subsequent parameterisation increases the models complexity and possible regionalisation effort.

While Markov chain models may be adequate for some sites and some seasons, taking the above issues into account and additional concerns through regionalising Markov chains due to their inherent site specific nature and the lack of discernable physical meaning in their parameters, a more appropriate rainfall model structure was sought.

2.4.4 Alternating Renewal Models

Alternating renewal models (or Wet-Dry Spell models) are event based models which replicate the occurrence process of rainfall by simulating wet and dry storm events. These events are continuous periods of mostly wet or continuously dry observations that are assumed independent and separated from previous observations by a minimum independence criterion (typically a minimum period of no rain). It follows that the event occurrence process for an alternating renewal model is completely defined by the probability distributions that describe the lengths of these wet (t_a) and dry (t_d) events (see Figure 2.6). The simulation proceeds by sampling alternatively from the dry and wet spell probability distributions, until the required length of record is reached. An additional probability model is then used to describe the intensity (i) or depth of each wet storm event.

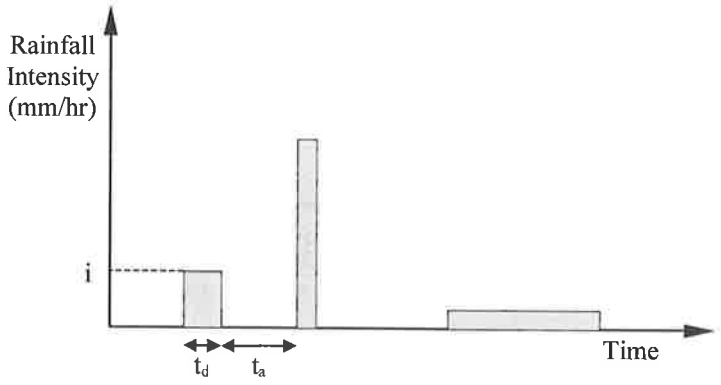

Figure 2.6: Schematic of Alternating Renewal Process

A major advantage of the alternating renewal structure is the ability to calibrate to data sets with sections of missing historical data. As the model is event based, missing sections of data will only influence the calibration if these periods consistently contain rare storms not found anywhere else throughout the record. Given that missing sequences from data records are typically random, this is usually not an issue. The ability to use data sets that contain missing data is a major advantage over models that calibrate to aggregation statistics (see for example Poisson models Section 2.4.1) or rely on a continuous data set to determine transitions between rainfall types (Markov models). Once the simulated time series and gross rainfall amounts has been simulated, each individual storm event can then be disaggregated to the required time scale providing a model that is capable of providing synthetic high resolution rainfall data.

Green (1964) compared an early alternating renewal model to the earlier Markov model of Gabriel and Neumann (1957, 1962). It was assumed that the sequence of wet and dry days formed an alternating renewal process and Exponential distributions were chosen for the lengths of dry and wet spells. The renewal model out performed the Markov structure for data at Tel Aviv, however neither model worked well for data from Chesire, England. The work of Green provided one of the first definitions of what constitutes a 'storm' event.

In an Alternating Renewal Model (or Wet-Dry model) the historical series must be divided into independent wet and dry events. The independence criterion is usually a

set length of dry period which is used to separate consecutive events. For the case of daily data (as analysed by Green (1964)) a single dry day is often chosen as this minimum dry or minimum inter-event time. Thus a sequence of consecutive wet days is considered to belong to the same wet storm event until a dry day is observed. Any subsequent sequence of dry days is then part of the current dry event until a wet day is observed and so forth. It is understood that 'real' physical storms may have periods of no rain or dry periods as part of their temporal structure, however due to the coarse nature of daily data, this fact is overlooked. For the purposes of high resolution data (5, 6, 10 minutes) a minimum dry time in minutes or hours is usually chosen and used to define independent storm events. Once a wet event begins, the event continues until a dry period is observed greater than this minimum dry time. This allows the definition of storm events for high resolution data to proceed in a similar fashion to the simpler definition for daily records; however it does have the capacity to incorporate dry periods less than the minimum inter-event time to be included in wet events thus providing a more realistic representation of a storm event.

Grace and Eagleson (1966, 1967) extended the earlier daily work of Green (1964) by introducing an alternating renewal model calibrated to a ten-minute rainfall record. The minimum inter-event time used in their study was obtained by serially correlating the historical rainfall depth series. Through testing the rank correlation coefficient the lag at which there was no significant dependence was determined. A lag time of 140 minutes was found to provide this independence criterion for data in Vermont. The Weibull distribution was calibrated to the distributions of inter-event times and storm durations. A regression relationship between event depths and durations was developed with a Beta distribution fitted to the residuals to incorporate a conditional depth – duration relationship in order to simulate rainfall amounts. A similar technique for defining the minimum inter-event time was also applied by Sariahmed and Kisiel (1968) in their work on representing summer thunderstorm occurrence. A minimum inter-event time of 3 hours was chosen and the Weibull distribution used for the distributions of inter-event times and storm durations. These models were able to show satisfactory results. Grayman and Eagleson (1969) looked at hourly data from Boston and found the storm duration, depth and time between storms were exponentially distributed.

Restrepo-Posada and Eagleson (1982) provided an alternative approach to the definition of storm independence. In contrast to the earlier work of Grace and Eagleson (1967) and Sariahmed and Kisiel (1968), they argued that the choice of analysing rainfall depths for the purposes of independence was flawed and that it was more appropriate to investigate the independence of successive inter-event times themselves for this purpose. By assuming a Poisson arrival process and therefore inferring that the distribution of inter-event times was exponentially distributed, they were able to calculate an independence criterion to separate independent storms. An iterative procedure was introduced to calculate an optimal minimum inter-event time that ensured the resultant distribution of 'independent' inter-event times was sufficiently exponential. This simple algorithm given the assumption of an exponential distribution for inter-event times was shown to work well at three arid-climate sites.

Koutsoyiannis and Xanthopoulos (1990), Koutsoyiannis and Foufoula-Georgiou (1993) and Koutsoyiannis and Pachakis (1996) also developed rainfall models based on a Poisson process and followed the assumption of Sariahmed and Kisiel (1968) that the distribution of inter-event times from independent events must follow an exponential distribution. To satisfy this assumption, a minimum inter event time was calculated and later used to distinguish independent events. Koutsoyiannis and Xanthopoulos (1990) found this value ranged from five to seven hours for hourly data while Koutsoyiannis and Foufoula-Georgiou (1993) and Koutsoyiannis and Pachakis (1996) determined a time of seven hours, based on a Kolmogorov-Smirnov test of the exponential distribution.

Relationships and comparisons can be made between the alternating renewal model and models of the Markov Chain type. An alternating renewal model is conceptually similar to a Markov Chain where the probability of simulating a dry spell after a wet spell is equal to unity without correlation. Roldan and Woolhiser (1982) compared an alternating renewal model to a first order Markov Chain. The sequences of wet and dry days were simulated using a truncated geometric distribution of wet day intervals and a truncated negative binomial distribution of dry days. They found the Markov Chain to be superior at the four US cities studied. Small and Morgan (1986)

derived a relationship between a continuous alternating renewal model and a Markov chain for the occurrence of daily rainfall. The gamma distribution was used for the distribution of dry intervals and the wet intervals were assumed to be exponential. They found that the Markov model worked well in some areas of the United States, however in other areas the clustering structure apparent in rainfall occurrence could not be modelled using a Markov Chain and was more accurately represented by the alternating gamma model.

Foufoula-Georgiuo and Lettenmaier (1987) developed a Markov renewal model to analyse daily rainfall. In simple terms a Markov renewal process is different to the generic Markov chain as the probability of observing a wet day does not depend solely on the state of the previous day (wet or dry) but on the number of days since the last wet event. The model presented by Foufoula-Georgiuo and Lettenmaier (1987) assumed that times between wet events or inter-arrival times belonged to either one of two types. Geometric distributions were used to describe the inter-arrival times for each type, with an overlying Markov structure governing the transitions from one type to the other. Rainy periods followed each dry interval and within rainy periods, the model behaves exactly as a Markov chain. (i.e. regular transition probabilities are employed to determine whether the next day is wet and remains part of the current wet event, or a new dry period begins and is therefore sampled from the predetermined type 1 or 2 dry distribution). In order to distinguish the historical events from the record, wet events were defined as any day with measurable precipitation. Finally the distribution of rainfall amounts was described by an exponential distribution. The model was able to preserve the daily statistics for data in Washington, but no work was undertaken on data at finer time scales. It is also unclear as to the motivation behind the two types of dry periods. While the authors claim that one type may relate to dry periods between major storm fronts and the second to subsequent dry periods which occur in the aftermath of storm events, the authors note that the inter-arrival times from the observed record cannot be classified directly as belonging to one type or the other. Only probabilistic classification is available, ensuring these parameters have no physical meaning, are not easily identifiable and difficult to extrapolate or regionalise to other sites.

Hutchinson (1990) combined previous work on Markov chains and renewal models to develop a three-state continuous Markov model. The first state of the model represents a dry spell (state 1) which is always succeeded by state 2, state 3 is identified as a wet spell and is also always succeeded by state 2 and finally state 2 is denoted as a transition dry spell state which can be succeeded by either state 1 or 3. As the mean duration of dry events from state 1 is greater than that of dry and wet events in states 2 and 3, the resulting simulation produces rainfall events which consist of a cluster of showers (consecutive periods of state 2 and 3), similar to the those of cluster based models. However as durations are incorporated directly, the model structure ensures that showers occur sequentially unlike cluster models which allow showers to overlap. This provided advantages over the previous cluster models in terms of physical interpretation of parameters, mathematical tractability and parameter estimation. A mixed geometric distribution was used to describe the durations of dry periods and the duration of overall rainfall events. Intensity was included via an exponential process, which was auto-correlated for the duration of each overall event, but independent from one event to the next. In effect, this provides a representation of the temporal pattern of rainfall and was considered an improvement from previously accepted ideas of assuming intensity to be independent of the duration of the shower and independent of the intensities of other showers (Rodriguez-Iturbe *et al.*. (1987a,b)). Incorporating a correlation between intensities within an overall storm provides a similar structure to alternate models that employ a temporal pattern to disaggregate a uniform intensity pulse. However the model of Hutchinson (1990) still ignores any possible correlation between event duration and intensity as the intensity of an overall storm event is only correlated to other event intensities within the storm and not the duration of the storm. While the model had difficulties replicating the observed wet run statistics, it was shown to outperform the earlier Bartlett-Lewis model of Rodriguez-Iturbe *et al.* (1987b) for non-calibrated statistics.

Lall *et al.* (1996) used non-parametric techniques to describe the distributions of dry spells, wet spells and rainfall amounts in their alternating renewal model of daily rainfall. Their model also incorporated extra information in terms of the joint and conditional distributions of wet and dry spell durations. While their model was able

to reproduce some of the daily statistics for data in Utah, the sample sizes required for their model is significant and limits its application.

Wong (1996) used an alternating renewal model structure with monthly parameters to simulate a synthetic 6-minute rainfall record. The first order Markov equation with a transformed gamma distributed random variable was used to generate inter-event times and storm durations. The generation of event intensity was based on the intensity-frequency-duration curves for the site. Curves were estimated from the historical record for storm durations of 0.5,1,3,6,9,12,18,24 and 72 hours. Other durations were linearly interpolated from these values. Presented results for simulated inter-event times and storm durations showed a reasonable comparison to the observed values however the model generated some unrealistic long storms prompting the proposition of an upper limit on the simulation. A discrepancy in the number of storms (7%) between simulated and observed was also noted.

Lambert and Kuczera (1996) undertook an extension of the earlier work of Eagleson (1978) and described the distributions of inter-event times using a gamma distribution, storm durations using an exponential distribution and the corresponding intensity by a generalised Pareto distribution (Rosjberg *et al.* (1992)). They described an intra-storm disaggregation scheme based on a constrained random walk through dimensionless depth-time space. This disaggregation scheme was developed to circumvent one of the disadvantages of wet-dry modelling in that once a storm has been identified, the internal characteristics of that storm need to be reproduced during the simulation. The disaggregation scheme proposed by Lambert and Kuczera (1996) ensured the internal rainfall patterns remained, however an underestimation of IFD data was evident for longer durations.

Lambert and Kuczera (1998) continued to work in this area and introduced a generalized exponential probability model. The structural development of this model enabled the authors to remove bias arising from the binned nature of observed rainfall. Typically rainfall records are recorded as discrete depths (mm) over a finite time period (minutes, hours or days). This binned nature of recorded rainfall leads to assumptions and uncertainty when using the resultant data. For example, if a storm lasts for 5 bins in a 6-minute format, i.e. there are 5 'wet' bins in a sequence, the

duration of this storm sits somewhere between 30 minutes and 18 minutes in reality. This can be seen in Figure 2.7 where it could have been raining for the exact length of the 5 bins or it may have started raining just before the second bin and finished just after the fourth bin. Thus the exact storm duration is indeterminate and the exact beginning or end of the event can only be determined to the degree of accuracy of the bin widths. This was taken into account in the model and likelihood function developed by Lambert and Kuczera (1998).

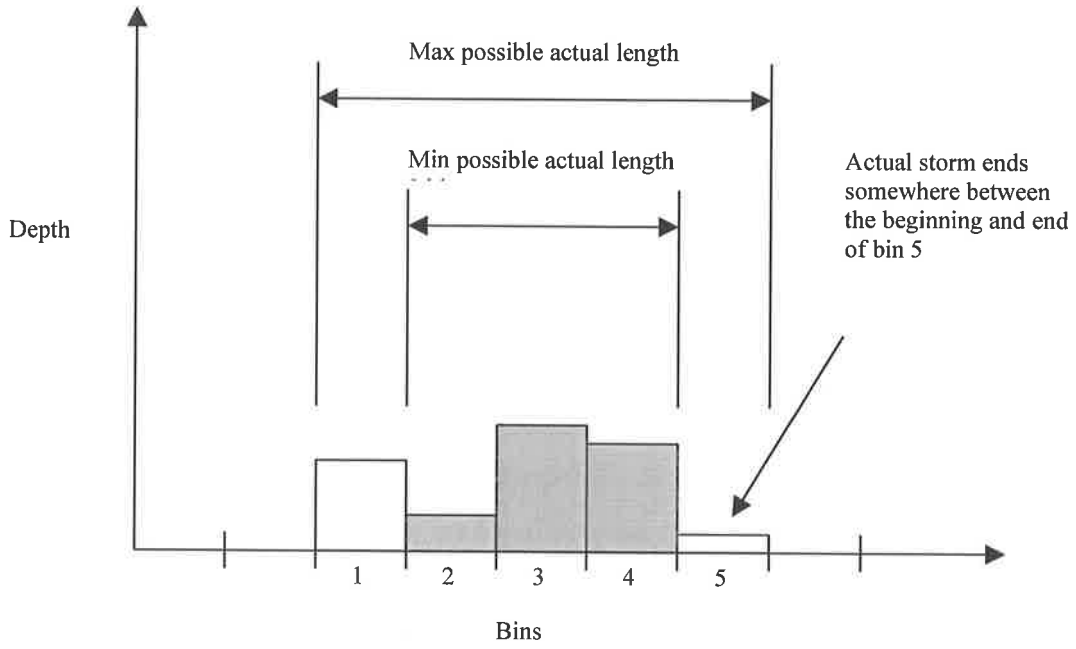

Figure 2.7: Description of the binned nature of rainfall

Heneker *et al.* (2001) extended the previous work of Lambert and Kuczera (1996, 1998) and other event based models to overcome some of their shortcomings. Storm events were extracted from six-minute historical records and again the generalised exponential distribution was used to describe the distributions of inter-event times and storm durations. The intensity of each wet event was described by a Pareto distribution with the corresponding parameters conditional upon storm duration. It was shown that this model was not only capable of reproducing the distributions of dry spell and storm durations that were used in the calibration, but

also statistics that were not introduced into the calibration procedure. Of significant interest were the simulated and observed Intensity-Frequency-Duration curves which showed good comparison for various cities, and resulting aggregation statistics. These values were not used in the calibration of the model and the reproduction of these values supports the credibility of the model.

2.4.5 Discussion of Point Rainfall Models and the Selection of one for Further Development and Regionalisation

Point rainfall models have been developed to reproduce the structure of rainfall using various techniques. Earlier models were often based on the theory of Markov chains with their explicit dependence structure. More recently event based alternating renewal models and to a greater extent Poisson cluster models have received substantial coverage in the literature. Investigation into published models and previous results provides an insight into the selection of a suitable model for the purposes of this study,

Poisson models and in particular cluster models continue to receive substantial literature coverage. Their structure ensures that the conditional relationship between intensity and duration cannot be modelled explicitly even though it is generally accepted that this dependency exists (Grace and Eagleson, 1966, 1967; Acreman, 1990; Lambert and Kuczera, 1996). Another serious concern for the application of these models is when the historical record contains missing or erroneous data periods.

The length of most rainfall records ensures the chance of observing high ARI events is unlikely. Periods of missing data exacerbate this issue by increasing the chance that high ARI events have not been recorded. While the amount of missing data can significantly affect the application of a chosen model, very few researchers publish how to adequately deal with this. This is particularly an issue for models which rely on aggregated statistics for calibration such as the cluster models. Cowpertwait (1991a) described a replacement strategy to handle missing data in his work on the

Neyman –Scott model. Missing daily data within a particular month was replaced with data from the same month within a different year but where the monthly totals were similar. If many months were missing then the whole year was deleted or if only a few data were missing those values were replaced by zero. Although this approach may be adequate for some data sets, there is no confidence that this approach will be adequate to compensate for significant missing or rejected data.

Numerous authors claim that Bartlett-Lewis and Neyman-Scott based models are able to reproduce a variety of rainfall statistics over different levels of aggregation. Frequently published results indicate an inability of these models to replicate statistics not used during the calibration process. The application of these models to historical data sets with missing or erroneous data periods is also a concern. While techniques have been suggested to circumvent this problem (Cowpertwait (1991), Gyasi-Agyei (1999)) questions remain as to the influence and effectiveness of these techniques given the quality of historical records. In addition to this, many researchers have indicated that parameters for the cluster models are difficult to estimate. This is partly because they are not intuitive or easily observed from the historical data. Foufoula-Georgiou and Guttorp (1986) suggested that since the N-S model does not provide an adequate description of the underlying rainfall generating process, no physical meaning should be attached to the parameters. Even though some parameters of the cluster models have been given a physical interpretation, such as the duration of a rain cell, they are not readily determined directly from a rainfall record. Onof and Wheater (1994) showed for their random-parameter Bartlett Lewis Model that as the number of parameters increased, identification of these parameters became more difficult and alternative identification strategies gave significantly different values for the same parameter. Given this parameter sensitivity, doubts remain as to the robustness and stability of the model and its results. Velghe *et al.* (1994) further confirmed that cluster models are very sensitive to the selection of calibration equations used. Calibration with one set of equations provided a poor reproduction of the percentage of dry intervals. In order to improve this result one of the equations was replaced but this in turn produced a poor reproduction of zero depth probability.

Not only are cluster models difficult to calibrate and contain parameters with little physical interpretation, the underlying process during calibration is of concern. Cluster models are calibrated using aggregated rainfall data, where the methodology is based on a continuous process. As a result, the discreteness of the data is not taken into account. Foufoula-Georgiou and Lettenmaier (1986) suggested that the problem is the result of the inappropriate assumption that a continuous-time point process, can be calibrated against an aggregated record. Foufoula-Georgiou and Lettenmaier (1986) further showed this to be the case for all continuous-time stochastic models. They concluded that using aggregated rainfall data to calibrate continuous-time point process models introduced biases in parameter estimation that can result in misleading interpretations regarding observed rainfall clustering. Several researchers have also indicated that the inferred description of the underlying process is dependent upon the scale at which the model is fitted (Rodriguez-Iturbe *et al.* (1984), Valdes *et al.* (1985)). This provides limitations on the models ability to extrapolate to other time scales and the ability to model the properties of the continuous process. In contrast, parameter estimation is straightforward for event-based models, which if continuous, allow the incorporation of estimation procedures to account for aggregated data.

Markov models provide a simple description of the rainfall occurrence process. Typically these models have been applied to represent daily rainfall however some authors have extended this to finer time scales. Markov chains have been shown to work at specific sites and for specific seasons but have not been shown to consistently model the rainfall structure or process. When applied to short time increment rainfall, a large number of parameters are introduced while still being unable to describe the clustering effect (higher chance of a wet increment occurring during a storm event than during a dry period) evident in historical records. In addition to this, some existing models use a number of wet states corresponding to different types of rain during a wet increment. These states require probability distributions introducing more parameters into the model. Large number of parameters with little or no physical meaning provides problems during the calibration process creating a model which is difficult to apply. If a Markov chain is applied to a historical record with significant missing data periods, the determination of state transitions and hence model parameters is hampered. This can also be a

problem at numerous Australian sites where there are a limited number of wet events in a given season or month.

Alternating renewal models have been shown to perform well over numerous levels of aggregation and preserve statistics that were not used during the calibration process. This is an important result particularly when considering a rainfall model for regionalisation as it provides credence to the model structure and introduces a level of confidence on the outcome of the regionalisation process. Alternating renewal models are unique as they are not restricted or hampered in their application due to erroneous or missing data periods. Given that missing data sections are random (i.e. missing event durations and depths are not consistent), then no significant bias is introduced through the removal of these periods from the record. Indeed a significant advantage of models of this type is the ability to include good data sections of a month where missing data is present. This ensures more information is available for calibration to the same length of record in comparison to cluster based models.

Like Markov models, parameters are easily defined and have a definite physical meaning. This ensures model parameters are intuitive and can easily be estimated from the historical record. Even if the renewal model is continuous, estimation procedures can be incorporated to account for the aggregated and binned nature of observed rainfall records. Two drawbacks in the use of alternating renewal models are the requirements of separating independent storm events and the need to disaggregate wet events down to the time scale of interest. Previous research has shown that these issues are easily handled by the selection of an appropriate model and or structure and numerous authors have presented model results that compare favourably to observed records.

The above discussion led to the selection of an alternating renewal model for further development in this study. The systematic and extensive model development provided by Lambert and Kuczera (1996, 1998) and Heneker *et al.* (2001) coupled with good results for calibrated and non-calibrated statistics has ensured this particular alternating renewal model was chosen. The original Heneker *et al.* (2001) model is presented in further detail at the beginning of Chapter 3. It is important to

note that while the Heneker *et al.* (2001) model was chosen for further development in this thesis, most of the techniques presented in this study can be applied to other alternating renewal models or alternatively to cluster based models with little or no adjustment.

2.5 Regionalisation Techniques

Regionalisation has for many years been a standard hydrological tool, used to facilitate extrapolation from sites at which records have been collected to others at which data is required but unavailable (Riggs, 1973). Typically regionalisation of hydrologic models has been focused on linking the parameters of the model to physically based measurable quantities and the development of regional parameter sets through the use of homogeneous regions. The purpose of these techniques either explicitly or implicitly is to identify areas that exhibit similar hydrological properties, ensuring that a calibrated model that works for a specific site can be applied to other sites in the area. Before considering the regionalisation of the rainfall model presented by Heneker *et al.* (2001) it is appropriate to investigate various techniques and evaluate their applicability to the problem of event based rainfall modelling directly.

2.5.1 Identification of Homogeneous Groups (or Clusters)

Traditionally the first step in many regionalisation processes has been the identification of homogeneous regions. If a number of sites (rainfall or stream flow etc) can be placed together into a group that exhibit similar hydrological processes or statistics, then they can be defined as belonging to a homogeneous group or region. After the definition of the group boundaries, model parameters can be developed from the combination of data sites inside the homogeneous group. These parameters can then be applied or moved to sites of interest within these boundaries that contain little or no historical data. The development of homogeneous regions and groups of sites has received substantial literature coverage with particular reference for use in regional flood frequency analysis.

Tasker (1982) noted that many investigators initially identified regions and sub-regions subjectively based on the residuals from a regression analysis. The residuals were used as a guide for drawing homogeneous regional boundaries on a map of the area (see Wandle (1977), Guetzkow (1977)). A more objective method of defining these homogeneous regions of similar hydrologic or basin characteristics was presented by DeCoursey (1973) and later by DeCoursey and Deal (1974). They applied a technique known as cluster analysis (Cooley and Lohnes (1971)). Cluster analysis is the organisation of observed data records to identify groups or clusters of sites that are similar *within* clusters and dissimilar *between* clusters. Typically when employing cluster analysis, the criterion used to determine similarity between sites and the classification of a new site into an existing cluster must be defined.

The early work of DeCoursey (1973) and DeCoursey and Deal (1974) introduced the use of the simple Euclidian distance as a measure of similarity between sites or clusters of sites. The Euclidian distance is given by

$$d_{jk} = \left[\sum_{i=1}^{p} (x_{ij} - x_{jk})^2 \right]^{\frac{1}{2}} \qquad (2.1)$$

where d_{jk} is the 'distance' between stations j and k, x_{ij} is the i^{th} hydrological or basin characteristic at station j and p is the total number of characteristics being considered for the cluster analysis.

In order to use cluster analysis, an initial grouping of the sites into groups must be undertaken. DeCoursey and Deal (1974) arbitrarily divided their (N) sites into two groups and performed a discriminant analysis to determine whether any sites do not belong to their current group. Sites that are shown to be in an incorrect group are switched and the iterative analysis continued. Additional clusters can be incorporated by dividing existing clusters into two and so on. An alternative technique referred to as the complete leakage algorithm (Soakl and Sneath 1963), the farthest neighbour (Lance and Williams 1967) or the maximum method (Johnson 1967) has also been applied as an alternative for the first step in the cluster analysis process. All these methods are based on initially nominating each site as a cluster of

one site. Distances between this site and every other site can then be calculated using the Euclidian distance formula. At each step, the two closest sites are combined to form a single cluster. Once clusters contain more than one site, (i.e. there is more than one value of (x_{ij})as there are two or more sites in cluster j) then the largest distance between all sites in the cluster and the object site k can be adopted as the corresponding distance d_{jk} between the cluster j and the site (or cluster) k. This was similar to Tasker (1982) who was able to show that the adaption of the earlier cluster analysis of DeCoursey and Deal (1974) has the ability to define sites into homogeneous groups.

Cluster analysis was also investigated by Mosley (1981) when attempting to identify regions of catchments in New Zealand which have similar hydrological regimes. Mosley (1981) was able to show that when there are a number of factors that have an equal influence on the hydrological regime of a catchment; homogeneous regions are difficult to define. It was also noted that cluster analysis should not be used independent of subjective decisions for the purposes of defining these regions, but that it is a useful tool for interpreting the available data sets. Once cluster analysis has been performed, then local knowledge of the climate, topography etc should be incorporated to describe why certain sites are similar in their catchment response. Unlike some of the earlier work, sites in a given cluster were required to be spatially continuous. This enabled homogeneous zones to be located on a map of New Zealand. It is not a requirement of cluster analysis that sites be geographically contiguous in a given grouping. Indeed for studies that focus on catchment responses such as stream-flow or floods, it is possible that sites considerable distances apart geographically are more similar in terms of their hydrological response than those nearby. However as noted by Burn *et al.* (1997), having regions that are largely geographically contiguous provide an advantage when ungauged catchments must be assigned to a specific region.

Burn and Boorman (1993) continued to use the idea of cluster analysis when grouping hydrologically similar catchments in their study of 99 catchments in the UK. They used what is referred to as the K-means clustering algorithm to minimise their objective function, an extension of the simple Euclidian distance equation given by

$$F = \sum_{k=1}^{K} \sum_{i \in I_k} \sum_{m=1}^{M} W_m \left(X_m^i - C_m^k \right)^2 \qquad (2.2)$$

where (W_m) is the weight applied to feature (characteristic) m in the Euclidian distance measure, (X^i_m) is the value of feature m for site i, (C^k_m) is the centroid coordinate for feature m of cluster k, (K) is the total number of clusters, (I_k) is the set of objects in cluster k and (M) is the total number of features. In this fashion, sites that have similar characteristics to their counterparts in the cluster will provide a low score and hence a better fit to the current grouping set-up of the sites.

Hosking and Wallis (1993) incorporated a different approach using L moments (Hosking (1986, 1990)) to provide an objective test as to whether certain groupings of sites actually belong to a homogeneous region. Theoretically in a homogeneous region all sites should have the same L-moments, however owing to sampling variability this will not necessarily be the case. Making use of this fact L-moments for all sites in a proposed homogeneous region were calculated. A heterogeneity measure is used to indicate whether the observed variability in these statistics was simply due to sampling variability or due to the incorrect classification of sites into the proposed homogeneous region. Sites could then be re-classified and the process repeated to improve results.

Burn *et al.* (1997) provided an "agglomerative hierarchical" clustering technique combining previous techniques to determine homogeneous regions. First a region forming process is adopted which determines an initial set of clusters. This process follows the previously introduced steps of cluster analysis. At each step the dissimilarity or distance measure is calculated between each pair of objects (an object is still either a single site or a grouping of sites) and the union of the two closest objects forms a new cluster. The distance measure is now recalculated from the new cluster to all other objects and again the two closest objects form a new cluster. This process is repeated until the required number of clusters has been formed. Burn *et al.* (1997) then use the heterogeneity measure of Hosking and Wallis (1993) to determine whether the sites now placed within a cluster are sufficiently hydrologically homogeneous. For a region that fails this test, further subdivision of the region is

undertaken through the re-application of the original cluster analysis. When all regions are either sufficiently homogeneous or contain too few sites to be further subdivided, the process stops. As part of the cluster analysis, a geographical distance measure was incorporated ensuring that largely geographically contiguous regions were identified. This provides an advantage for classifying ungauged catchments as the location of the site on a map determines which cluster the site belongs to. The technique was successfully applied for a set of 217 catchments in West-Central Canada.

DeGaetano (1998) developed a clustering algorithm for specific application to extreme rainfall data. The basis of this work was that the largest rainfall events at all stations within a subregion or cluster could be represented by the same theoretical extreme value distribution. Smirnov tests calculated for each iteration were used as a measure of whether the distributions of sites in a proposed cluster were similar. Like Burn *et al.* (1997) a measure of the geographic proximity between sites was also incorporated. The application of cluster analysis on the distribution of the climate variable, rather than the variable average is noticeably different to previous techniques and provides a significant advantage. Not only are stations grouped on the location (mean) of a particular variable, but also on the other parameters that define the distribution. This becomes important if the regional assumptions are to be used for extrapolation of values outside those contained within the original historical records. The presented results of DeGaetano (1998) compared favourably when tested with the heterogeneity measure of Hosking and Wallis (1993), however the author concedes that cluster analysis by its nature contains significant bias. These are a result of the type of clustering selected, the inclusion of redundant or irrelevant data and by the variable nature of climate data.

In addition to the negative discussion of DeGaetano (1998), Acreman and Wiltshire (1989) challenged the idea and use of homogeneous regions developed through cluster analysis or similar techniques. They argued that the estimated regional relationship (distribution or curve) would only be valid for a site situated at the centroid of the group of sites used in the analysis. It was also argued that certain sites could belong to a given region more than others and that some sites may influence the regional relationships more, introducing additional bias. To circumvent

this problem, they suggest that instead of defining each site as part of a group, each site should be allocated its *own* group consisting of sites that have similar hydrological properties. These sites could then be used to estimate the required characteristics at the site of interest.

Burn (1990) further developed the ideas of Acreman and Wiltshire (1989) and referred to it as the region of influence approach. The region of influence approach has the advantage of eliminating the distinct boundaries developed by defining homogeneous regions. In order to define a target sites region of influence, a distance measure was calculated from the target site to every other site. A surrounding site that recorded a distance measure less than a nominated threshold value or cut off point is then included in the region of influence for the target site. These sites are then combined to form a regional flood frequency curve, only applicable at that specific target site. A modified Euclidian distance measure is used to determine the distance from the target site to the other sites and is defined as

$$D_{ij} = \left[\sum_{m=1}^{P} W_m \left(x_m^i - x_m^j \right)^2 \right]^{\frac{1}{2}} \qquad (2.3)$$

Where

D_{ij} is the weighted distance between site i and station j, W_m is the weight applied to the attribute m to reflect its relative importance, P denotes the number of attributes and x_m^i is the standardised values of the measure of attribute m for site I. The standardisation is applied to remove the problem of units and is calculated by subtracting the sample mean from the value and dividing by the sample standard deviation.

Ribeiro-Correa *et al.* (1995) investigated using canonical correlation analysis to identify regions of influence for drainage basins instead of a threshold style procedure. They were able to identify regions for 55 catchments in Canada. The region of influence approach has also been applied by El-Jabi *et al.* (1998) for the regionalisation of 100 year floods in Canada with some success.

Nathan and McMahon (1990) continued to raise and discuss the apparent problems associated with the identification of regions using techniques based on cluster analysis. Of major concern was the selection of important characteristics and their corresponding weighting which are used to determine similarity between catchments and/or sites. This can often be a compromise between available data, the judgement of their importance by the modeller and model computation time constraints. The measure used to assess similarity is also highly dependent on the scale of data used. A clustering algorithm working with rainfall in millimetres for instance would obtain a different set of similarities to those working with rainfall in metres. As noted by Nathan and McMahon this can be removed by scaling the variables so that they all have a mean of zero and a unit variance; however information on the variability of individual measures is lost.

An additional problem associated with these techniques is the abundance of different algorithms and distance measures available. It is evident that these different measures and techniques can produce different groupings of sites based on the same data. Even if the use of several techniques provides a set of similar clusters, there will still remain sites that are not continually allocated to the same cluster set and thus must be classified on an arbitrary basis. This fact was noted by DeGaetano (1998) who suggested that numerous cluster analyses with different techniques should be undertaken and then the best groupings of sites can be chosen when comparing solutions with physical and climatological considerations.

The region of influence approach appears to circumvent numerous identified problems with the use of traditional cluster analysis. Cluster analysis relies on including variables in the analysis which have a significant impact on the required output. This selection is dependent on the expectations of the modeller and introduces a bias into the region identification process. The inconsistent results produced when using different clustering algorithms or data recorded at different time scales also introduces doubt on the validity of cluster analysis as an accurate method of determining homogeneous regions, if they exist. The region of influence approach removes any boundaries and associated problems with sites that are located on or near these boundaries. However the problem of selecting catchment attributes and the associated bias is still evident. This approach also requires the determination

of a threshold value, which alters the number of sites that have an impact on the calculation of data at the target site.

The application of homogeneous regions through cluster analysis and the region of influence approach has been widely adopted for regional flood techniques; however their usefulness for rainfall modelling is questionable. The selection of influential variables for use in a rainfall cluster analysis is problematic at best due to the influence of unmeasurable quantities such as local and global atmospheric fluctuations. While it is conceivable to group sites for the purpose of regional flood analysis by using influential variables such as catchment conditions, observed rainfall, elevation, slope etc, it is difficult to measure and distinguish similar indicators for rainfall. The atmospheric processes and fluctuations that influence rainfall are difficult to determine directly. However, it may be reasonable to assume that a regional shape distribution of rainfall exists, but the determination and use of this distribution is the primary concern. The idea that sites within a region may have a consistent distribution was the motivating idea behind developing the new methods presented in this study. Given the idea of a consistent regional distribution for rainfall, a technique to apply this idea and successfully regionalise the rainfall model is required.

2.5.2 Regional Flood Analysis

A significant amount of literature on the regionalisation of hydrological models has been focused on flood frequency analysis and the prediction of stream flows. Generally regional flood frequency modelling can be classified into two groups. The Index Flood method introduced by Dalrymple (1960) and the Multiple Regression method attributed to Benson (1962). While these procedures are used to provide regional predictions of extreme values rather than regional continuous simulations as is required for this study, an investigation into their use provides further understanding of the general procedures available and possible extensions to rainfall modelling.

Boyd (1978) provided a good explanation of regional flood frequency techniques employed in Australia and in particular the use of multiple regression as a

regionalisation tool when investigating 79 catchments in New South Wales. In applying multiple regression on flood peak discharges to various catchment characteristics, Boyd noted that the inclusion of a large number of physical and hydrological variables in the regression procedure ensures the resulting regression equations are applicable to a very wide range of catchments. This in turn ensures that a high degree of regional homogeneity is not required removing the need for extensive work in defining homogeneous regions using cluster analysis or a similar procedure. Regional regression models have also been used to develop relationships between catchment physical quantities and low-flow statistics [Thomas and Benson (1970), Thomas and Cervione (1970), Vogel and Kroll (1992) and Kroll and Stedinger (1998, 1999)] and for flood flows [Matalas and Gilroy (1968), Tasker *et al.* (1996)]. Given the number of options available, Valdés *et al.* (1979) provided a technique for choosing between various alternative streamflow regression models in the literature. The use of regression models which link physical descriptors to hydrological models can be an advantage over alternative regionalisation techniques as the model can be applied instantly to any site where the physical descriptor can be determined or measured. However using this physical descriptor to model link for rainfall modelling is problematic. Rainfall is a complex process which is dependent on local and global atmospheric physical properties. These properties are not easy to measure directly and have resulted in the development of statistical based rainfall models in comparison to physical process models for rainfall. The ability to link a rainfall model to physical atmospheric condition measurements and successfully predict rainfall volumes or events is questionable at best.

Boughton (1984) presented an alternate approach to regionalisation with a simplified water balance model that could be applied to ungauged catchments. Instead of applying a regression procedure, physical catchments were classified into one of three types depending on simple field observations. The parameters of his water balance model were fixed dependent on the given classification and the estimation of parameters for ungauged catchments was simply based on this site classification. Given an observed daily record, the resulting water yield could be estimated for the ungauged catchment. While this simplified model provided good results for the estimation of stream flow, it was not rigorously tested under different catchment and climate regimes and it is unlikely that a similar classification strategy with fixed

regional parameters could be undertaken for rainfall modelling. Again the model would rely on a physical input record at the site of interest to assist in modelling the site specific behaviour. (In the case of his stream flow model, the physical record is the observed daily rainfall) Given the complex nature of rainfall, it is unlikely that a similar single physical descriptor could be found to adequately describe the changes in observed rainfall from one site to the next nor is it likely that sites could be grouped together and have identical model parameters.

Sefton and Howarth (1998) established relationships between physical catchment descriptors and the dynamic response characteristics obtained from the output of a rainfall-runoff model. Given the driving variables of rainfall and temperature, and the measurement of the descriptors (topography, soil type, climate and land cover) flow can be simulated for any catchment in the region. They applied their multiple regression model to two test cases in England and Wales and were able to satisfactorily reproduce the daily flows. An advantage of the model was that no attempt was made to group or define homogeneous regions. However measuring physical descriptors that influence rainfall as opposed to stream flow is significantly more difficult and provides a limit to the application of simplified techniques such as this and the previously discussed Boughton (1984) model.

Seibert (1999) also analysed a region containing 11 catchments by establishing functional links to catchment characteristics. A two-parameter regression function was used (one of either a linear, exponential, power or log function). Unlike previous works, Seibert concluded that while the results appeared acceptable, the use of this technique itself is questionable. The uncertainty in the relationship between model parameters was a major concern. The regression of parameters requires a certainty about the model values. If this knowledge is not certain, the value of regressing these against the physical characteristics may be in question. Seibert found that for a certain catchment, a number of various parameter values might provide the same goodness of fit. It is therefore difficult to relate this range of parameters to physical properties to provide a regional estimate. The inter-relationship between physical quantities was also discussed. If it could be shown that two physical based quantities were correlated, then this also could induce problems into the regression routine.

A final problem associated with models that link physical properties to model parameters arises when the required site of interest lies some distance away from the calibrated sites. Seibert (1999) recognised this fact and stated
"Obviously, using larger regions increases the number of gauged catchments. However at the same time the variation of climate and physiography between the catchments increases as the sampling area increases, i.e., more variables have to be included into the regression analysis".
From these results, it appears that regression may be a useful tool for localised regionalisation; however when the problem is over a larger scale for instance the process of rainfall, an alternate method must be found.

Index flood procedures use the underlying assumption that the sites in a given homogeneous region share an identical frequency distribution apart from a site-specific scaling factor. This scaling factor is called the index-flood. The term "index flood" is a remnant of the early work of Dalrymple (1960) who applied this procedure to flood data, however theoretically it can be applied to any data set.

Given data at (N) sites, with site (i) having sample size (n_i) and observed data at these sites (Q_{ij}, $j=1,\ldots,n_i$). Let the quantile function of the frequency distribution be ($Q_i(F)$, $0<F<1$). Given the assumption of the existence of an identical regional frequency distribution with localised scaling from an index flood then we can write

$$Q_i(F) = \mu_i q(F), \qquad i = 1, \ldots, n \tag{2.4}$$

(μ_i) is usually taken as the mean at site frequency distribution but any location parameter of the distribution can be used instead. ($q(F)$) is the regional growth curve, a dimensionless quantile function common to every site in the homogeneous region.

A number of researchers have used the annual maximum series in conjunction with the index flood procedure for modelling extreme hydrological events (see Hosking *et al.* (1985), Lettenmaier *et al.* (1987), Hosking and Wallis (1988)). Birikundavyi and Reusselle (1997) use the index flood technique with the Partial Duration series while Madsen *et al.* (1997) suggests that using the Partial Duration series with the index

flood technique provides a more efficient technique (than using the annual maximum series) for estimating regional floods. The index flood method has had significant success for the purposes of stream flow regionalisation.

Jin and Stedinger (1989) developed a maximum likelihood technique which took into account both the regional information and the local at-site historical information with a regional index flood distribution. They concluded that regional flood analysis can be improved by using both a good regional model and good historical data provided it is used carefully.

Regional frequency techniques have previously been interested in estimating extreme values or flow quantiles at ungauged catchments. It is conceivable that methods such as the index flood could be applied to estimate extreme rainfall values however it would still not be able to directly calibrate a rainfall model in order to provide a continuous rainfall record. Physical regression of rainfall model parameters to catchment characteristics could provide a regionalised rainfall model, however the choice of appropriate physical descriptors is difficult and problems associated with leaving out influential descriptors is evident. Given the physical variables that influence rainfall are not only local (and potentially) measurable quantities such as elevation, distance to coast, surrounding hills etc but also atmospheric conditions such as global circulation, cloud development, atmospheric wind patterns etc, the availability of applicable data and as a consequence the usefulness of models which require this information is a major issue. Rainfall patterns are influenced by atmospheric fluctuations and the complex interaction between various meteorological parameters and at this time the data required to undertake a successful regression analysis to these physical descriptors is not available.

2.5.3 Rainfall Model Regionalisation

The application of regionalisation procedures to rainfall models specifically has had limited coverage in the literature. Typically models that rely on daily data for calibration do not require the use of regionalisation techniques as there is an abundance of daily rainfall records available for use. However, models that attempt to reproduce extreme values or are calibrated to pluviograph data need to

incorporate regionalisation techniques in one form or another to ensure wide spread application.

Cao (1974) provided an early technique for estimating short duration-depth-frequency curves at sites which only record daily data. Sites which record short-time increment rainfall in the area were grouped into homogeneous regions and a common regional depth-duration-frequency curve estimated. A target daily site is then allocated to one of these groups by a comparison of the one-day depth distribution between the site and the groups of sites. Once allocated to a group, the target site adopts the regional frequency curve, providing an estimate of the required duration-depth-frequency curve. While only presenting results for Sardinia, a good estimation of these curves was achieved.

Cong *et al.* (1993) extended the earlier frequency curve similarity assumption of Cao (1974) by assuming that the distribution form of rainfall at all stations in a study area are the same. By applying various mathematical and statistical tests, they were able to ascertain a) the form of the regional distribution and b) the probability that this determination was in fact true. This result could then be used to assume a distribution for any site within the region. They applied their technique to annual daily maxima in Pennsylvania and West Virginia. While this technique can be applied to develop a regional distribution curve, the authors note that it cannot be applied to determine the distribution at any individual site.

Hay *et al.* (1991) developed a regionalisation method by introducing a 'weather state' model in which measurements of synoptic atmospheric information are used to classify each day into one of a small number of states. The weather state effectively acts as an automatic classifier of atmospheric conditions. Historical data is then used to fit a model relating these weather states to the observed daily rainfall. As the model is fitted to atmospheric information, it can be applied in any region so long as the climatic driving force of precipitation does not change and the controlling synoptic variables are incorporated into the fitting procedure.

Hughes and Guttorp (1994) also applied a similar weather state technique, by relating site precipitation to synoptic atmospheric patterns and in particular to sea level

pressures. A major assumption in models of this type is that the relationship between the climatic state and the precipitation does not alter over time due to climate change and other influences that have not been included directly in the model. While this type of model provides a solution to the regionalisation issue, it requires the selection of appropriate atmospheric variables which introduces uncertainty and relies on the understanding and assumptions from the modeller. Models of this type are also reliant on the availability of good quality climate data at the time step of interest which may also be an issue. This method does have the potential to develop into a useful tool for predicting effects due to climate change (Hughes *et al.* (1993)) if it can be incorporated into atmospheric forecasting models however its applicability to fine time scale rainfall modelling is questionable. Typically in the literature these models were focused on reproducing the occurrence process of daily rainfall and as a result no simulation of rainfall amounts was presented.

Arnbjerg-Nielsen *et al.* (1996) again dealt with extreme rainfall and studied the regional variation of extreme values of peak intensity in Denmark. Interested in identifying variables that could describe the regional fluctuations; they analysed the correlation structure between various extreme values and possible covariates. They found that the annual average precipitation and the 0.2yr return period for depth per day at a nearby gauge with daily resolution could be used to describe the regional variations in the maximum 10-minute intensity of a rain-event and the total depth of rain events at certain return periods. While this was seen as a success, these regional variables could not explain the variation in longer return periods. Their analysis was also hindered by the availability of only 15 years of data.

Cowpertwait *et al.* (1996) developed a methodology to enable the application of the Neyman-Scott (NS) model to ungauged catchments. Similar to the later rainfall runoff work by Sefton and Howarth (1998) and Seibert (1999), the model was calibrated to 112 sites scattered throughout the United Kingdom. A set of explanatory variables was then developed which was used to describe and relate the changing physical process of rainfall at each of these sites. By using linear regression on the model parameter estimates against these explanatory variables, a relationship was developed that could be applied to any site for which the explanatory variables

could be calculated. The general regression model for each of the NS model parameters can be written as

$$\ln(\psi_i) = a_0 + a_1 X_{i1} + a_2 X_{i2} + a_3 X_{i3} + ... + a_n X_{in} + \varepsilon_i \qquad (2.5)$$

Where ψ_i is the NS model parameter for ith station-month, X_{ij} is the jth explanatory variable for the ith station-month, a_j is the least-squares regression parameter for the jth explanatory variable, ε_i is the residual error in the regression model for the ith station month and n relates to the number of explanatory variables that were found to contribute significantly to the prediction of the NS parameter.

For the purposes of their work, the explanatory variables selected were Altitude (A), North Ordinance Survey (OS), Grid Reference (N), West-East effect (W) and the distance from the Coast (C). The West-East effect was a result of the well-known 'rain-shadow' that the Pennine mountain range causes on rainfall in the UK. Depending on the location of the site in relation to this range, the rainfall statistics vary significantly. Wigley *et al.* (1984) proposed the use of an east-west dividing line to delineate areas that were affected by this range. Cowpertwait *et al.* (1996) adopted this line as the criteria for defining the east-west effect. As the values (A,OS,N,W and C) are all measurable at any point in the UK, Cowpertwait *et al.* (1996) had provided a regionalised Neyman-Scott rainfall model.

While the regionalisation method of Cowpertwait *et al.* (1996) was claimed to be a success, the authors cautioned that errors on selected individual sites are due to the microclimate effects that produced precipitation variation from site to site. To remove this error, other explanatory variables would need to be introduced to develop a relationship that enables the microclimate effects to be included. This illustrates the difficulty faced when relating stochastic rainfall model parameters to physical quantities. The assumptions made by including some physical descriptors and not others due to modeller input or a lack of available data induce errors in the final model.

Jones and Thornton (1999) introduced an alternate method linking a third-order Markov rainfall model to climate surfaces. Climate surfaces are usually developed for

monthly rainfall and minimum and maximum temperature. These climate surfaces are calculated by dividing the study region into pixels of a nominated size. Jones and Thornton (1999) used a simple interpolation algorithm based on the inverse square of the distance between the station and the interpolated point to calculate the surface values for each pixel. This is given by

$$x_{pixel} = \sum_{i=1}^{5} d_i^{-2} \times \sum_{i=1}^{5} \frac{x_i}{d_i^{-2}} \tag{2.6}$$

Equation (2.6) ensures that the climate surface passes exactly through each station point an advantage over alternate techniques of interpolation. As the weight distance $\left(\frac{1}{d_i^2}\right)$ approaches infinity as (d) approaches zero, the value of a pixel containing a station observing climate data tends towards the value of that station. These climate surfaces were then used to calibrate the rainfall model. Initially daily records from across the world were grouped into similar climatic clusters. The rainfall model can then be calibrated at any pixel point by considering how the pixel climate surface adjusts the parameter values within each cluster relative to the cluster climate means. Presented results indicated the model performed well over long periods of time (crop growing seasons) but discrepancies were more common over shorter periods. One site showed consistent deviations from the observed and it was concluded the poor results were a factor of its complex climate, the size of the pixels not reproducing small variations in climatic variables and the groupings during the cluster analysis. This technique requires the calculation of climate surfaces for any required value at the time period of interest. A model calibrated to monthly data therefore requires 12 monthly rainfall surfaces. For a model that uses a finer time scale, the computational burden increases.

Gyasi-Agyei (1999) extended the earlier jitter rainfall model presented by Gyasi-Agyei and Willgoose (1997, 1999) and identified regional model parameters from daily rainfall statistics. A total of 13 sites in central Queensland were used for the study. After simplifying the parameters of the model by removing correlated parameters and those that remained constant over the year, the model was calibrated to one (Rockhampton) of the sites in the region. The data from all 13 sites was then

combined and a regional parameter set identified. Gyasi-Agyei (1999) concluded that while there were differences between the parameter sets, a single set of regional parameters could be used for all sites. These regional parameters and the observed daily data at the target site of interest are then used to estimate model parameters at the required simulation time scale used in the generation of short-time increment rainfall. This technique was shown to produce hourly rainfall but was not adequate for generating a synthetic 6-minute record. Significant variations in one of the assumed constant parameters at this time scale led to the development of an alternate technique for 6-minute generation. The jitter model was employed to simulate at the hourly level and the resulting hourly rainfalls disaggregated into 6-minute periods. Gyasi-Agyei (1999) assumed that the distribution of fractional rainfall proportions in wet 6-minute bins was uniform. Therefore given 5 wet bins in an hour of 50mm of rain, a uniform distribution can be used to generate fractional weights. Each weight is divided by the sum of the five values and these are then multiplied by the hour rain depth to obtain 6-minute rainfall values. It was not shown how the number of wet bins in an hour was calculated. Presented results showed good agreement between the observed and simulated dry probabilities, mean, variance and lag-1 autocovariance at selected months for certain levels of aggregation. Extreme rainfall results were not presented and the author proposed further research to improve the reproduction of second order moments.

Wotling *et al.* (2000) provided another physical link model to regionalise the extreme precipitation distribution in Tahiti. A limited set of variables that described the topographical environment were linked to the parameters of the rainfall intensity distribution through the use of a stepwise regression with 20 rainfall sites. These regression estimates were then applied to 300 fixed points on a grid and the resulting intensity estimates interpolated to provide an approximation of extreme rainfalls over the entire island. Wotling *et al.* (2000) concede that the linking of model or distribution parameters to topographical descriptors, while working well in Tahiti is restricted in its application to mountainous areas where the relationship between rainfall and the topographical descriptors is very strong.

Smithers and Shulze (2001) developed a regional estimation of short duration design storms using L-moments. Using 172 rainfall stations in South Africa, 15 relatively

homogeneous clusters were developed. The index storm method was used and defined as the mean of the annual maximum series. A relationship between this index storm and the mean annual rainfall was derived enabling short duration design storms to be estimated at ungauged sites in South Africa. Except for two clusters, results indicated the ability of this technique to estimate the 24 hour design storm. Results presented for smaller durations showed a progressively worse comparison to observed values.

2.5.4 Summary of Rainfall Model Regionalisation

The regionalisation of hydrological models is often a required element in the application of these models for engineering analysis. Historical records are often short and can only provide a small amount of direct information about the site in question. This leads to models that are poorly calibrated and in turn provide questionable results. This is particularly relevant when utilising short time increment (6 minute) rainfall data as an input into hydrological studies.

While significant work has been presented on regionalisation methods, work in this area has primarily been focused on stream-flow and flood frequency rather than rainfall specifically. However the underlying theory used for regionalisation of catchments can generally be applied to regionalisation of rainfall models and may even be applied with more confidence as noted by Cong *et al.* (1993) who wrote

"... the assumption that the distribution form (not the distribution itself which involves parameters) of the rainfall is the same for all stations in a homogeneous region is more reasonable than a similar assumption for floods, because the effect of the ground surface condition on rainfall is much less than floods."

Numerous techniques presently exist that enable the development of a relationship between model parameters and measurable physical quantities and therefore provide a regionalised hydrologic model. While these techniques are different, the same inherent weakness is present particularly when looking specifically at regionalising rainfall models. These types of models require the measurement of physical properties that affect observed rainfall values. As rainfall is linked to atmospheric

dynamics, annual and inter-annual global climate variations and smaller area fluctuations, the ability to include enough physically measurable quantities in the development of these relationships must be questioned. The inclusion of certain physical parameters over others is also generically flawed and is an individual bias introduced by the expectations of the modeller.

The application of homogeneous regions for the purposes of developing contiguous areas that exhibit identical hydrological properties has merit and has been successfully applied in numerous studies for the purposes of regional flood frequency. If a sufficient network of long term historical rainfall records were available to enable such an analysis, a parameter contour map could be produced across the country. This would enable the simulation of rainfall at any point of interest at the required time scale. However given the complex atmospheric interaction in the development and production of rainfall, the lack of significant data records and the cost of acquiring available data sets, this method is not viable both economically and computationally. Alternative methods such as compiling climate surfaces at the daily or finer time scale also require extensive resources prohibiting their use directly.

The shortcomings of existing regionalisation techniques and the lack of quality long historical data sets for calibration have motivated the development of an alternative approach to simulating short time increment rainfall at ungauged sites. This new approach adapts the underlying theories of regional flood frequency and in particular the index flood method and applies it to the parameter distributions of a high resolution point rainfall model. Initially a master set of calibrated parameters are determined based on a nearby pluviograph record and once this master parameter set has been calibrated, any site specific short pluviograph or daily data records are then used to adjust the master parameter set and provide model parameters at the target site of interest. This enables the existing model to improve its application to short historical pluviograph records and use the abundance of daily data records around Australia as the basis for a new regionalisation approach, providing a rainfall simulation tool capable of simulating accurate high resolution rainfall data across the country.

In order to develop and apply a new regionalisation technique, existing models were investigated with the view of selecting an appropriate model for further development. For the purpose of this study, the selected model was the event based model presented by Heneker *et al.* (2001). The Heneker *et al.* (2001) rainfall model is an alternating renewal model and calibrated to independent storm events. The calibration of the model to independent storm events ensures the model can be calibrated to records with missing periods of data. This was seen as an important attribute when considering models that are calibrated to short time increment rainfall records which often contain large periods of missing or erroneous data. In addition to developing a new regionalisation technique, this study also presents significant improvements to the model in terms of parameter calibration, identification and the inclusion of parameter uncertainty. These ideas were lacking in the original model and are a significant improvement to the model when comparing results to observed statistics. The key ideas behind the Heneker et al (2001) model and the improvements are presented in the following chapters.

CHAPTER 3

IMPROVEMENTS TO THE HIGH RESOLUTION POINT RAINFALL MODEL

3.1 Introduction

The existing rainfall model presented by Heneker *et al.* (2001) and selected for further development as part of this study belongs to the group of wet-dry alternating renewal models. Models of this type are characterised by their calibration to independent storm events. The initial development of the model, while adequate for its desired purpose, contained a number of deficiencies in its original form which warranted attention prior to the development of a regionalisation approach for applying the model at sites with little or no calibration data. It was important that the model produced favourable simulation comparisons to observed statistics, was robust during calibration and required minimal user input. In order to verify its performance, the original model was comprehensively investigated and then modified ensuring the improved final model is easy to use, accurate and robust.

3.2 Description of Original Rainfall Model

The original point rainfall model developed by Heneker *et al.* (2001) is an event based model where the event series is completely defined by the probability distributions of inter-event time, storm duration and conditioned storm intensity. In its original form given an adequate record available for calibration, the model is capable of reproducing various rainfall statistics including certain values that were not used during the calibration process. This was an important consideration when selecting

this model for further development and the ability to replicate non-calibrated statistics is an indication that the model is theoretically well structured. The structure of the model and the results presented by Heneker *et al.* (2001) motivated improvements and the new regionalisation approach proposed in this thesis.

3.2.1 Model Structure

The Heneker *et al.* (2001) model is based on the alternating renewal process introduced by Green (1964). As discussed in section 2.4.4 the simulated time series is completely defined by a sequence of wet events (storms) interspersed with dry events. For the Heneker *et al.* (2001) model this structure was characterised by three main variables, the dry periods or inter-event times t_a, the wet periods or storm durations t_d and the average intensity i and is shown schematically in Figure 3.1. Probability distributions were used to describe the observed populations of inter-event times and storm durations, while a third conditional probability distribution was used to describe the relationship between storm intensity and storm duration. This provided a simulation which was able to generate a rectangular rainfall pulse for each storm event dependent upon the duration of the event. Finally a temporal pattern model takes the simulated rectangular rainfall pulse and disaggregates the storm event down to the time step required (typically 6 minutes). In this fashion the model was able to provide a synthetic pluviograph record which compares favourably to various observed data statistics.

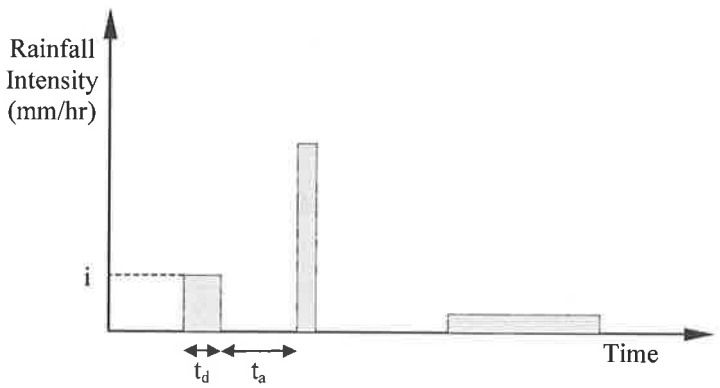

Figure 3.1: Schematic of the Heneker *et al.* (2001) model

3.2.2 Probability Model of Inter-Event Time & Storm Duration

In order to fit a probability distribution to the observed inter-event time t_a and storm duration t_d populations, a procedure was employed to extract independent events from the continuous historical record. After extensive analysis of correlation results, Heneker *et al.* (2001) adopted a minimum inter-event time of 2 hours to distinguish between independent storms. This value provides a balance between ensuring consecutive events are sufficiently independent and the need to have as much calibration storm data as possible with a fixed length historical record. While this minimum inter-event time differs between researchers, 2 hours was shown to be sufficient criteria for the definition of independence across numerous Australian sites.

Once the criteria defining storm independence has been set, the historical record can be examined. A historical wet storm event begins with any recorded observation of rain and continues until a dry period is observed that exceeds the minimum inter-event time. When a dry period greater than 2 hours is observed, the previous wet event is complete and a new dry event begins until the next observed period of rain. Using this definition of storm independence and the resulting extraction procedure ensures dry periods of no rain can be present during wet events, a phenomenon that is readily observed in real storms (see Figure 3.2). Upon completion of this storm processing step, probability distributions can be fitted to the resulting populations during the calibration process.

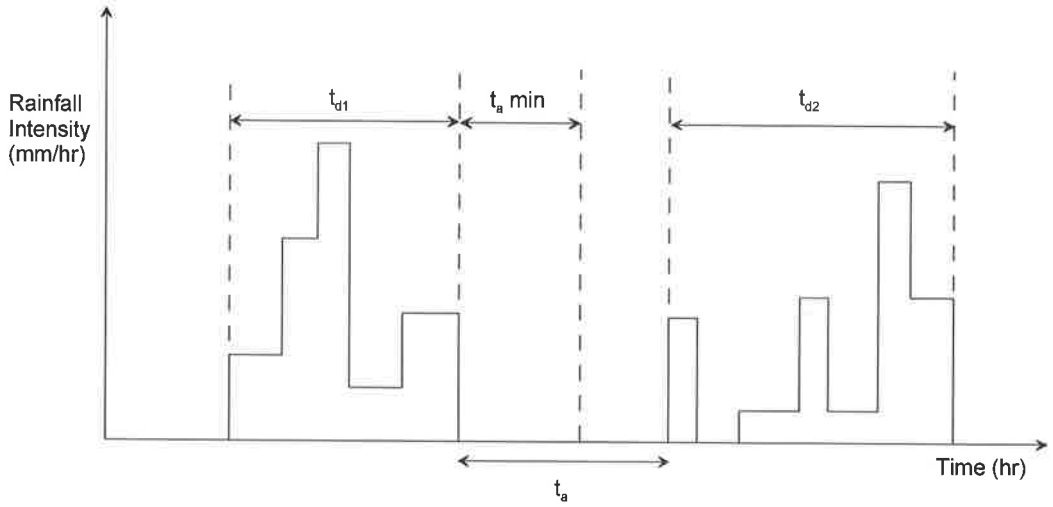

Figure 3.2: Schematic of Calibration Procedure

3.2.3 Calibration of Inter-Event Time & Storm Duration

In order to model the distributions of inter-event time and storm duration a combination probability kernel was used. Lambert & Kuczera (1998) introduced a generalised exponential distribution which is the basis for the calibration of inter-event time and storm duration. The generalised exponential distribution takes the form

$$F(x \mid \theta_t) = P(X \le x \mid \theta_t) = 1 - \exp[-g(x, \theta_t)] \qquad x > 0 \qquad (3.1)$$

where X is the independently distributed random variable, t relates to the time at the start of the storm or inter-event time. (If parameters are calibrated monthly, this relates to the month at the start of the event. Parameters can also be calibrated using harmonics across the year, in this case t is expressed as a fraction of the year that has passed.), θ_t is a model parameter vector dependent on t, and $g(x \mid \theta_t)$ is a kernel function. Re-expressing this equation as

$$\ln[1 - F(x \mid \theta_t)] = -g(x, \theta_t) \qquad x > 0 \qquad (3.2)$$

and using the appropriate kernel allows the distribution to be plotted on exponential probability paper. The kernel chosen by Heneker *et al.* (2001) to best fit the data was a combination of the Generalised Pareto Distribution (GPD) (Rosjberg *et al.* 1992) and the power law kernel shown respectively below.

$$g(x,\theta_t) = -\frac{1}{\theta_1}\ln\left(1-\theta_1\frac{x}{\theta_2}\right) \qquad \theta_1 < 0, \theta_2 > 0, \tag{3.3}$$

$$g(x,\theta_t) = \theta_3 x^{\theta_4} \qquad \theta_3 > 0, \theta_4 > 0, \tag{3.4}$$

The combination of equations 3.2, 3.3 and 3.4 gives the complete kernel used to describe the distributions of inter-event times and storm durations:

$$\ln[1-F(x \mid \theta_t)] = -g(x,\theta_t) = \frac{1}{\theta_1}\ln\left(1-\theta_1\frac{x}{\theta_2}\right) - \theta_3 x^{\theta_4}$$
$$\theta_1 < 0, \theta_2, \theta_3, \theta_4 > 0, \tag{3.5}$$

The parameter vector θ_t is calculated using maximum likelihood techniques.

3.2.4 Method of Maximum Likelihood

The method of maximum likelihood was chosen by Heneker *et al.* (2001) and is also used during this study for estimating the optimal model parameter set.

Given a set of observations X (x_1, x_2, ..., x_n), the method of maximum likelihood finds the parameters θ_t of a model that are most consistent with these observations. Consistency is measured by the probability of a model generating the observed value. The method is intuitively appealing as it tries to find the values of the parameters that would have most likely produced the observed data. If the observed samples have a density function $f(x;\theta_t)$, the likelihood of observing a particular value x_i can be assumed to be proportional to the value of the probability density function evaluated at x_i. Therefore the likelihood of observing a set of observations X is given by

$$L(\theta_t \mid x_1, x_2, \ldots, x_n) = f(x_1;\theta_t) f(x_2;\theta_t) \ldots f(x_n;\theta_t) \quad (3.6)$$

(Independence assumed)

The value of θ_t which maximises the likelihood function can be obtained either by differentiating the likelihood function with respect to θ_t, or by numerically maximising the likelihood. Typically the shuffled complex evolution method (SCE_UA) described by Duan *et al.* (1992) and the simplex search method introduced by Nelder and Mead (1965) are used to find the maximum likelihood parameters.

Since the natural logarithm function ln is strictly increasing, the maximum value of $L(\theta_t|x_1,x_2,\ldots,x_n)$, if it exists, will occur at the same point as the maximum value of $ln[L(\theta_t|x_1,x_2,\ldots,x_n)]$. This log likelihood function was the form used during this study. The numbers produced by the multiplicative nature of the likelihood function become too small for current computers to distinguish from zero and therefore the log likelihood function becomes easier to work with.

The parameters can be fitted as constants over the entire year, to individual months or allowed to vary smoothly over the year via harmonics in order to capture any seasonality in the data. Validation of parameter values can be undertaken prior to and after the simulation. Probability plots comparing observed and predicted storm events indicate the success of the search routine and the distribution assumptions. At the completion of the required simulation, plots comparing observed and simulated event distributions and aggregated statistics for various time periods can be produced to provide evidence of the quality of the simulation. Example calibration plots for selected months in Melbourne can be seen in Figure 3.3 and Figure 3.4. These plots clearly indicate that the inter-event times and storm durations are not exponentially distributed (exponential distributed variables would plot as a straight line) and that the combination kernel employed by Heneker *et al.* (2001) provide an excellent fit to the observed distributions.

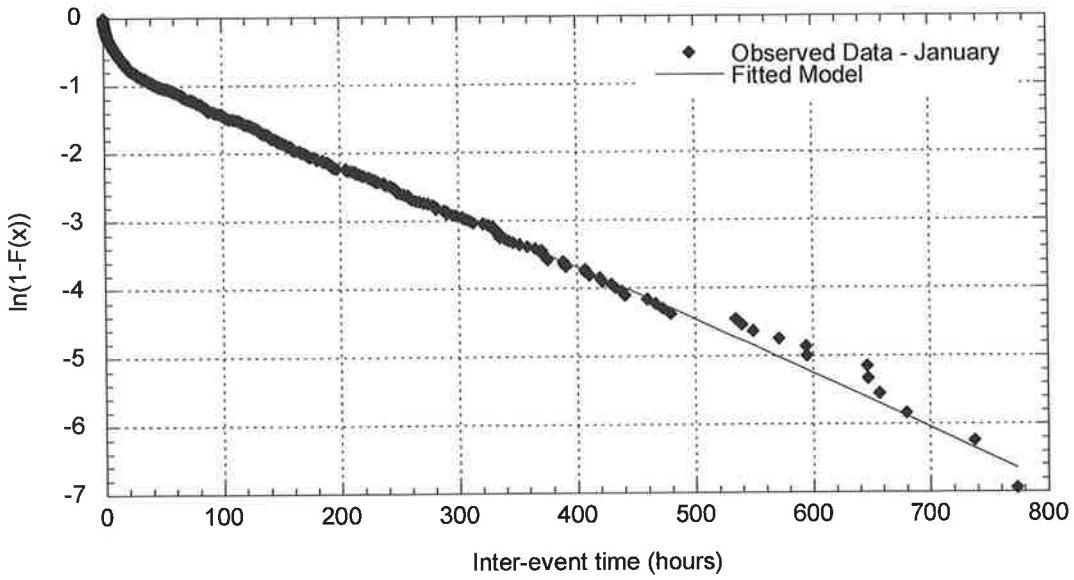

Figure 3.3: Heneker *et al.* (2001) model fitted to monthly inter-event time data for Melbourne in January

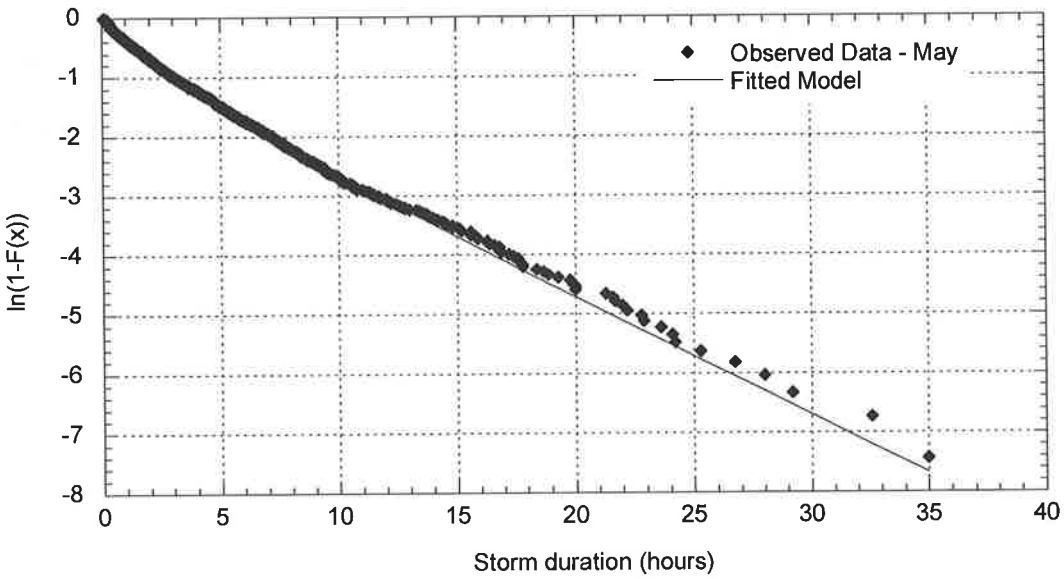

Figure 3.4: Heneker *et al.* (2001) model fitted to monthly storm duration data for Melbourne in May

3.2.5 Probability Model of Average Rainfall Intensity

The description of rainfall intensity is an integral part of the alternating renewal model. Scatter plots clearly indicate that the rainfall intensity is dependent on storm duration. Figure 3.5 displays such a scatter plot for Adelaide. It is evident from this plot that the calibration and simulation of rainfall intensity corresponding to each storm event is not a trivial task. Short rainfall events tend to have a slightly higher intensity pulse in comparison to longer duration events. This correlation between the intensity of a rain event and the corresponding storm duration requires the incorporation of a conditional relationship. In addition to this, a probability distribution must be chosen to describe the population of rainfall intensities.

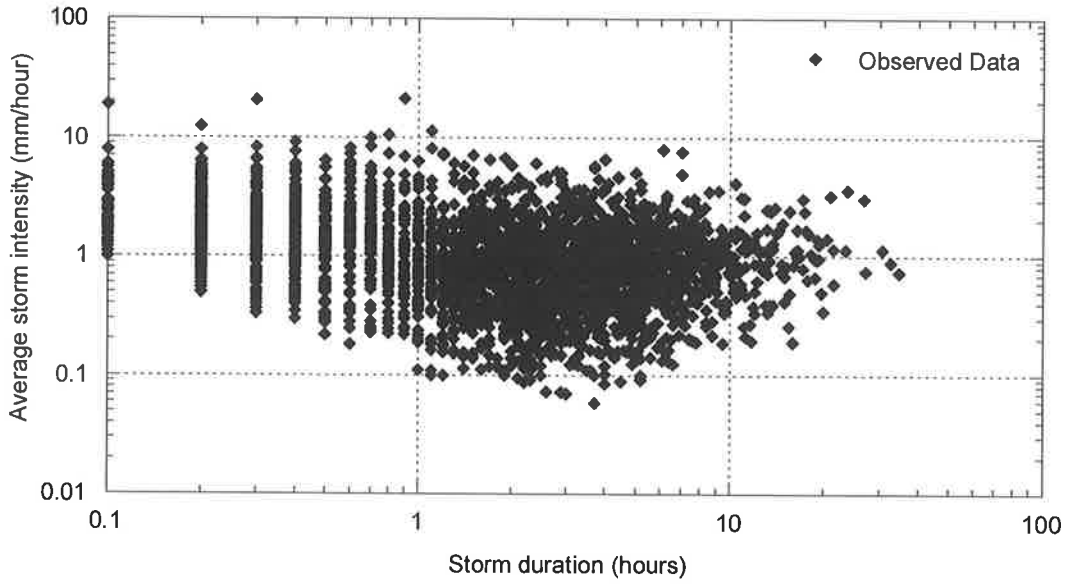

Figure 3.5: Average storm event intensity v duration for Adelaide

Heneker *et al.* (2001) used the GPD distribution to represent average rainfall intensity. The GPD has a number of advantages which provide significant benefits when it is used to describe average rainfall intensity. Most notable of these advantages is the existence of an upper bound when the standard deviation has a value higher than the mean. This ensures long duration events will not have unrealistically large intensities associated with them producing large abnormal storm depths, a problem which was observed and identified by Heneker (2002).

Another advantage of the GPD is that its two parameters can be directly related to the mean and standard deviation of the observed population. The resulting model of conditional intensity is given by the following equations

$$\ln(1-F(x)) = \frac{1}{\theta_1}\ln\left(1-\theta_1\frac{x}{\theta_2}\right) \tag{3.7}$$

$$\theta_1 < 0, \theta_2 > 0$$

where the parameters of the GPD θ_1 and θ_2 are dependent upon the mean and standard deviation and are given by

$$\theta_1 = \frac{1}{2}\mu\left(\left(\frac{\mu^2}{\sigma^2}\right)-1\right),\ \theta_2 = \frac{1}{2}\mu\left(\left(\frac{\mu^2}{\sigma^2}\right)+1\right) \tag{3.8}$$

and the mean μ and standard deviation σ are conditional on the storm duration denoted as

$$\mu, \sigma = f(\ln(t_d)) \tag{3.9}$$

where t_d is the corresponding storm duration.

This relationship between the GPD parameters and the sample event statistics ensured the conditional relationship between storm duration and intensity could be modelled. To gain an insight into the relationship between the mean and standard deviation of intensity against duration, the historical intensity-duration pairs are ordered from shortest to longest duration and then divided into groups of 50 consecutive events. The mean and standard deviation in each group can be calculated and these are then plotted against storm duration. A typical result for the mean average storm intensity versus duration can be seen in Figure 3.6 for Adelaide.

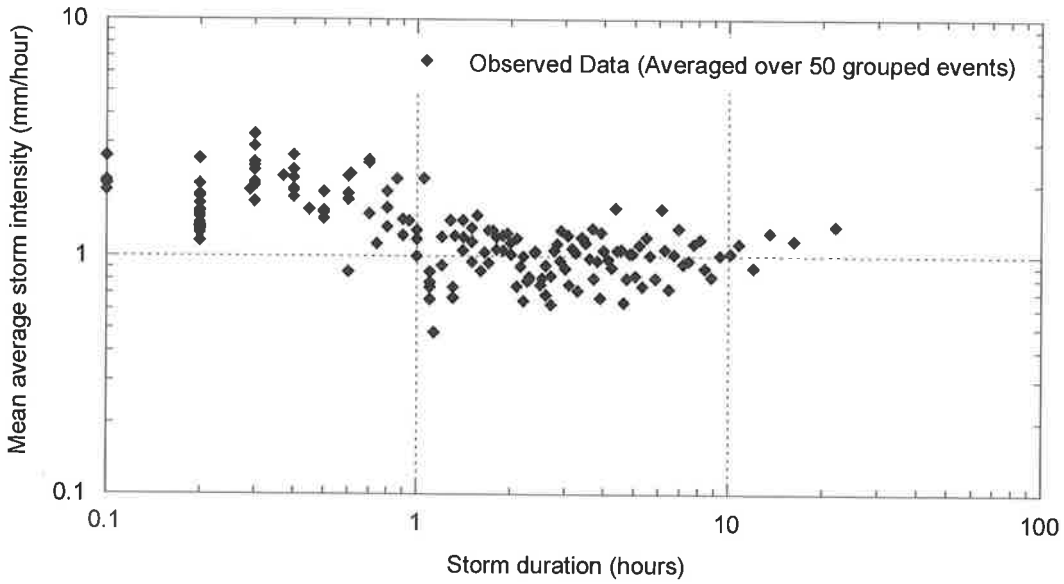

Figure 3.6: Mean average storm event intensity against duration for Adelaide

The values of μ, θ correspond to the mean and standard deviation of the average storm event intensity. In order to model these changes of μ, θ with storm duration a piece-wise linear relationship was developed. A series of straight line segments link the changes in slope of the conditional relationship. Breakpoints are manually included at certain durations by the modeller to ensure sections between breakpoints are predominantly linear. For the above Adelaide data, breakpoints were set at 0.1, 0.2, 0.3, 1.0, 3.0, 10, 18 hours. The resulting piece-wise linear model of mean average storm intensity can be seen in Figure 3.7. During simulation, the values of μ, θ can be calculated from this continuous piece-wise linear function and are then used to estimate the GPD model parameters in order to generate an average intensity. In this fashion the complex conditional relationship between intensity and duration was modelled successfully. Seasonality can again be incorporated by the use of monthly parameters or harmonics and maximum likelihood techniques are employed to find the optimum parameter set.

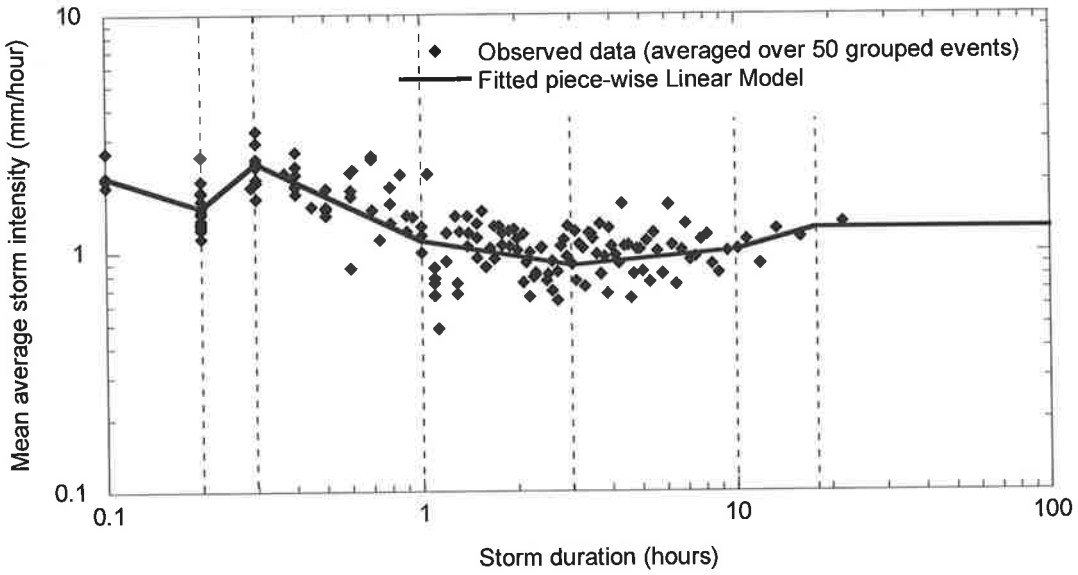

Figure 3.7: Fitted piece-wise linear model of mean average storm event intensity v duration for Adelaide

3.2.6 Disaggregation of Rainfall Events

Rainfall models that generate rainfall at time scales in the order of minutes require a method to ensure the intra-storm variability of rainfall is reproduced. Each observed rainfall event has a corresponding temporal pattern i.e. internal periods of varying rainfall intensity over time mixed with periods of no rain. Once the rain event series and corresponding intensity has been simulated, a disaggregation procedure can be included to replicate these internal storm characteristics.

The main idea behind the disaggregation scheme introduced by Heneker *et al.* (2001) is that the temporal storm pattern can be conceptualised as a conditional random walk on a dimensionless mass curve. If the rainfall trace of a storm event is considered, when moving from one time step to the next the trace can either move upward corresponding to a rainfall period, or remain horizontal indicating an internal dry period. This is shown diagrammatically in Figure 3.8.

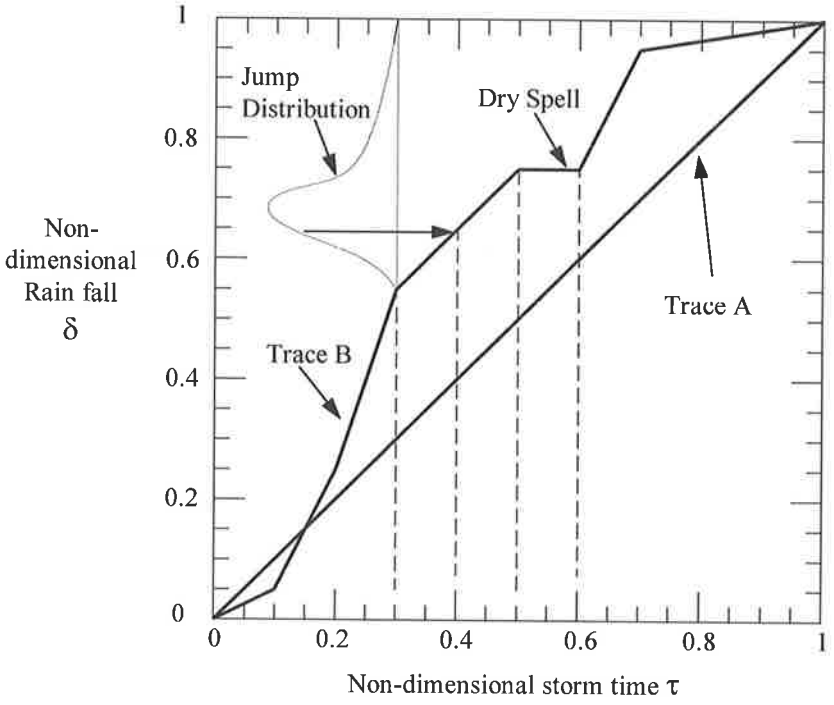

Figure 3.8: A non-dimensional description of the rainfall temporal pattern

The process used to describe the progression in this dimensionless mass space is assumed to be a self-similar, discrete stochastic process. The self-similarity concept where storms are assumed to exhibit similar internal properties despite differences in storm duration and depth provides the ability to simulate high resolution temporal patterns for long duration storms. This has been used with success previously by Woolhiser and Osborn (1985) and Koutsoyiannis and Foufoula-Georgiou (1993).

The disaggregation procedure developed by Heneker *et al.* (2001) separates the internal wet and dry periods of a rainfall event and considers them separately. During the simulation process, the required wet and dry periods are generated and then using a non-replacement-sampling scheme, the dry periods are interspersed within the wet periods to produce the final temporal pattern. These procedures are now investigated in turn.

3.2.7 Intra-Storm Rainfall

The treatment of the internal event rainfall periods begins by removal of any intra-storm dry periods. The rainfall periods are then consolidated into a continuous sequence of varying rainfall intensity or increasing depth providing a rainfall trace.

For each observed rainfall trace, storm duration is non-dimensionalised (from Heneker *et al.* 2001) by $\tau = t/t_d$ where t is the time since the start of the storm and t_d is the total storm duration. Depth is non-dimensionalised by $\delta = d(t)/d(t_d)$ where $d(t)$ is the cumulative rainfall up to time t and $d(t_d)$ corresponds to the total event depth. All rainfall traces therefore lie between (0,0) and (1,1) and have a non-negative slope.

In order to describe the progression of the rainfall trace, the non-dimensionalised space τ is initially divided into ten finite intervals. A jump distribution is used to describe the progression through the non-dimensionalised space from one interval to the next. By analysing the histogram of all observed jumps in each tenth of the dimensionless event duration space, an assumption for this jump distribution can be made. As shown in Heneker (2002) and reproduced in Figure 3.9, Figure 3.10 and Figure 3.11 (for Melbourne, Adelaide and Sydney respectively), the shape of the histogram suggests that a log-normal distribution can be used.

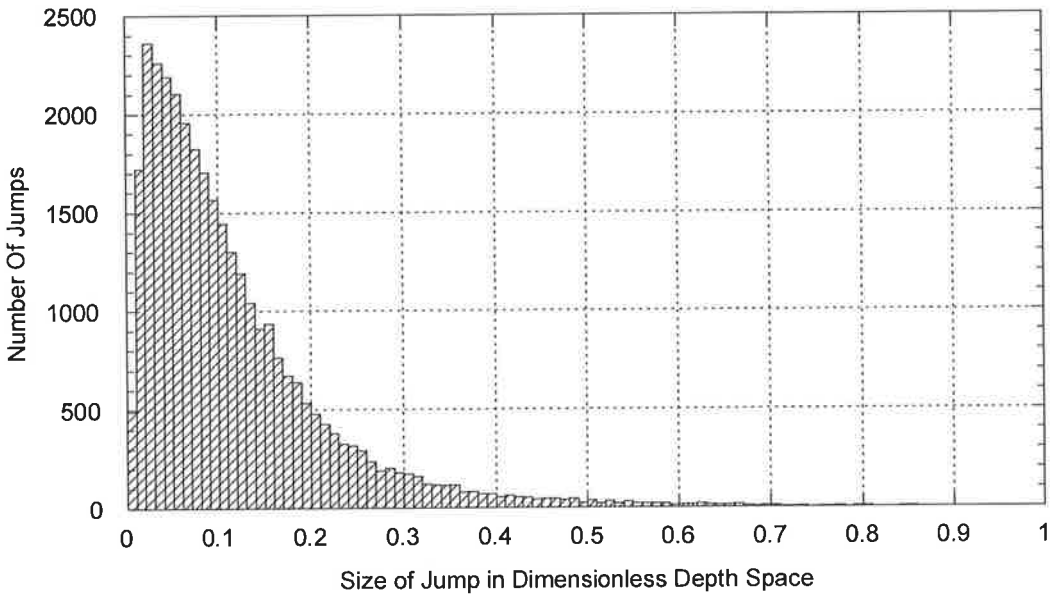

Figure 3.9: Melbourne Jump Distribution

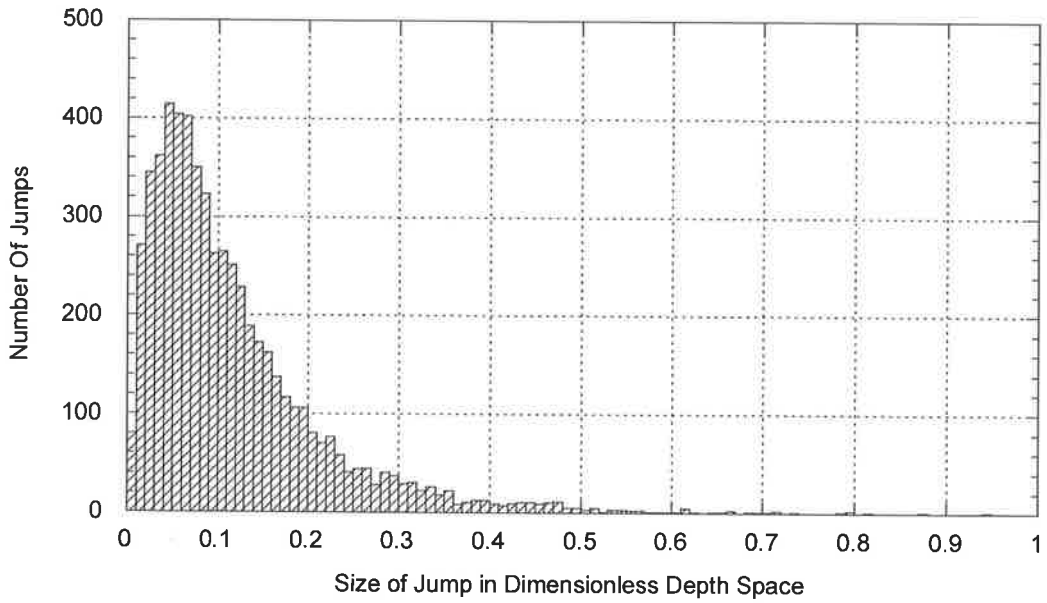

Figure 3.10: Adelaide Jump Distribution

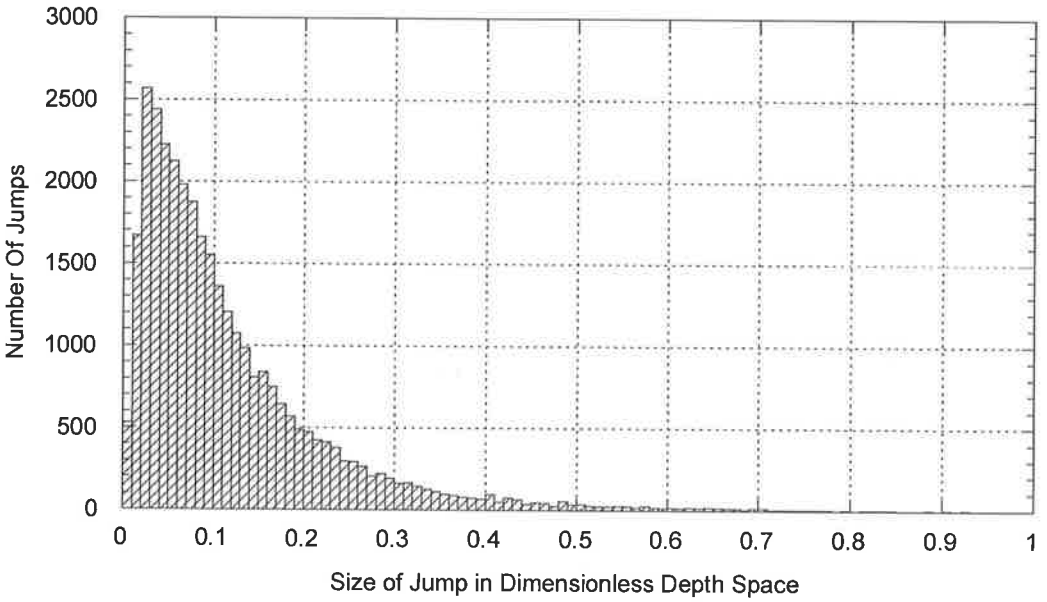

Figure 3.11: Sydney Jump Distribution

The first jump from the origin (0,0) to the first duration interval of $t_d/10$ can be described by the marginal distribution of all initial jumps in the observed record. However subsequent jumps must be conditional on the current location in the non-dimensionalised space. This ensures that the rainfall trace always ends up at (1,1) after disaggregation. A truncated log-normal distribution with a jump mean m and a

jump standard deviation s was used to model the observed jumps of the rainfall trace inside each interval. To ensure the mean and standard deviation of these jumps tend towards zero (for constraint) as the rainfall trace approached the end of the dimensionless storm, the following parabolic curves were used to describe the parameters m and s. These were then calibrated to the observed rainfall traces.

$$m=(1-\delta_{i-1})(m_1+m_2\delta_{i-1}) \tag{3.10}$$

$$s=(1-\delta_{i-1})(s_1+s_2\delta_{i-1}) \tag{3.11}$$

$$\text{where } \delta = \frac{d(t)}{d(t_d)}$$

The distribution of internal rainfall jumps used during the disaggregation process is therefore predominantly a conditional relationship based on what has transgressed previously throughout the storm. For the purposes of simulation, rainfall events are non-dimensionalised and divided into 10 initial intervals each one tenth of the total storm duration in length. Rainfall depths are calculated and assigned through the disaggregation process into each of these intervals. The assumption of self similarity then allows the disaggregation scheme to be applied to each of these 10 intervals separately and so on. This process continues until the length of these intervals are less than or equal to the required time resolution.

3.2.8 Internal Storm Dry Periods

Once the wet intervals of the storm event have been simulated and disaggregated down to the required time scale, the insertion of dry intervals produces typical historical temporal patterns. To do this, a dry spell fraction P is introduced, which defines the number of dry increments in any given storm. Heneker *et al.* (2001) characterises storm events as belonging to one of three groups with regards to the P value. Storms with a total duration of less than 0.5 hours are assumed to have no dry periods and are assigned a P value of 0. During the simulation process, any storm where $t_d < 0.5$ did not require the insertion of internal dry periods.

After the treatment of all short storms ($t_d < 0.5$ hours), Heneker *et al.* (2001) divided the remaining storms into two groups. Through an analysis of the distribution of intra-event dry fractions (see Figure 3.12), they found a significant number of storms had a low P (P_L, $0<P<0.05$) value between 0 and 0.05. The shape of the distribution of all remaining storms with a higher P value (P_H, $P > 0.05$), indicated that it could be modelled using a Beta distribution with a lower limit of 0.05 and an upper limit (P_{MAX}) defined as the maximum P observed in the data. The storms allocated a low P value (P_L) where modelled using a uniform distribution.

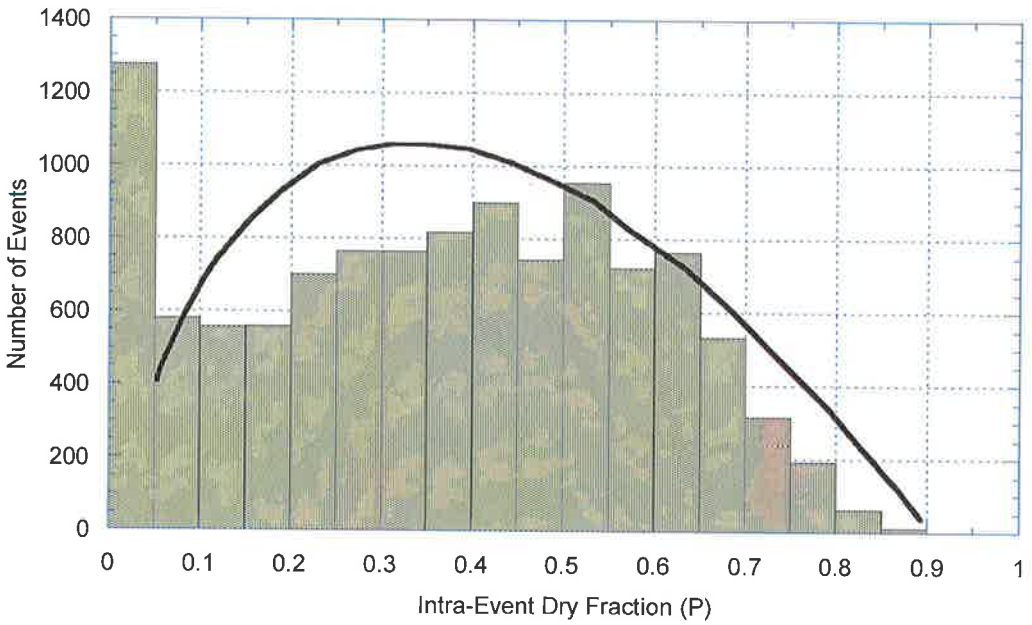

Figure 3.12: Distribution of intra-event dry fractions for Melbourne.

The calibration process requires the calculation of four parameters from the observed data. These relate to the two parameters of the Beta distribution, the probability of P_L and P_{MAX}. During a model simulation, the sampled P value determines the number of dry intervals that are interspersed into the rainfall trace. To insert these dry periods, a random insertion without replacement scheme is employed to intersperse these dry intervals into the simulated rainfall trace, producing the temporal pattern of the rainfall event. Two constraints are introduced to control this insertion process. Firstly, any dry period within the storm must not exceed the minimum inter-event time of two hours, which was initially used to determine storm independence. Secondly, the first and last interval of the storm must always be wet to ensure the storm duration remains correct.

3.2.9 Summary

The model results presented by Heneker *et al.* (2001) show the model to be capable of generating synthetic rainfall data down to time resolutions in the order of minutes. The reproduction of short duration IFD values at most sites gives an indication of the effectiveness of the disaggregation procedure. Observed and simulated inter-event times and storm durations compare well, as does the mean of annual rainfall. Indicative plots comparing observed and simulated data from Melbourne for IFD, monthly and annual rainfall are shown in Figure 3.13, Figure 3.14 & Figure 3.15 respectively and display the ability of the model to capture these statistics.

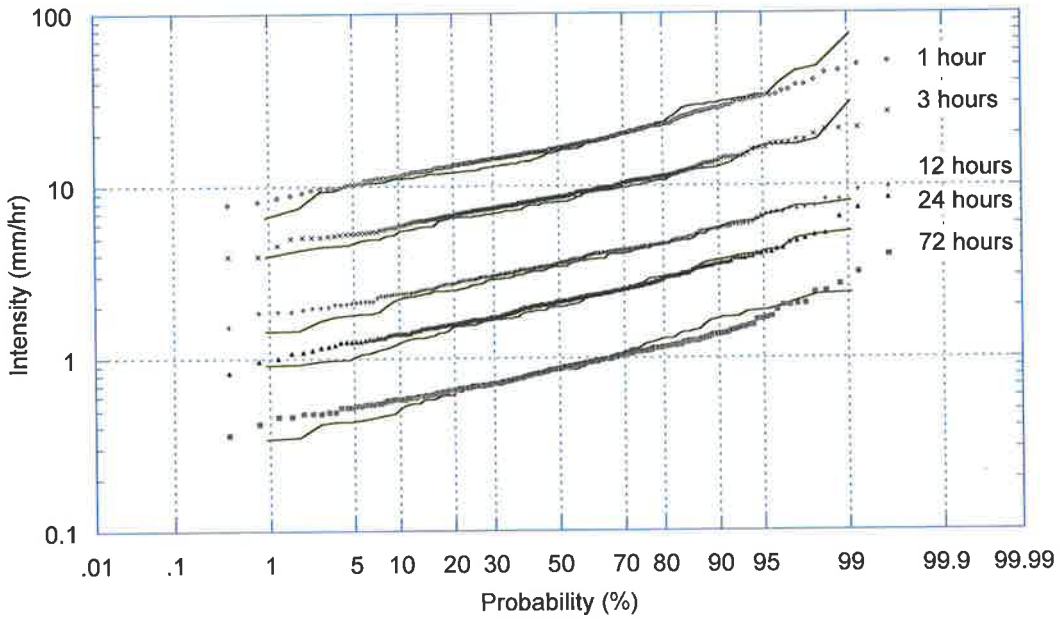

Figure 3.13: Simulated and observed IFD probability distributions for Melbourne

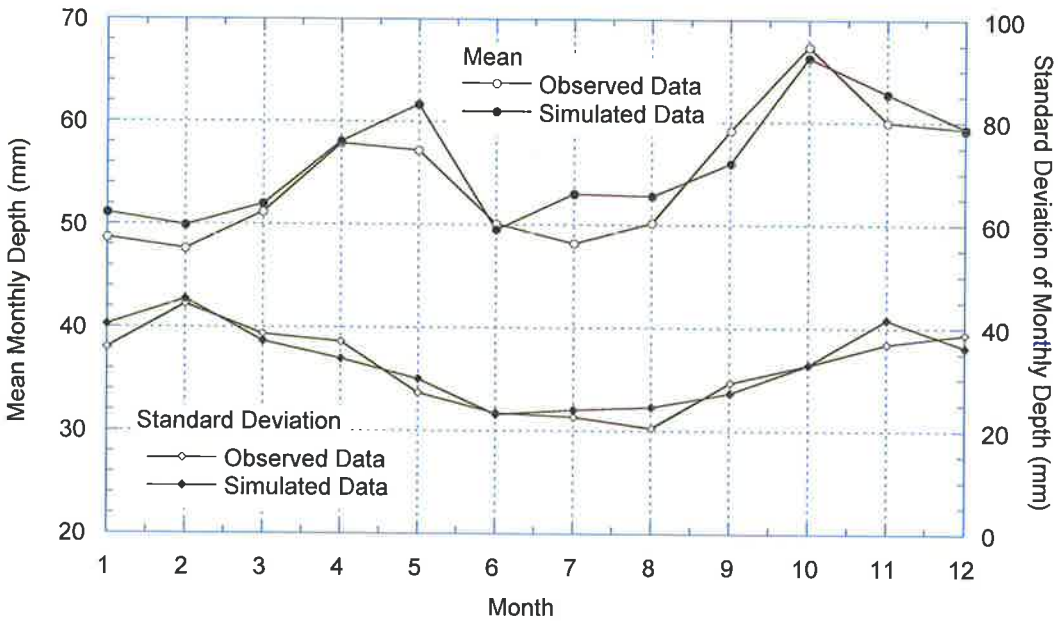

Figure 3.14: Mean and standard deviation of monthly depth for Melbourne.

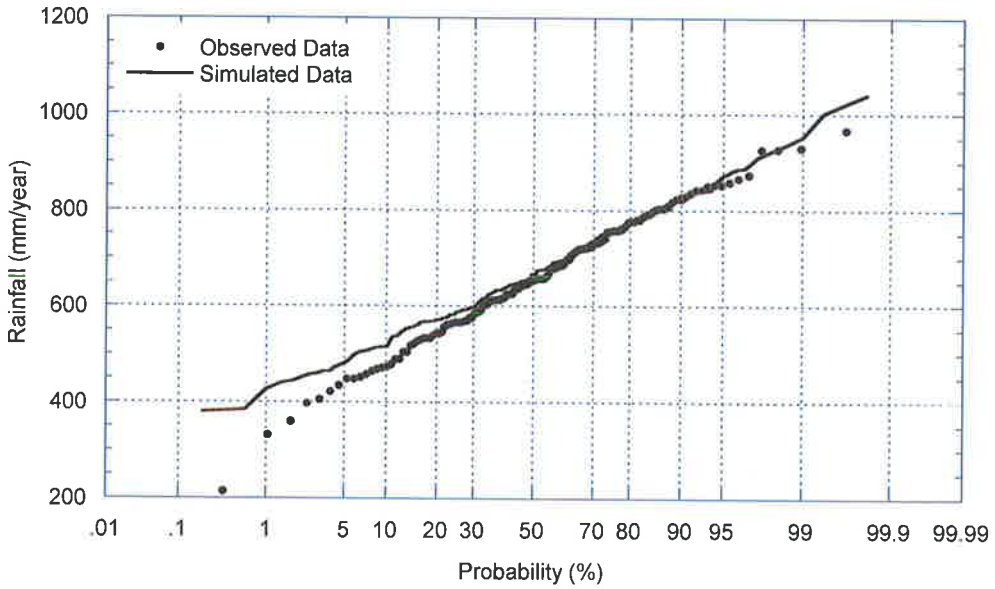

Figure 3.15: Probability distribution of annual rainfall for Melbourne.

Even though the model was deemed a success, the model in its original form contained areas that could be significantly improved. While investigating the performance of the original model and calibrating to numerous rainfall sites across Australia, it became apparent that model parameters were not always well determined by a global search routine. Further analysis has indicated the existence of significant

parameter correlations within the adopted probability structure used to describe inter event times and storm durations. The existence of these parameter correlations can be shown to induce 'flat' areas in the objective function space (regions where the likelihood is almost identical over a vast range of parameter combinations). This makes it difficult for global search routines to determine the optimum parameter values. For the purposes of regionalisation, it is important that model parameters are well defined and not significantly correlated so master calibrations can be adapted and used at other sites with confidence.

To improve the estimation of model parameters and provide an easier model to regionalise, an investigation into the relationships between model parameters has been undertaken. A well known technique known as the Metropolis algorithm (Lee (1989), Gelman *et al.* (1997)) (a Monte Carlo Markov Chain) has been incorporated into the new model and has allowed an insight into these parameter relationships. The result of this work is the complete removal of one superfluous parameter. Replacing the original 4 parameter distribution with a modified 3 parameter version has removed the significant parameter correlation, resulting in optimum model parameters that are well identified and assists in understanding the relationship between parameters in the model.

Another issue warranting attention was the procedure introduced by Heneker *et al.* (2001) to calibrate storm event intensity. The original calibration process required modeller input at various stages, resulting in the procedure being time consuming and potentially biased on the expectations and interpretations of the modeller. This was not the case for the calibration of storm inter-event time or storm durations which were operator independent. This manual intervention introduces an inconsistent calibration process which has the possibility to vary significantly between data sites and modellers. Analysis of the intensity-duration relationship has led to the development of an automatic calibration procedure which removes the need for modeller input and simplifies the calibration process. This has improved the reliability of the model and in turn its usefulness and functionality as a hydrological tool.

These ideas and results behind these two major improvements to the model are presented below.

3.3 Identifying and Removing Correlated Parameters

The ability to successfully identify optimum model parameters is a critical step in the application of hydrological models. With substantial improvements in computing power and improved searching algorithms, the identification of model parameters continues to become easier to accomplish. However, the ability to determine the true optimum parameters is particularly important for a model being used for regionalisation. This was summarised by Kuczera (1983) when he said:

"If the full potential of the regionalisation approach is to be realized, it is desirable that optimized parameters be close to their true values; that is, they should be precisely inferred or well determined."

The importance of accurate calibration was also discussed by Gaume *et al.* (1998) who noted that it (an accurate calibration) can not be over estimated and that correlations between parameters provides a major obstacle in the ability to accurately identify model parameters. The ability to analyse and remove correlated parameters would improve the calibration process and provide further evidence of a well developed model.

The existence of parameter correlations can influence the calibration process by yielding objective function surfaces that are complex with numerous flat areas and localised valleys. These surface structures can become problematic for search routines (and modellers) interested in finding the optimum parameter values. Correlations in the parameter structure can also result in an uncertainty in the fitted parameter values, which in turn may produce a poor simulation result. If parameter correlations can be identified, the model can be reconstructed/re-structured to remove the need for these parameters, thus improving the model.

For these reasons, improving the Heneker *et al.* (2001) model by removing parameter correlations was seen as an important step towards providing a robust model capable

of further development into a regionalised rainfall model. In order to analyse the Heneker *et al.* (2001) model parameters an estimation of the distribution of these parameters after calibration ($p(\theta|y)$) is required. Unfortunately, it is difficult in complex models with numerous parameters to obtain samples from this posterior parameter distribution directly. To circumvent this problem, the Metropolis algorithm, a member of the Monte Carlo Markov Chain simulation (MCMC) family, was incorporated into the original model. A brief description of the workings of the Metropolis algorithm is provided below. For a more detailed description of Metropolis, mathematical background and alternative MCMC's the author recommends Lee (1989), Gelman *et al.* (1997) and Gamerman (1997).

3.3.1 The Metropolis Algorithm

The Metropolis algorithm was developed initially by Metropolis *et al.* (1953) to deal with the calculation of chemical substance properties that are determined by the equilibrium of potential energy and the vector position of the chemical molecules. In its presented form, the algorithm is able to sample from the posterior distribution of parameters in a complex model such as this, where direct calculation of these distributions is not possible.

The idea of all MCMC's, is to simulate a random stepped walk through the parameter space whereby samples drawn at each step in the process eventually converge to the stationary posterior distribution of the calibrated model parameters. As samples are drawn sequentially and the next sampling distribution is dependent on the previous value, the process forms a Markov Chain. The required posterior parameter distribution is denoted $p(\theta|y)$ and is often called the target distribution. This refers to the fact that in reality the distribution is only inferred from the accepted simulation samples and in essence is targeting the true parameter distribution.

Initially samples are drawn from an approximation of the parameter distribution referred to as the jump distribution. As the algorithm proceeds, controlled adjustment of this jump distribution ensures subsequent samples are more likely to be similar to those that would be drawn from the real target distribution. During the process, drawn samples are accepted or rejected as a sample of the true target

distribution based on a certain test criteria. This technique allows the algorithm to converge towards the true target distribution. Mathematical proofs are available (see Gelman *et al.* (1997)) that show samples drawn after convergence are equivalent to drawn samples from the required target distribution. The adjustment of the approximate distribution at each step, which allows the simulation to converge to the target distribution, is the primary reason that this technique can be used with success in higher dimensional problems. This is not the case for one/low-dimensional approaches such as importance sampling where the distributions remain the same.

A key to Markov Chain simulation is running the simulation long enough to ensure that the distribution of current draws from the simulation are close to the required stationary distribution. The question in practice becomes just how long should the algorithm run to ensure convergence? If the algorithm is launched with several sequences each an independent Markov chain simulation, then the 'mixing' of these independent paths (occurs when each independent path is sampling from the same parameter region) provides evidence that the algorithm has converged. Various other parametric and non-parametric techniques exist which are available to test for convergence (see Gelman *et al.* (1997)).

To gain an understanding of the workings of the algorithm, an example situation with 4000 randomly generated data points from a normal distribution with a mean of 10 and a standard deviation of 10 is presented. An SCE_UA search was conducted initially, assuming an underlying normal distribution and using an objective function based on Maximum Likelihood criteria. This search resulted in returned optimum parameters of 10.012 (mean) and 9.914 (Standard Deviation). Due to the large number of data points (4000) and a two parameter distribution (normal) it is reasonable to expect the calibrated parameters are well defined and concentrated around the optimum modal parameters. The steps involved in then applying the Metropolis algorithm are as follows

1. A starting point θ_0 is drawn which can be randomly selected or centred about the modes/best fit of the distribution providing that $p(\theta_0 | y) > 0$ (i.e. the initial point is feasible).

2. A candidate point θ^* is sampled from the approximate or jump distribution $Jt(\theta^*, \theta_{t-1})$ at time t.
3. The ratio of the densities at these points is calculated, that is $r = \frac{p(\theta^* | y)}{p(\theta_{t-1} | y)}$
4. A random number is generated between 0 and 1. If this value is less than r, then the candidate point is accepted as a sample of the posterior distribution. If this value is greater than r the candidate point is rejected. Once a point has been accepted, the algorithm shifts to this point for the next iteration. If the candidate point is rejected, then the algorithm stays in its current position. The acceptance/rejection process ensures that the algorithm will always accept candidate points which have a greater density than that of the previous time step while also sometimes accepting points which have a lower density. This enables the algorithm to explore the entire parameter space and eventually converge on the correct posterior distribution.
5. Repeat the process iteratively until convergence.

After assigning random starting points inside the parameter space, Figure 3.16 shows the first 30 iterates of the algorithm. The solid squares are included to indicate the starting points of the 4 independent path sequences. It can be seen that the random walks have each traced a path through the parameter space, however it is clear at this early stage that convergence has not yet been reached.

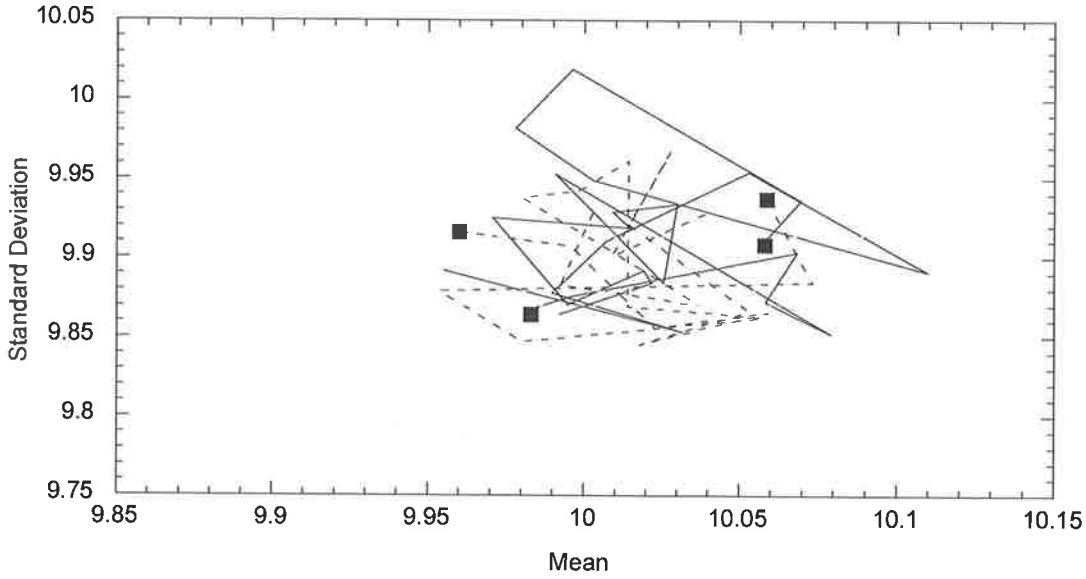

Figure 3.16: 30 Independent Sequences of a Markov Chain Simulation of a Normal Distribution. (Modal Parameters of 10.012 and 9.914)

Figure 3.17 presents a more mature simulation after the first 1000 iterations. The sequences are much nearer to convergence indicated by each independent path spending more time around the same position in the parameter space. This mixing of the 4 paths gives an indication that the model is converging and therefore drawing samples that are more likely to be from the underlying posterior distribution. It is also evident that the independent sequences have been able to successfully explore the parameter space as evidenced by accepted samples near the extremities. The sequences now almost have a common stationary distribution of $p(\theta | y)$.

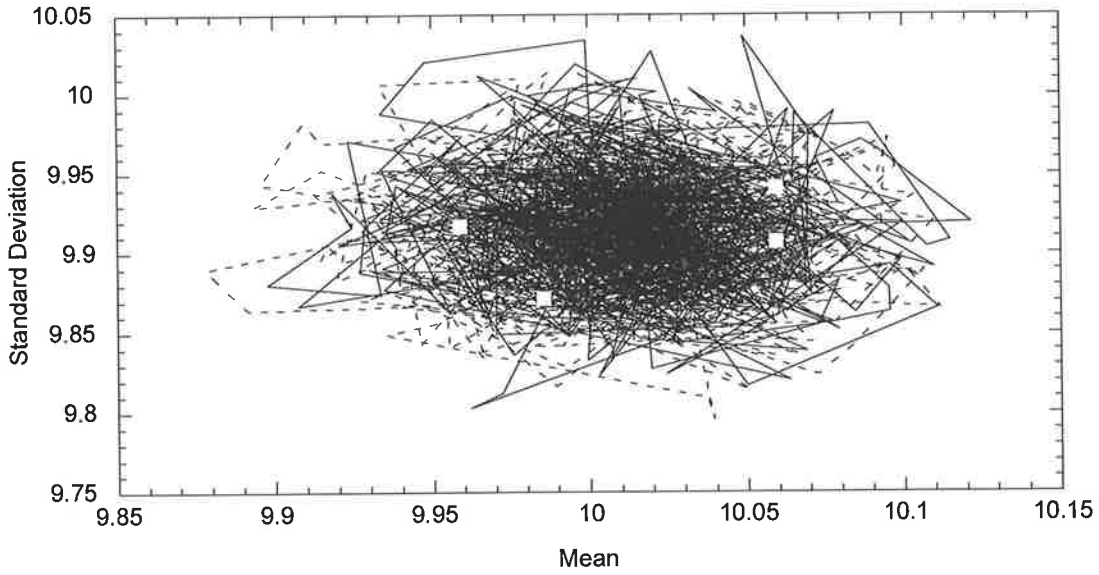

Figure 3.17: 1000 Independent Sequences of a Markov Chain Simulation of a Normal Distribution. (Modal Parameters of 10.012 and 9.914)

Given a mature converged simulation (Figure 3.17), the posterior parameter distributions can now be inferred from subsequent samples. Figure 3.18 displays the iterate cloud of the final 2000 samples. These samples can be considered as draws from the target distribution and thus give the parameter distributions relating to the calibrated mean and standard deviation of the assumed normal model.

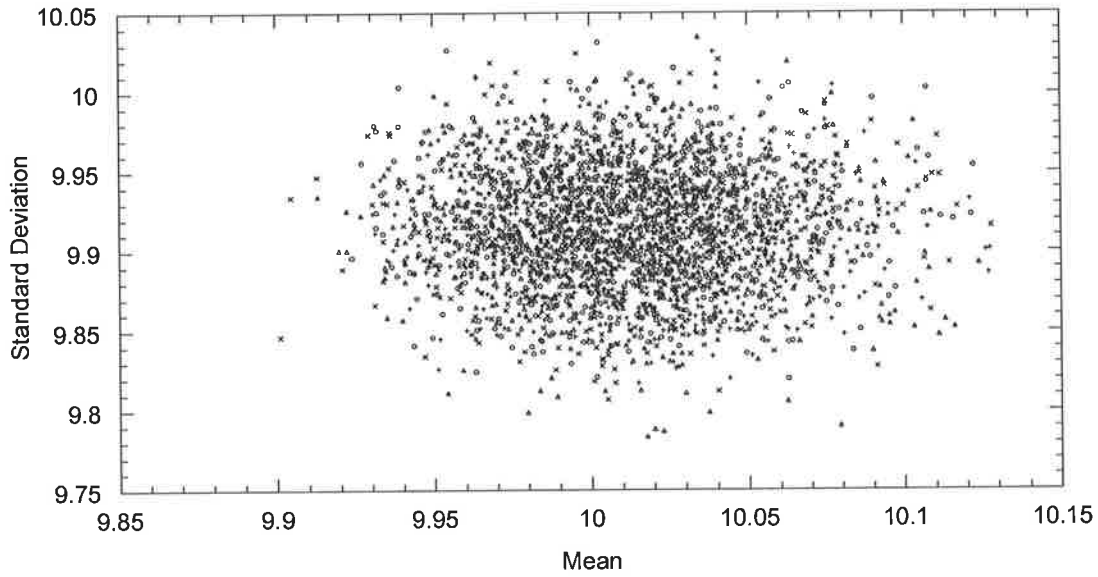

Figure 3.18: Final 2000 Independent Iterates of a Markov Chain Simulation of a Normal Distribution. (Modal Parameters of 10.012 and 9.914)

The efficiency of the Metropolis algorithm is related to the selection of the jump distribution. Primarily a good jump distribution must be easy to sample from. It must also ensure that each jump goes a reasonable distance in the parameter space so that the random walk does not explore the parameter space too slowly and that candidate points are not rejected too frequently so that the algorithm doesn't waste time standing still. Gelman *et al.* (1997) provide suggestions to improve the efficiency of algorithms which are progressing slowly. Typically for a multivariate normal distribution, the optimal jumping rule should have an acceptance rate around 0.44 in one dimension or 0.23 in high dimensions. If the simulation is proceeding with an acceptance rate significantly different to these values, then the following adjustments can be included into the algorithm to improve efficiency.

1. After a certain number of iterations, the covariance of the jumping distribution can be adjusted to be proportional to the posterior covariance matrix estimated from the accepted simulation samples.
2. The scale of the jumping distribution can be increased or decreased if the acceptance ratio is too high or too low respectively.

Monte Carlo Markov Chain simulation provides a useful technique to determine the parameter distributions in complex models. The algorithm is relatively simple to implement and relies on the ability to draw samples from the jumping distribution, the ability to calculate r and the ability to generate random numbers.

3.3.2 Incorporating The Metropolis Algorithm into the Rainfall Model

The Metropolis algorithm in the form presented above was incorporated into the structure of the Heneker *et al.* (2001) model to allow parameter uncertainty to be examined. The likelihood functions in the model provided an ideal objective function for determining the required value of r. In order to implement the Metropolis algorithm, the only remaining requirement was the selection of an appropriate jumping distribution. Accepting the requirement that the jumping distribution must be symmetric for implementing the basic Metropolis algorithm, the selection is somewhat arbitrary. For the purposes of this study, a multivariate normal

distribution was chosen. The multivariate normal is symmetric and has the advantage that samples from this distribution are easily obtained. Generalisations of the Metropolis algorithm (namely the Metropolis-Hastings algorithm see Hastings (1970)) can handle the selection of jumping distributions which are not symmetric but this has not been required during this investigation.

The following values were used for all Metropolis simulations. Three parallel independent paths were simulated each with 11500 samples. The first 1500 samples of each path were discarded leaving a total of 10000 per path. Initially these independent sequences were seeded by sampling from a multi variate normal distribution centred at the optimum or modal parameters.

3.3.3 Dry Spell and Storm Duration Parameter Analysis using the Metropolis Algorithm

As previously discussed the calibration of inter-event times and storm durations are undertaken with a generalised exponential with a kernel function given by

$$\ln(1-F(x)) = \frac{1}{\theta_1}\ln\left(1-\theta_1\frac{x}{\theta_2}\right) - \theta_3 x^{\theta_4} \qquad (3.12)$$

$$\theta_1 < 0, \theta_2, \theta_3, \theta_4 > 0,$$

The four model parameters for both the inter-event and storm duration calibration can be referred to as the shape θ_1, location θ_2, constant θ_3 and exponent θ_4 parameters. A pair-wise comparison of the resulting Metropolis posterior parameter distributions shows the existence of significant parameter correlations in the original model structure.

A typical result where two parameters exhibit very little to no correlation is shown in Figure 3.19 (in this case the shape and location of an inter-event calibration for Sydney data). As can be seen from the bi-variate plot, movements away from the mode of the parameter produce a significant decrease in the goodness of fit as evidenced by a decrease in the density of accepted samples. The bi-variate normal

shape that is evident in Figure 3.19 is a good indication that this pair of parameters is well defined and has little correlation.

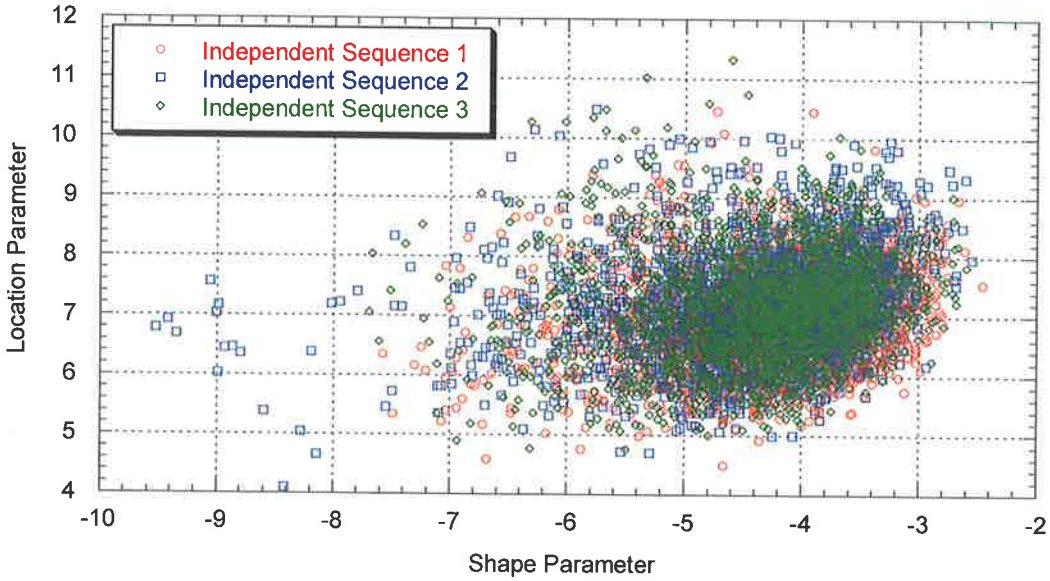

Figure 3.19: Scatter Plot from Metropolis Output Comparing Shape θ_1 and Location θ_2 Parameters (Calibrated to Sydney Inter-Event Data January)

A similar plot of the constant θ_3 and exponent θ_4 parameters from the same distribution and with the same data shows a very different result. Figure 3.20 is an example of two parameters that exhibit a classical 'cigar' shape and thus have significant correlation. Changes in the fit of the model due to a change in one parameter have been compensated by a corresponding adjustment in the second parameter which ensures a similar goodness of fit. The density of accepted samples along the 'cigar' shape is similar suggesting that any point on the curve is as good a fit as any other. This can provide difficulties for the search routine in finding the optimum parameter values, and provides parameter estimates that have a high degree of uncertainty. An analogy of this problem is the simple sum of two variables, where there are an infinite number of combinations that can be chosen for values of x and y such that $x + y = 10$. Generally the existence of highly correlated parameters may not be a major problem, however for the purposes of regionalisation it is desirable that the master parameter values are well defined. Figure 3.20 through Figure 3.23 show similar plots for other data sets, other months and for the calibration of storm

duration indicating that this is a generic problem with the adopted hybrid probability distribution.

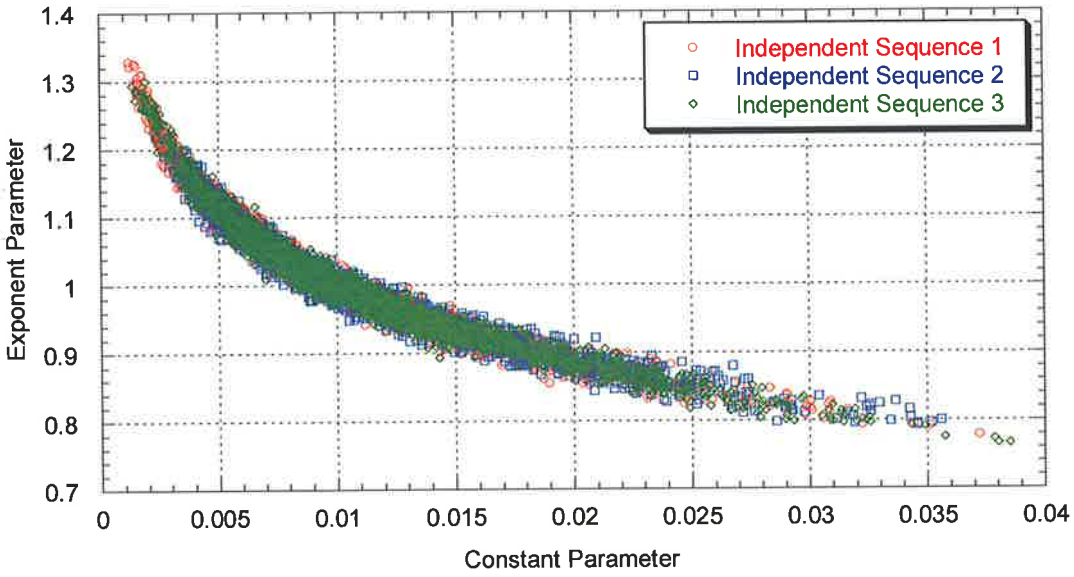

Figure 3.20: Scatter Plot from Metropolis Output Comparing Constant θ_3 and Exponent θ_4 Parameters (Calibrated to Sydney Inter-Event Data January)

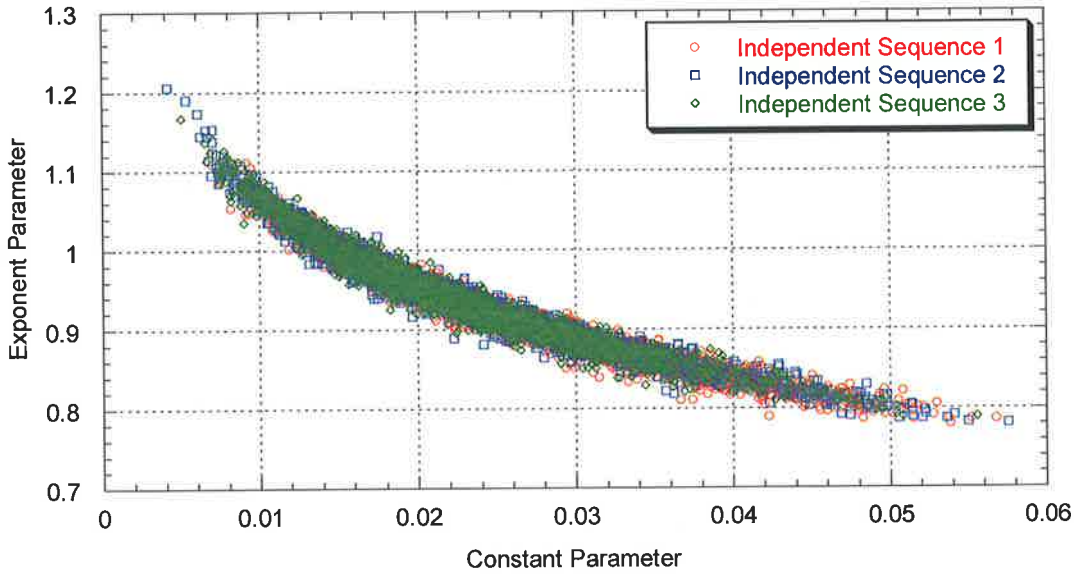

Figure 3.21: Scatter Plot from Metropolis Output Comparing Constant θ_3 and Exponent θ_4 Parameters (Calibrated to Melbourne Inter-Event Data June)

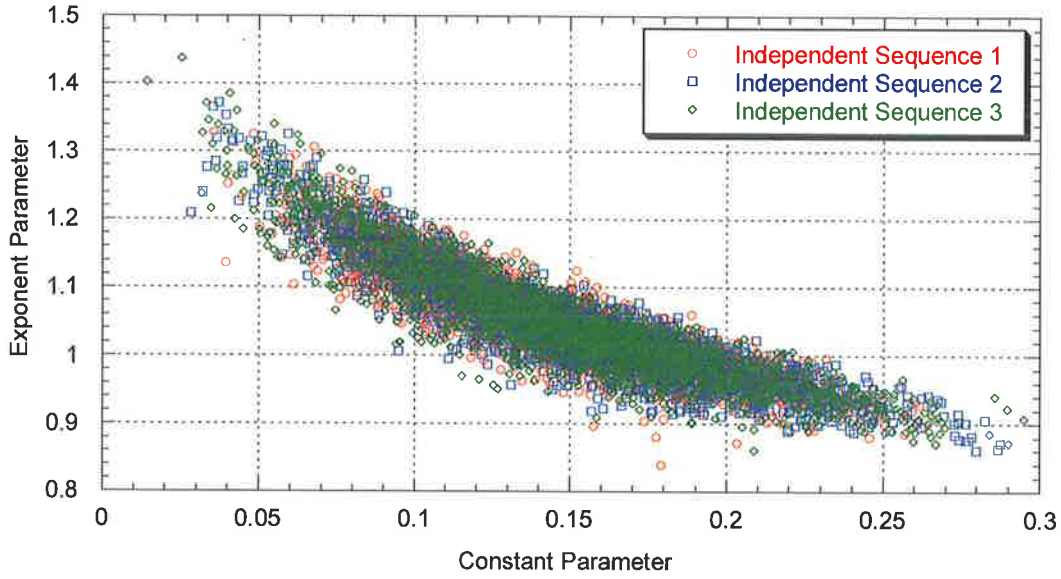

Figure 3.22: Scatter Plot from Metropolis Output Comparing Constant θ_3 and Exponent θ_4 Parameters (Calibrated to Melbourne Storm Duration data June)

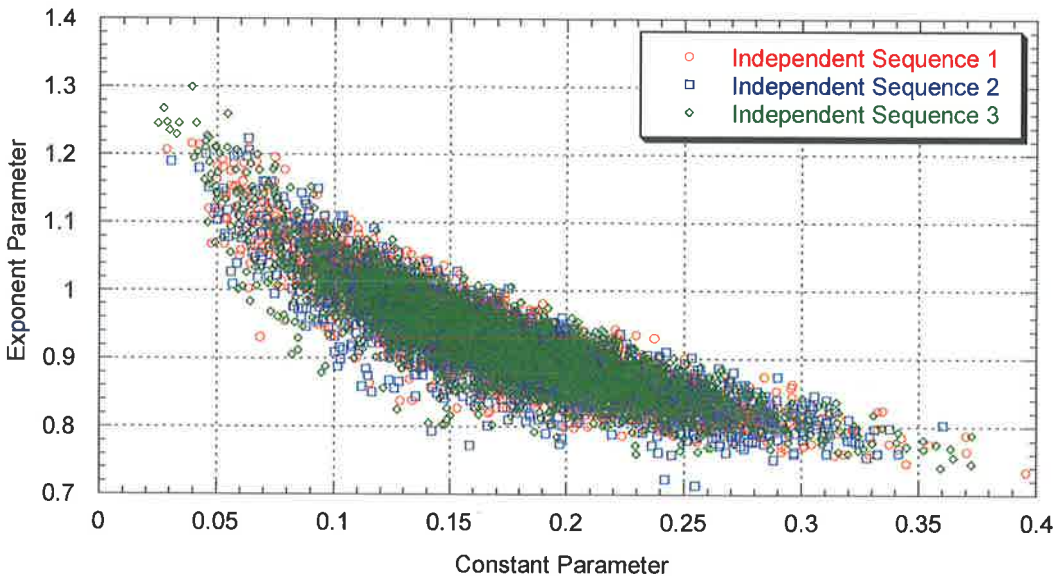

Figure 3.23: Scatter Plot from Metropolis Output Comparing Constant θ_3 and Exponent θ_4 Parameters (Calibrated to Brisbane Storm Duration data April).

3.3.4 Improvements to the Calibration of Dry Spell and Storm Durations through a 3-Parameter Model

A simplification to the hybrid four parameter distribution was chosen to describe the distribution of inter-event times and storm durations and remove the high parameter correlations displayed above. The nature of the cigar shaped posterior distribution suggests that an equally good fit could be obtained by fixing one parameter (either the exponent or constant parameter) at a certain value and leaving the other parameter variable to compensate. This would remove the parameter correlations while still providing an adequate description of the observed distributions.

To implement this proposed change, the model was initially calibrated using the original 4 parameter model. The average value of the exponent parameter was then calculated from this monthly calibration and the model re-fitted to the observed data keeping the exponent parameter constant. A comparison between this new 3 parameter model and the original 4 parameter version was undertaken by way of best likelihood estimates and plots of observed versus predicted values.

Figure 3.24 compares the maximum likelihood estimates for both the 3 parameter and 4 parameter model. While a slight decrease in the likelihood is noted (thus a worse fit), this minimal difference does not decrease the performance of the model (the *maximum* change in likelihood estimate from one model to the other was 6350.6 to 6354.7). Figure 3.25 compares the distributions of observed and calibrated data using both the 4-parameter and the 3-parameter model. As evidenced in these plots, there is very little difference in calibration accuracy between these two options. Analysis of the parameter distributions (Figure 3.26 and Figure 3.27) which are produced with the 3-parameter model indicate that all model parameters are now well defined and are without significant correlation.

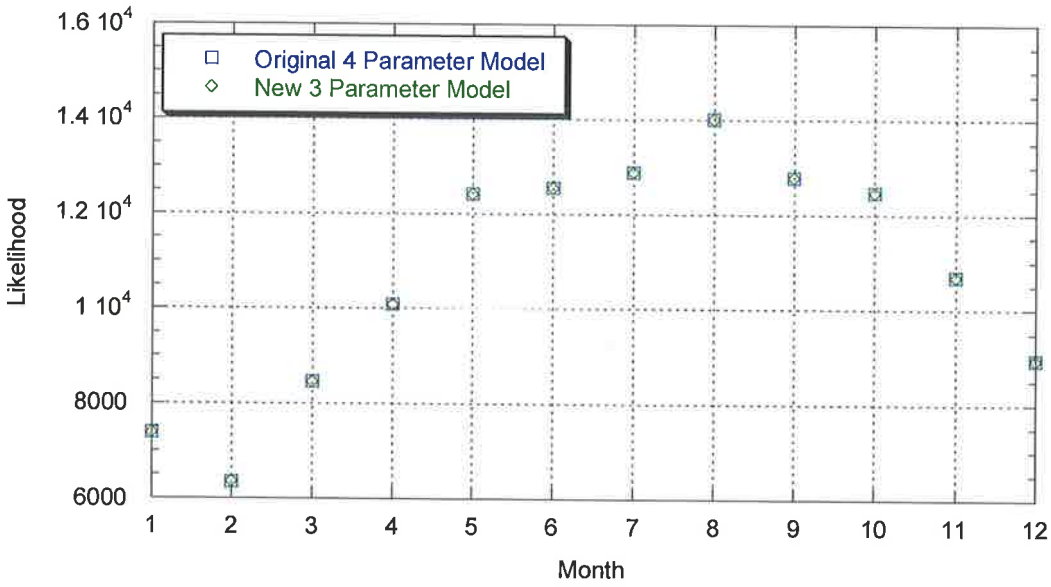

Figure 3.24: Comparison between Maximum Likelihood Function Values for the Original 4 Parameter Model and the New 3 Parameter Model (Calibrated to Melbourne Inter-Event Data)

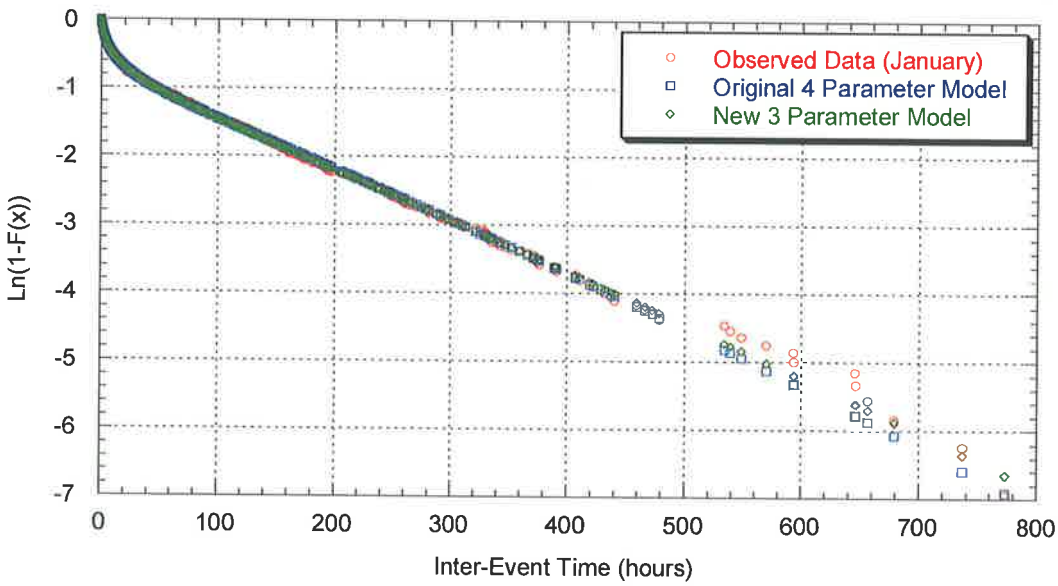

Figure 3.25: Calibration Plot Comparing Observed Data to Calibrations from the Original 4 Parameter Model and the New 3 Parameter Model (Calibrated to Melbourne Inter-Event Data January)

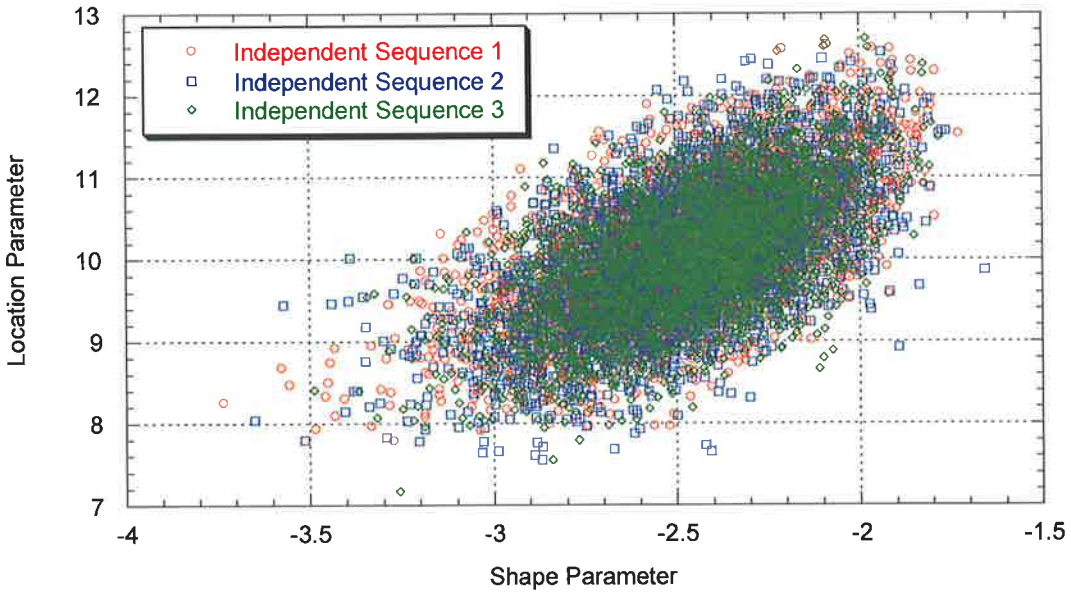

Figure 3.26: Scatter Plot from Metropolis Output Comparing Shape θ_1 and Location θ_2 Parameters (Calibrated to Melbourne Inter-Event data June).

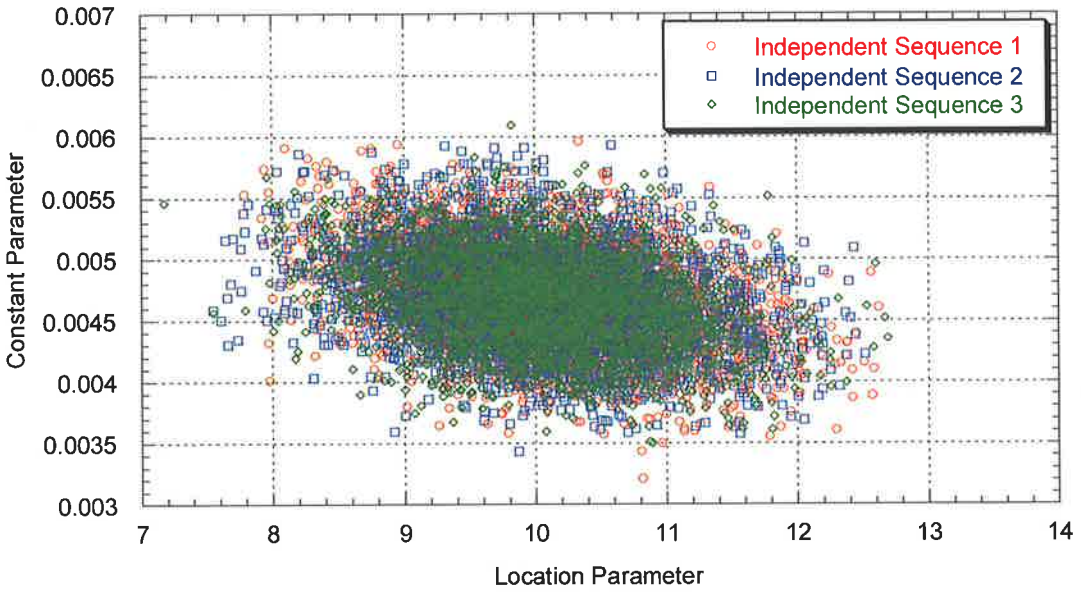

Figure 3.27: Scatter Plot from Metropolis Output Comparing Location θ_2 and Constant θ_3 Parameters (Calibrated to Melbourne Inter-Event data June).

3.3.5 Intensity Parameter Analysis using the Metropolis Algorithm

The intensity model presented by Heneker *et al.* (2001) requires the calibration of the two parameters (namely the mean and standard deviation) of a regular Generalised Pareto Distribution to the population of conditional intensity through a piece wise linear model structure. These parameters are directly related to regular probability distribution parameters and as such it was believed that the possibility of encountering significant parameter correlations was small. However for completeness, the Metropolis algorithm was again incorporated into the calibration of event intensity model and similar parameter distribution outputs as that presented for dry spell and storm duration parameters were produced. Figure 3.28 indicates the typical correlation between the mean and standard deviation parameter at a given breakpoint, in this case for Melbourne data. As expected the scatter plot indicates that there is little to no parameter correlation between the two parameters within the intensity model at a given breakpoint.

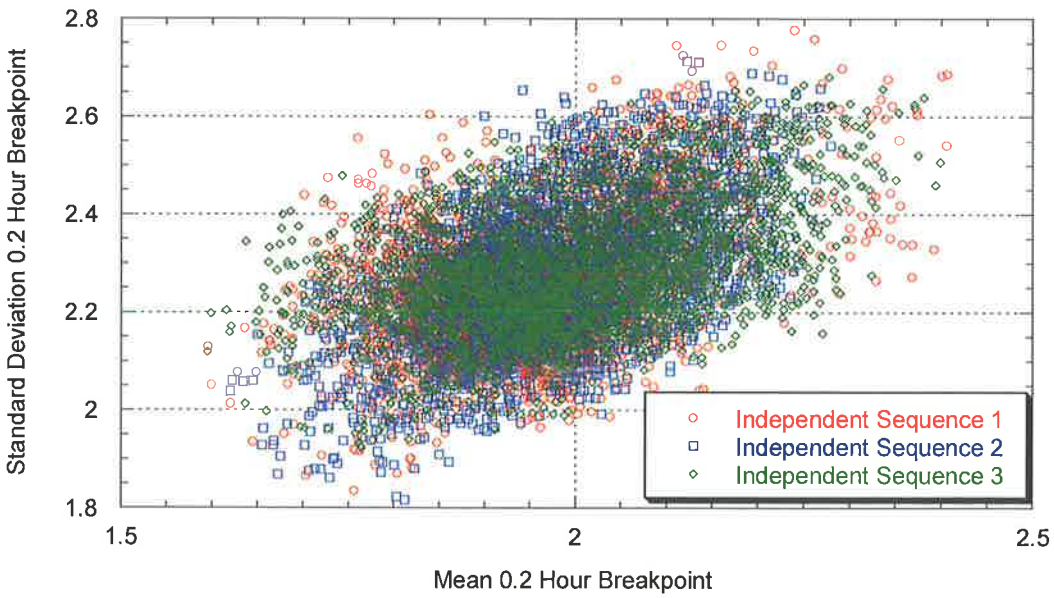

Figure 3.28: Scatter Plot from Metropolis Output Comparing the Mean and Standard Deviation Parameters, 0.2 Hour Storm Duration Breakpoint (Calibrated to Melbourne Storm data December).

The second comparison to check within the calibration of storm intensity is the possibility of significant correlation between parameters associated to different breakpoints. While the parameter values at subsequent breakpoints are intrinsically linked due to the structure of the piece wise linear model, does this relationship correspond to significant parameter correlations that make it difficult for the search routine to determine the optimum parameters? Figure 3.29 displays the correlation between the mean parameter at subsequent breakpoints for data from Melbourne. Again it can be seen that the shape of this scatter plot suggests there is little or no parameter correlation in the calibration of conditional storm intensity parameters.

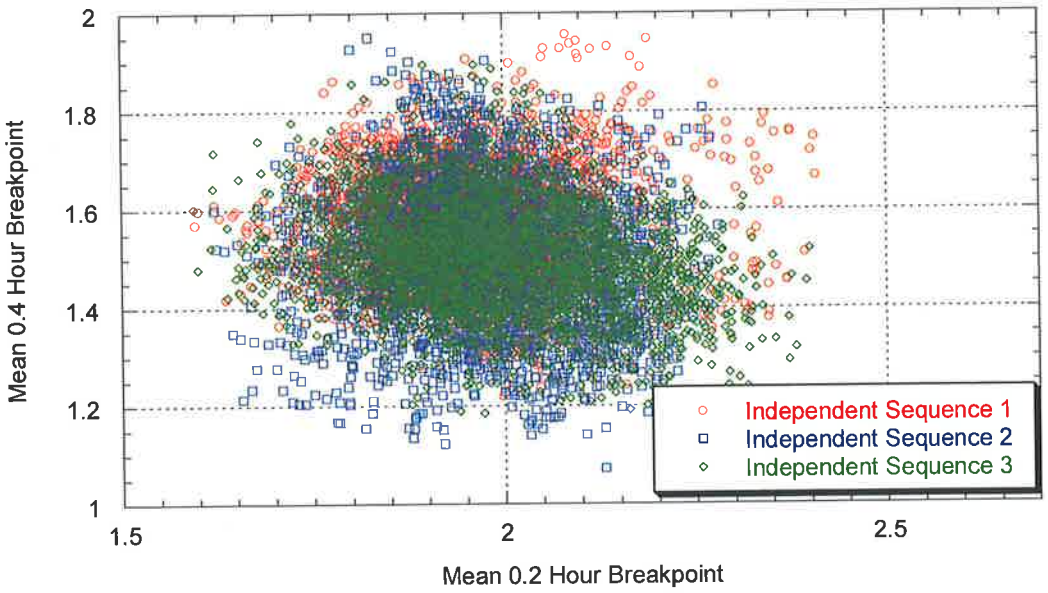

Figure 3.29: Scatter Plot from Metropolis Output Comparing the Mean Parameters, 0.2 Hour and 0.4 Hour Storm Duration Breakpoints (Calibrated to Melbourne Storm data December).

While these results from analysing the calibration of conditional intensity suggest that model parameters are clearly defined and the optimal parameter values can be found using an adequate searching algorithm, the process for developing the required piece wise linear relationship during the calibration process remained cumbersome and inconsistent. If the model was to be used within a regionalised framework, further improvements to this section of the calibration process were required.

3.4 Improvements to the Calibration of Storm Intensity

The original structure of the Heneker *et al.* (2001) rainfall model used a Generalised Pareto Distribution to model the distribution of event storm intensities. As discussed previously, the probability model was governed by the following equations

$$\ln(1-F(x)) = \frac{1}{\theta_1}\ln\left(1-\theta_1\frac{x}{\theta_2}\right) \tag{3.13}$$

$$\theta_1 < 0, \theta_2 > 0$$

where the parameters of the GPD θ_1 and θ_2 are dependent upon the mean and standard deviation and are given by

$$\theta_1 = \frac{1}{2}\mu\left(\left(\frac{\mu^2}{\sigma^2}\right)-1\right),\ \theta_2 = \frac{1}{2}\mu\left(\left(\frac{\mu^2}{\sigma^2}\right)+1\right) \tag{3.14}$$

and the mean μ and standard deviation σ are conditional on the storm duration denoted as

$$\mu, \sigma = f(\ln(t_d)) \tag{3.15}$$

where t_d is the corresponding storm duration.

The complex conditional relationship between the mean rainfall intensity (μ) and storm duration (t_d) denoted in (3.15) was modelled using a piece-wise linear relationship. A series of straight line segments were used to map the changes in slope of the conditional relationship, while breakpoints located at certain durations defined the start/end of each linear segment (see Figure 3.30). A similar relationship was used to model the relationship between the standard deviation of rainfall intensity and storm duration. At each breakpoint, the mean and standard deviation parameters from Equation (3.15) are calibrated using maximum likelihood techniques with the observed storm data. By calibrating the parameters of Equation (3.15) at each of the breakpoint positions, the piece wise linear relationship can then be used to determine parameter values for any storm duration.

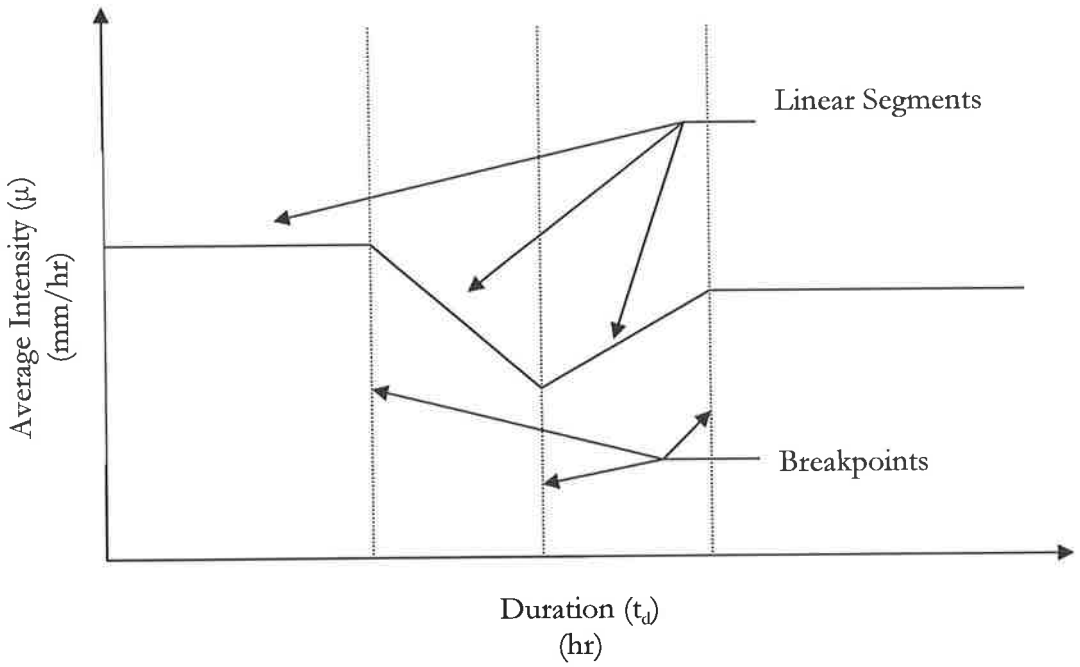

Figure 3.30: Schematic of Piece-Wise Linear Relationship

The original calibration procedure required significant modeller input during various stages of the process. Initially data pairs of intensity and the corresponding storm duration were ranked in order of increasing duration and then collected into bins of 50 consecutive points. The mean and standard deviation of these 50 storm intensities and the resultant average storm duration were then calculated and plotted on a log scale to produce a calibration plot. The initial position of the breakpoints within the piece wise linear model were then determined by analysing scatter plots of these average mean intensity against duration and average standard deviation of intensity against duration. Examples of these plots are shown in Figure 3.31 for mean intensity against duration and Figure 3.32 for standard deviation of intensity against duration.

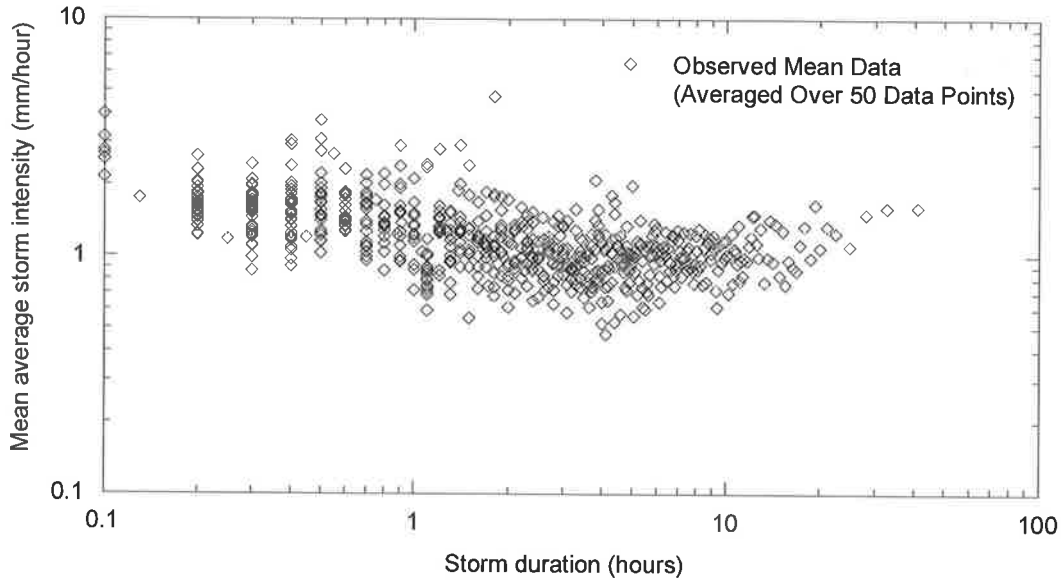

Figure 3.31: Calibration Plot for Mean Average Intensity (Data for Melbourne)

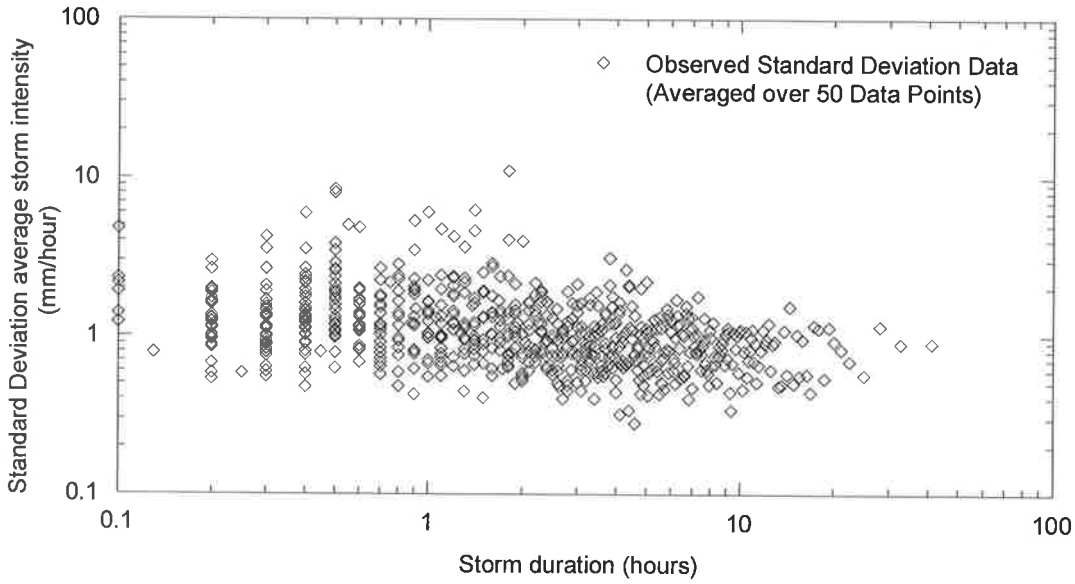

Figure 3.32: Calibration Plot for Standard Deviation of Average Intensity (Data for Melbourne)

The next step in the calibration process was to examine these plots and identify specific durations where the scatter cloud underwent a significant change of shape/slope. At these positions, a breakpoint was created in the piece-wise linear model in an attempt to describe the complicated conditional relationship. These breakpoints are indicated on the calibration plot (Figure 3.33) by a vertical straight line.

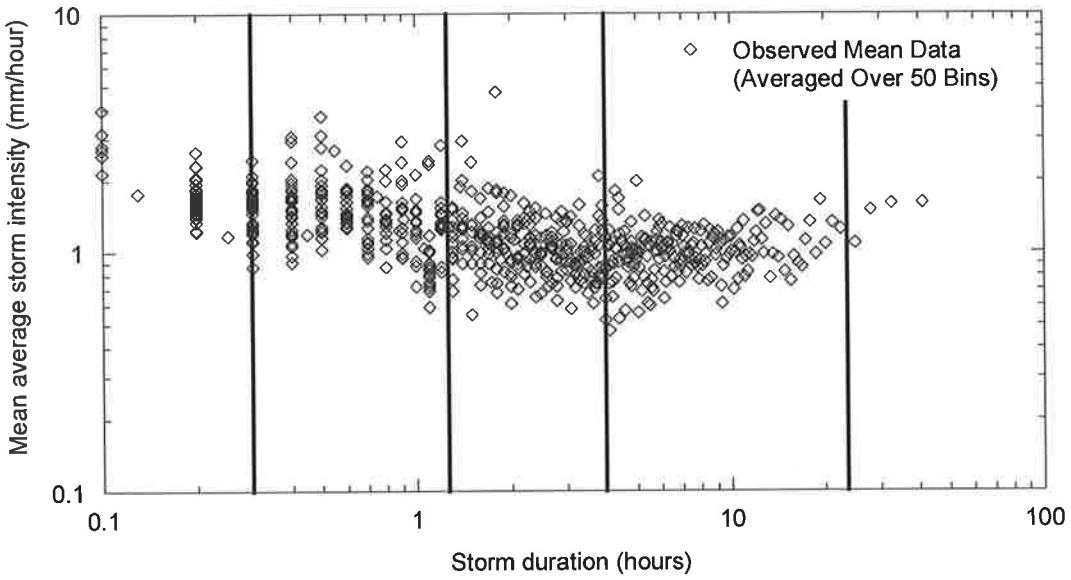

Figure 3.33: First Iteration Breakpoint Calibration

Calibration then proceeded by an iterative process where a simulation of the model would be run and extra breakpoints introduced to the relationship to improve the comparison between simulated and observed values. In

Figure 3.34 additional breakpoints have been introduced 0.5 and 0.9 hours with 0.2, 0.7, 2 and 10 hour breakpoints introduced in Figure 3.35. Analysis of the standard deviation of average storm intensity and storm duration introduced breakpoints in a similar manner. When the resultant simulation provided an accurate comparison to the observed statistics, the calibration ceased.

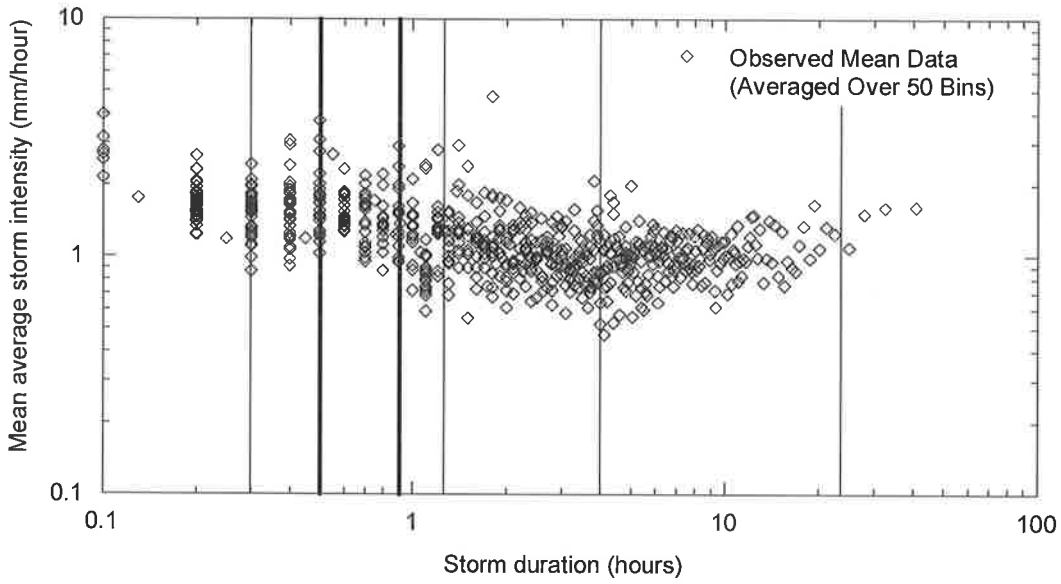

Figure 3.34: Second Iteration Breakpoint Calibration

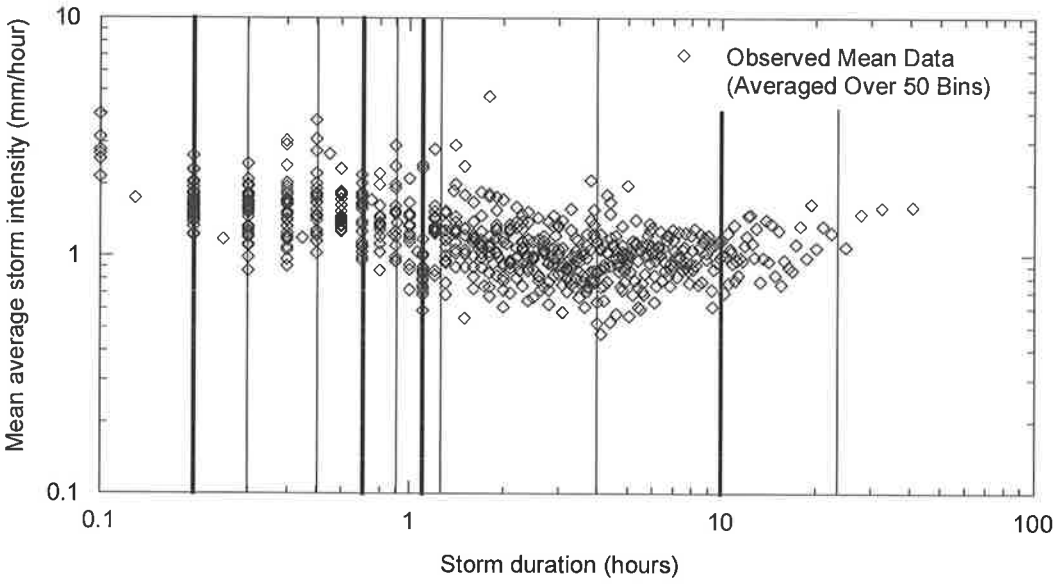

Figure 3.35: Final Breakpoint Calibration

The manual selection of breakpoints in its original form was inefficient and uncertain. The iterative process was time consuming and inconsistent as the breakpoint selection was dependent on each individual modeller's expectations and/or interpretations of the calibration plots. As a result the calibration may be poorly fitted by an inexperienced model user who does not select an adequate

number of breakpoints or over fitted by the introduction of additional breakpoint parameters which have little influence. Ideally, model results should be independent of the model user and should be designed to require no manual interaction reducing the chance of error and the time of calibration.

3.4.1 Investigation into the Intensity-Duration Shape

To develop an efficient automated calibration procedure, a better understanding of the shape of the conditional calibration relationship was required. If the shape could be shown to be consistent or predictable, then the breakpoint locations could be standardised across a number of sites reducing the need for manual intervention. There was also a possibility that a function could be used to describe the complex conditional relationship and remove the need to use the breakpoint linear segment approach altogether.

To investigate the shape of the conditional intensity – duration relationship, numerous sites across Australia were selected and their records adjusted to ensure each site contained concurrent data periods. Calibration plots were then produced for all sites and compared with the expectation that a similar conditional intensity-duration relationship would be observed across a number of sites. Such an expectation was not unreasonable given sites located within proximity of one and other should be influenced by similar rainfall patterns and events driven by large climatic factors. While there is likely to be small localised differences, a similar pattern should be observed overall. The results of the investigation provided below indicate that this is the case, and remarkably that the relationship is similar across sites a great distance apart.

A comparison between 5 capital cities in Australia illustrates the similarity in the structure of the intensity – duration relationship. Figure 3.36 presents the mean intensity - duration relationship for data from Perth, Adelaide, Brisbane, Sydney and Melbourne. If the relative scaling of sites is ignored (i.e. Brisbane generally sits well above other sites due to its higher rainfall), it can be seen that the intensity-duration relationship is consistent across a number of rainfall sites. While these Australian sites are a significant distance apart geographically, they still exhibit a similar

conditional intensity – duration relationship. Analysing a similar plot confirms the same result for the standard deviation of intensity in Figure 3.37.

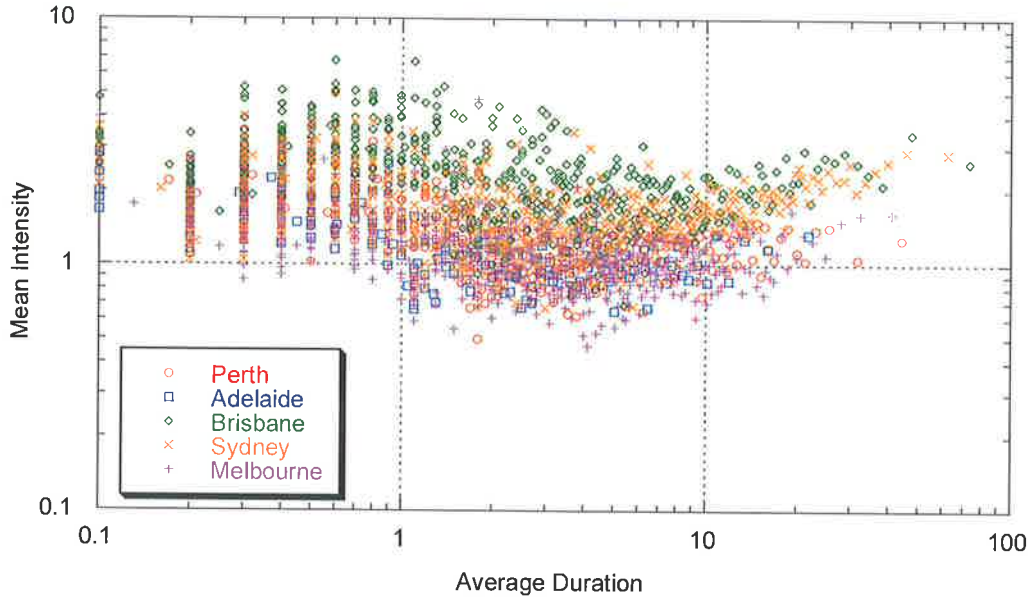

Figure 3.36: Comparison of Mean Intensity v Duration at Numerous Rainfall Sites

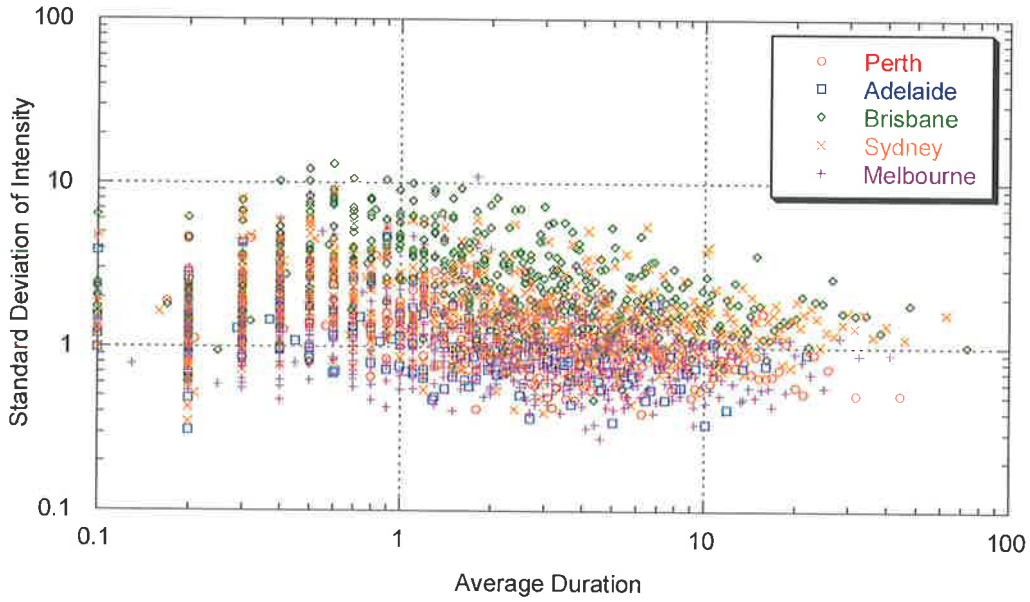

Figure 3.37: Comparison of Standard Deviation of Intensity v Duration at Numerous Rainfall Sites

If the investigation is refined further and local regional sites compared, the consistent shape is even more pronounced as is evidenced by Figure 3.38 and Figure 3.39 for sites in New South Wales where there is very little spread in the plotted data cloud.

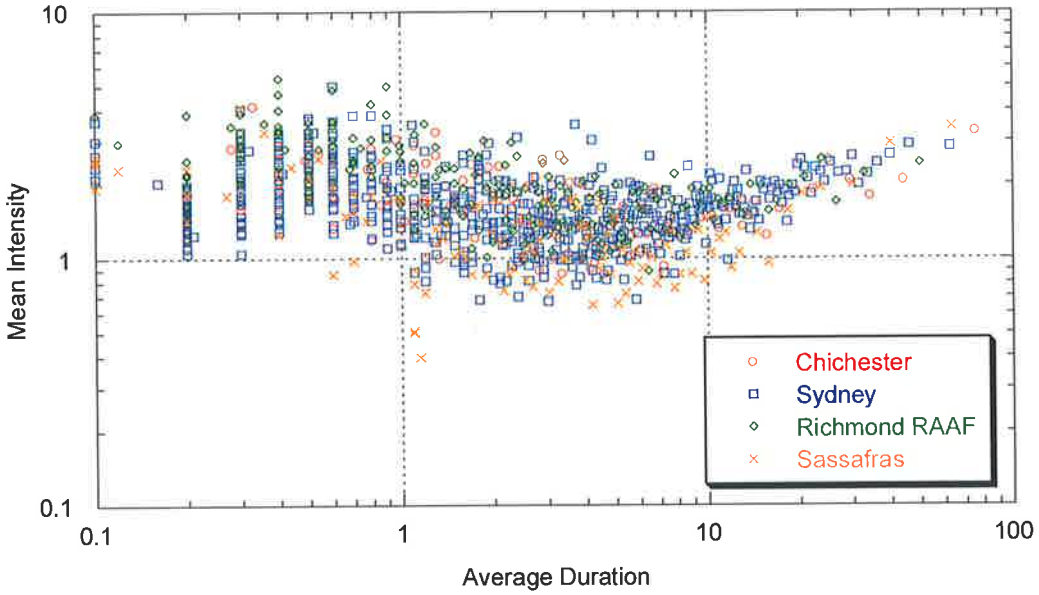

Figure 3.38: Comparison of Mean Intensity v Duration for New South Wales Rainfall Sites

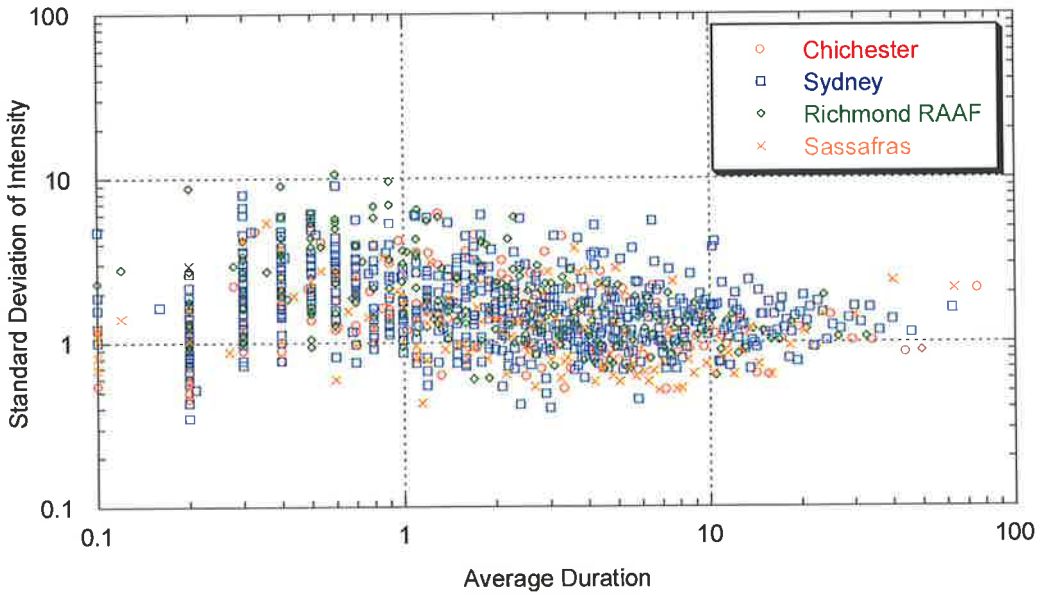

Figure 3.39: Comparison of Standard Deviation of Intensity v Duration at New South Wales Rainfall Sites

This observed consistency in results was encouraging and suggested that a similar breakpoint structure across all sites could be used to describe the conditional relationship. Significantly, this would remove the need to determine individual breakpoints for a given site therefore removing the need for modeller input.

Initial investigations into the use of such a blanket approach to the problem (whereby the number and locations of breakpoints were set and incorporated into the model removing the need for an iterative scheme), were not successful as the complex nature of the relationship and the model structure made it difficult for accurate calibration. Nine standard breakpoints were set up after analysis of numerous calibration plots and from user experience with their locations fixed at 0.2, 0.3, 1.0, 1.8, 2.4, 3.0, 9.0, 12.0 and 24.0 hours. Defining breakpoints at these locations provides sufficient structure to the piece wise linear relationship to capture the required variability however, attempts at numerous sites with this approach produced poor calibrations. While the calibration did estimate a set of model parameters, they were not able to locate the optimal set. Resultant simulations provided poor comparisons between observed and simulated statistics, significantly poorer than the original iterative manual calibration approach. This unfortunate outcome was suspected to be a result of the large search space introduced by the number of parameters (or the degrees of freedom) being calibrated concurrently (this is not an issue in the original model which employed an iterative process) and by the range of possible values each parameter could undertake. This was in direct contrast to the original iterative procedure which, though time consuming and requiring manual input, provided the chance to incrementally increase the number of breakpoint locations and by using an interpolation of the previous best fit also provided a significantly smaller search space at each step for the calibration process to manage.

Even though the blanket approach had not been an immediate success, it was obvious that the use of such an approach to the calibration of intensity had significant merit. Not only would it remove the need for an iterative process, it would also remove the need for modeller input reducing a major source of error and inconsistent results. However without additional modification to the calibration procedure it was evident that a blanket style approach could not be employed. With this in mind the possibility of developing some form of function to describe the conditional relationship warranted investigation. The consistent nature of the calibration plots (as discussed above) provided the ideal opportunity to introduce a continuous function to describe the required shape at all sites. This function

description could then be used to provide an initial estimate of the parameters within the standardised breakpoint piece-wise linear model, increasing the efficiency and robustness of the calibration process. There was also the possibility that development of such a continuous function may remove the need for a piece wise linear model, providing an alternate solution to the intensity calibration process.

3.4.2 Initial Duration – Intensity Relationship Description using a Continuous Function

To provide an initial mathematical description of the intensity – duration relationship, a continuous function was developed which contained enough variability to describe the observed intensity variations while still having significant constraint to ensure model parameters could be found quickly when using a global searching algorithm. While the observed shape of the intensity – duration relationship is complex, it was hoped that a function could be developed which described the structure adequately to provide an initial fit to the piece wise linear model in the worse case or in the best case provided an alternate option for the modelling of storm event intensity.

Initially considering data around Sydney, analysis of the average mean calibration plot suggest that the overall shape could be approximated by the combination of a constant or linear trend across the range of storm durations, and a curved deviation below this trend beginning at approximately 0.7 hours, and returning at 30 hours (Figure 3.40). This has been shown graphically in Figure 3.41. A similar description is present in Figure 3.42 for average standard deviation calibration plot, where deviation away from the linear trend begins earlier (0.3 hours) returns at around 3 hours and sits above the general linear trend.

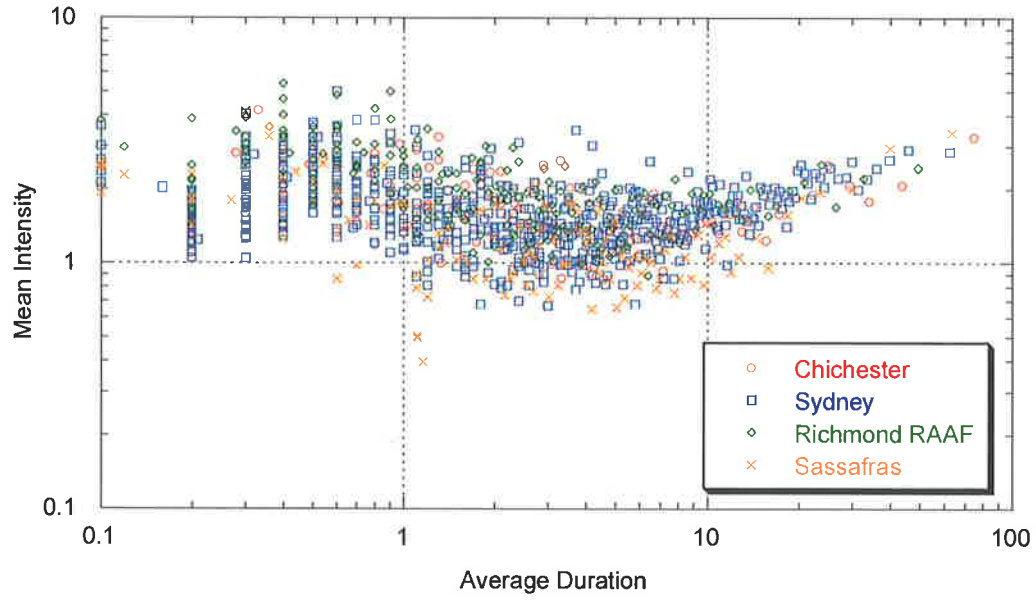

Figure 3.40: Comparison of Mean Intensity v Duration for New South Wales Rainfall Sites

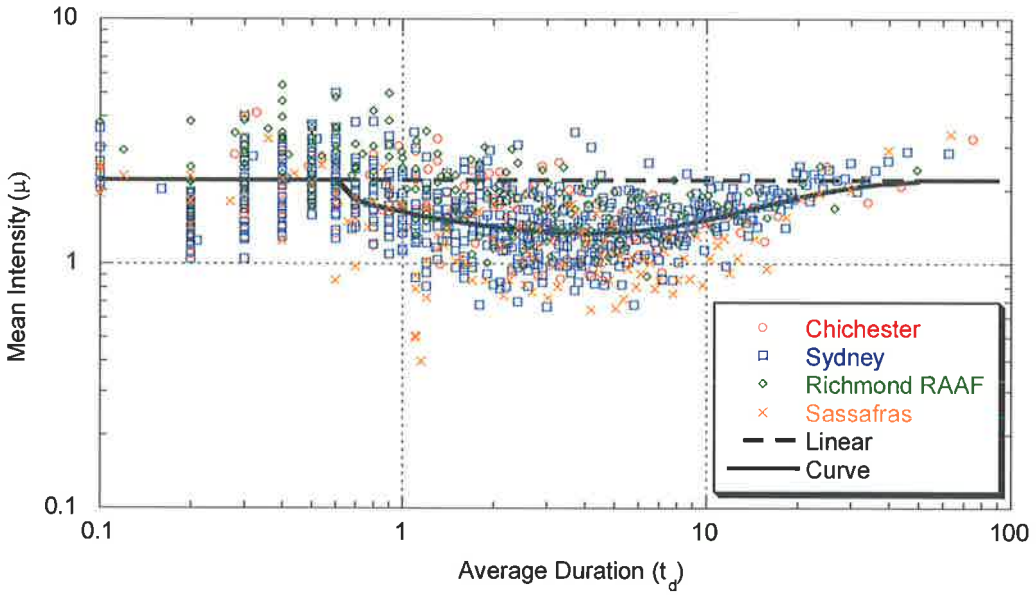

Figure 3.41: Schematic of Average Mean Intensity Calibration Plot Trend

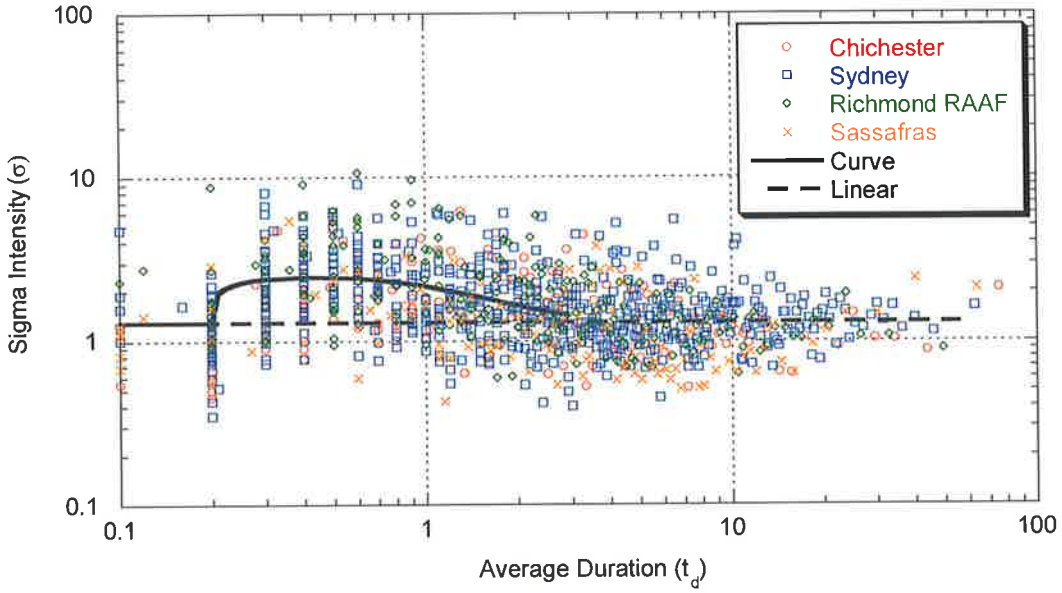

Figure 3.42: Schematic of Standard Deviation of Intensity Calibration Plot Trend

In order to provide an approximation of this shape, a hybrid function was developed which combined a simple linear trend with an additional function incorporated to describe the shape of the deviation away from the general trend. Describing this deviation is difficult in a consistent manner across numerous sites due to local variations. In some instances, as is the case for Sydney data, the deviation begins around 0.7 hours (for average intensity) and returns at 30 hours. Other sites this deviation can be more or less pronounced and occur at different times over different periods. In addition the adopted function must have the ability to adapt to the different shapes evident between the mean and the standard deviation calibration plots. After investigating a number of functions, the shape of the Beta distribution was chosen. This distribution is given by

$$p(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}(x)^{\alpha-1}e^{\left(\frac{-x}{\beta}\right)} \qquad (3.16)$$

where

$\frac{1}{\beta^{\alpha}\Gamma(\alpha)}$ is a scaling value to ensure the integral of the density equals 1, x relates to the storm duration, and α and β are the parameters of the Beta distribution.

As only the shape of the distribution is required, the scaling value of the Beta function can be disregarded, leaving

$$p(x) = (x)^{\alpha-1} e^{\left(\frac{-x}{\beta}\right)} \tag{3.17}$$

In this form the Beta function has the ability to form almost any shape imaginable to describe the deviation in the intensity – duration relationship. The interactions between α and β control the location of the deviation peak, the length of the decay, and to a lesser extent the size. A further control on the size of this deviation was provided by including a scaling value to the size of the deviation.

Combining this scaled Beta distribution with a generic linear equation, the final form of the continuous function used to describe the conditional intensity - duration relationship is given by

$$f(t_d) = \theta_1 + \theta_2 t_d - \theta_3 \left(t_d^{\theta_4 - 1} e^{\left(\frac{-t_d}{\theta_5}\right)} \right) \tag{3.18}$$

t_d denotes the storm duration and $f(t_d)$ provides an estimate of the average storm intensity. θ_i are the parameters of the hybrid function, with θ_1 and θ_2 providing the linear trend, and the deviation shape controlled by θ_1, θ_2 and θ_3.

The calibration of this function to Sydney data can be seen in Figure 3.43. The continuous function is able to provide a reasonable approximation of the conditional relationship, while being simple enough to enable a quick calibration. It is also obvious from this plot that the continuous function, as could be expected, smooths some of the local variability that occurs in the data.

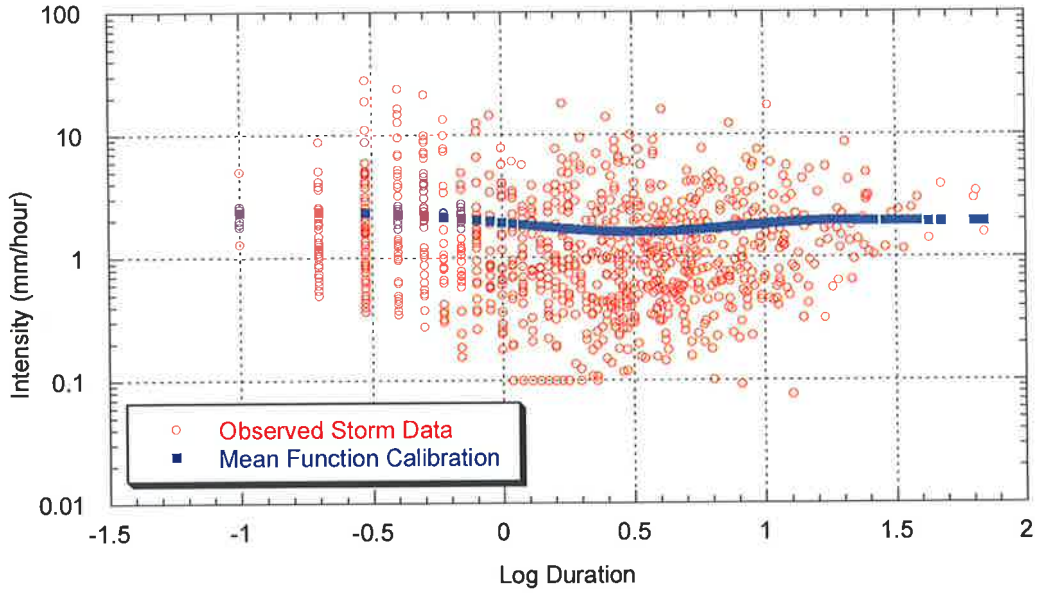

Figure 3.43: Example Hybrid Continuous Function Calibration (Data from Sydney, January)

A similar result is seen in Figure 3.44 for data from Melbourne where the piece-wise linear fit calibrated using the continuous function as a starting point is included for comparison.

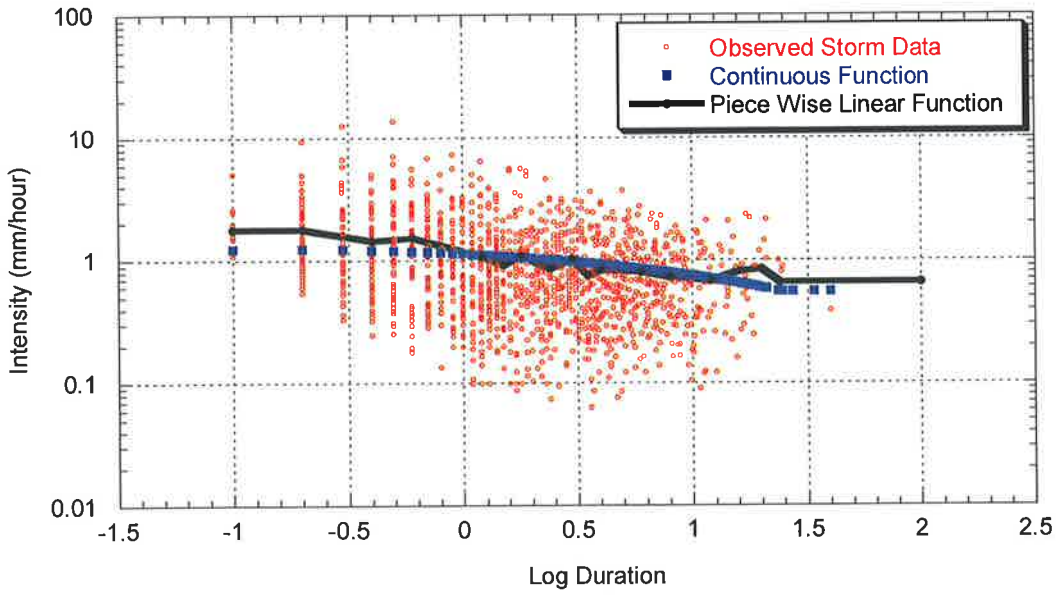

Figure 3.44: Comparison between Beta Continuous Function and Automatic Piece-Wise Linear Fit (Data from Melbourne, August)

The results presented in Figure 3.44 indicate that the piece wise linear fit captures the localised variations in the duration-intensity relationship more accurately in contrast to the continuous function on its own. Comparison between calculated likelihood values for the piece wise linear fit and the continuous function using optimal parameter values confirms this result. Table 3.1 and Table 3.2 present the likelihood values for Melbourne and Sydney Data respectively. In these tables the smaller the value, the better the fit to the observed data.

Table 3.1: Likelihood Function Comparisons for the Intensity Calibration Options (Melbourne)

MONTH	CONTINUOUS FUNCTION FIT	AUTOMATIC PIECE WISE FIT
January	1094.7	1068.1
February	1052.3	1015.7
March	1185.1	1161.8
April	1441.2	1413.8
May	1485.9	1449.8
June	1304.0	1283.6
July	1435.6	1381.7
August	1726.2	1673.8
September	1618.3	1577.7
October	1603.2	1572.5
November	1444.8	1426.4
December	1298.3	1274.7

Table 3.2: Likelihood Function Comparisons for the Intensity Calibration Options (Sydney)

MONTH	CONTINUOUS FUNCTION FIT	AUTOMATIC PIECE WISE FIT
January	1692.7	1656.4
February	1801.4	1778.5
March	2041.4	2019.6
April	1683.9	1653.7
May	1575.1	1556.6
June	1485.4	1459.6
July	1017.8	970.8
August	1100.3	1072.3
September	1228.9	1198.8
October	1479.8	1445.9
November	1644.3	1618.9
December	1634.2	1615.4

As shown in these typical results, using the continuous function calibration as a starting point for the automatic piece-wise linear model produces a better description of the conditional relationship than simply using the hybrid function on its own. This can be seen graphically in Figure 3.45, where the cumulative improvement to the model calibration (improvement to likelihood) over the year is displayed.

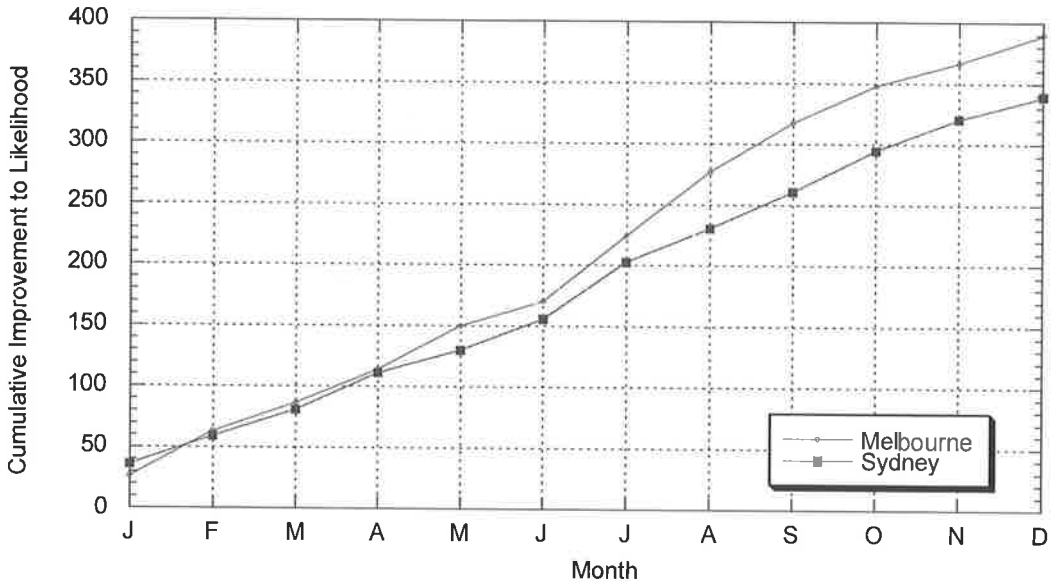

Figure 3.45: Cumulative Likelihood Improvement

While the continuous hybrid function is not able to describe the variations within the calibration shape adequately on its own, it is an extremely useful pre-calibration step within the calibration process. The continuous function provides both an initial starting point for each parameter within the parameter search space (which is close to the optimal value) and decreases the size of the search space by providing realistic limits on parameter values around these initial parameter estimates. The SCE search routine takes these initial parameter values and is able to refine them quickly within the limited search space to determine the optimal parameter set.

3.4.3 Verification by Simulation

Comparisons between observed and simulated data sets were undertaken at various sites to ensure that the new calibration technique were able to reproduce the same observed statistics and distributions as the previously used manual calibration technique. While the new calibration routine solves the issue of manual modeler input and improves the efficiency of the search routine, it should also perform to the same standard as the previously used manual approach if it is to be adopted as an alternative.

Results from the manual calibration model showed good reproduction of both aggregated rainfall totals and Intensity – Frequency – Duration curves. If the new automated calibration was shown to also reproduce these values, its validity as a calibration tool could be verified. Extensive validation of the improved model is presented in Chapter 5 incorporating all improvements however Figure 3.46 and Figure 3.47 display typical reproductions of monthly rainfall distributions using the automated calibration procedure. These plots indicate good agreement between observed and simulated values suggesting there is no reduction in the ability of the model to reproduce these non-calibrated statistics.

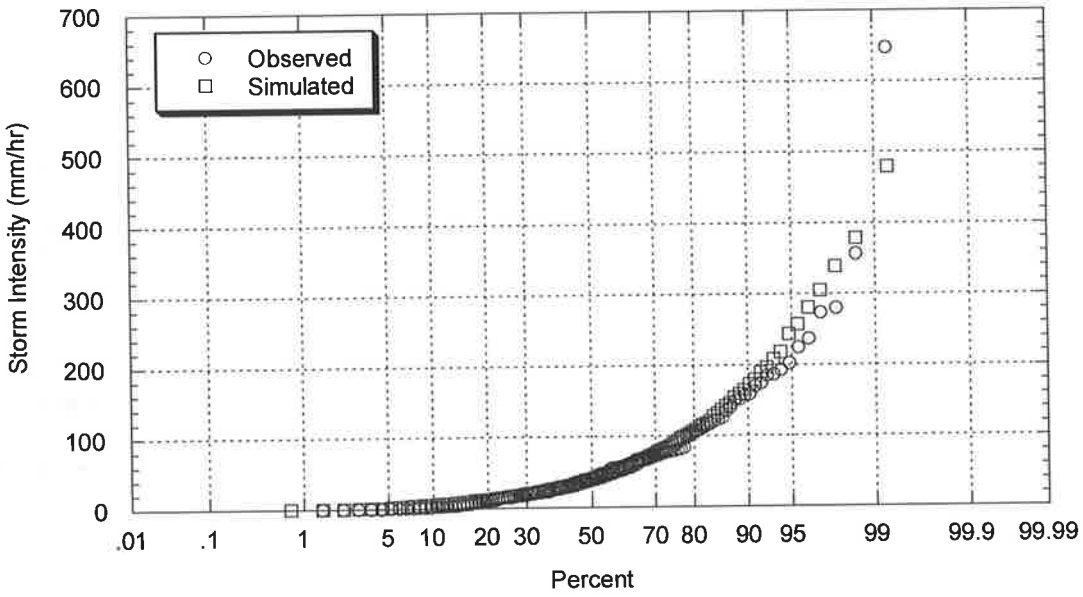

Figure 3.46: Comparison of Observed and Simulated Rainfall Intensity Values for Brisbane Data, July

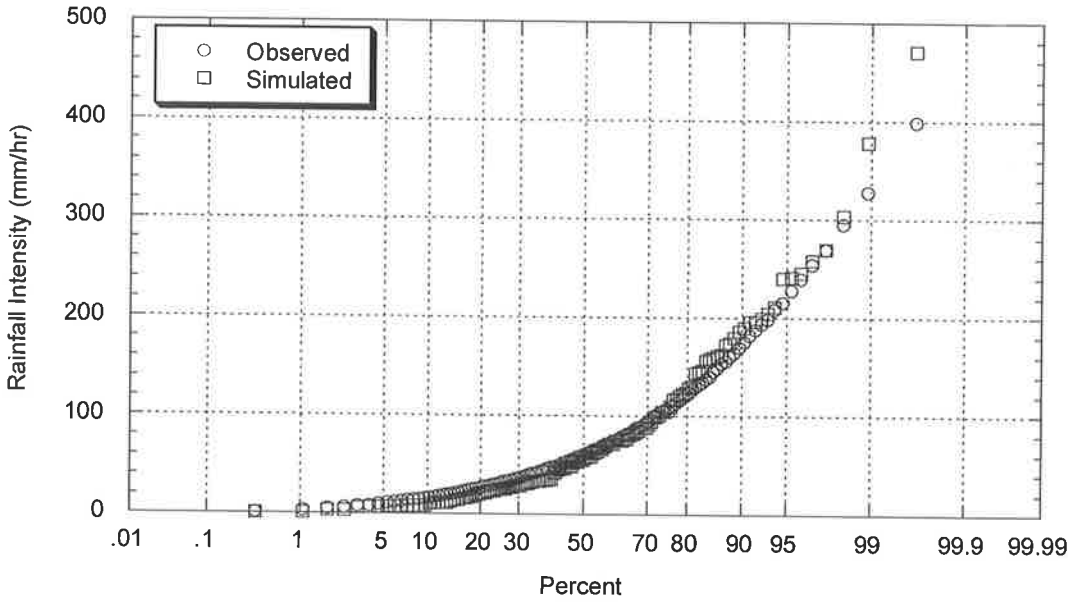

Figure 3.47: Comparison of Observed and Simulated Rainfall Intensity Values for Sydney Data, August

An additional check using Intensity – Frequency – Duration curves is provided in Figure 3.48 and Figure 3.49 below. Again, the automatic calibration model reproduces these statistics across various time scales with the advantage that there is no manual intervention in the calibration process.

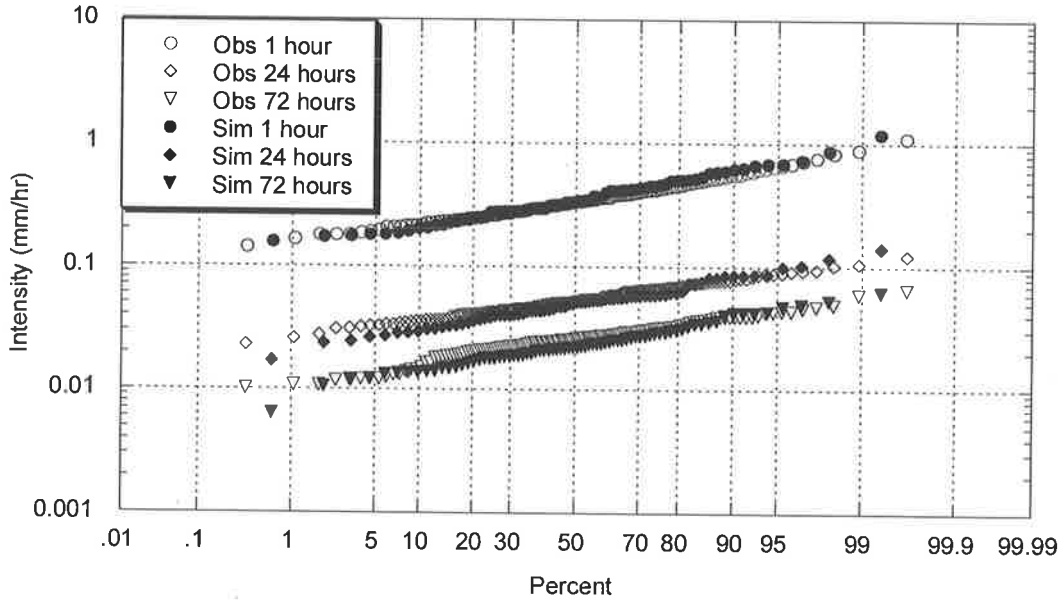

Figure 3.48: Intensity Frequency Duration Comparison for Various Durations, Sydney

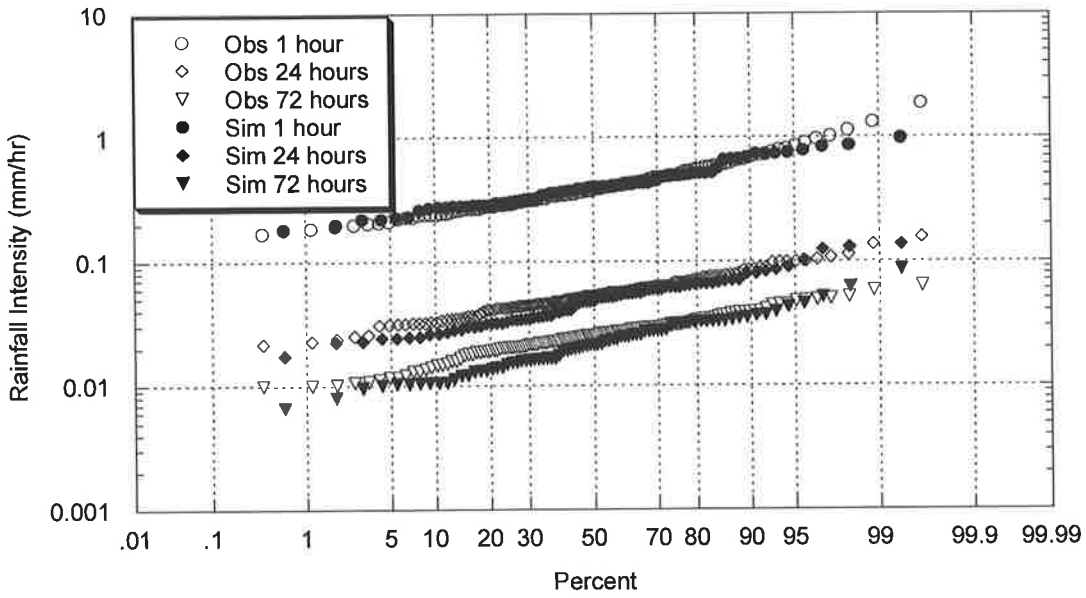

Figure 3.49: Intensity Frequency Duration Comparison for Various Durations, Brisbane

3.5 Summary

The existing rainfall model presented by Heneker *et al.* (2001) and selected for further development as part of this study contained a number of deficiencies in its original form which warranted attention prior to the development of a regionalisation approach for applying the model at sites with little or no calibration data.

Analysis of calibrated inter-event time and storm duration parameter distributions from the original model setup of Heneker *et al.* (2001) showed significant correlations were present which provided identification problems when searching for optimal model parameters. The incorporation of the Metropolis algorithm led to an investigation into the inter-parameter relationships and enabled one parameter variable being replaced with a constant numerical value. Checks between the altered and original model have shown this simplified description of the storm event distributions with one less parameter per month are as accurate as the original model. This has increased the confidence and stability in the maximum likelihood parameter estimations, which provides a benefit for future regionalisation work. A similar investigation into the calibration of conditional storm intensity has indicated that the

existing model does not contain significant parameter correlations and in its present form calibrated parameters are well defined.

Improvements to the calibration process for storm intensity were investigated and analysis across numerous sites indicate that a similar shape exists within the conditional intensity – duration calibration relationship that is well described by using a piece-wise linear relationship which has a constant set of breakpoints across all sites. This significantly removes the need for the manual selection of breakpoints. The automatic calibration process has been completed by using a hybrid continuous function to describe the conditional relationship and provide initial parameter estimates at the set breakpoint positions. The adoption of constant breakpoints and the inclusion of initial parameter estimates through the use of a hybrid function have ensured optimum parameter values are quickly and easily determined with no manual intervention.

Validation of the changes to the calibration procedure was undertaken through a comparison of simulated and observed statistics with selected plots presented here and a more detailed analysis of the overall model with all updates provided in Chapter 5. Favourable comparisons between observed and simulated values demonstrate the ability of the new intensity calibration to accurately describe the observed relationships, improving the robustness of this section of the rainfall model.

In order for the model to be used within any regionalisation framework it was important that improvements were made without decreasing the accuracy of the model itself. This has been achieved and the incorporation of the automatic calibration of event intensity and the removal of highly correlated superfluous parameters has provided a rainfall model that is efficient, well defined and easy to use. The remaining deficiency in the model structure warranting attention is the treatment of uncertainty. Understanding and defining uncertainty is important when considering the performance of the regionalised rainfall model and techniques which have been developed and incorporated into the model to describe uncertainty are presented in the next chapter.

CHAPTER 4

INCLUSION OF PARAMETER UNCERTAINTY IN SIMULATED TIME SERIES

4.1 Introduction

The development of any advanced rainfall model reliant on the estimation of model parameters from limited observed data should incorporate some description and or treatment of uncertainty. Uncertainty in model parameter estimation can arise due to the length of the observed rainfall data available for calibration or as a result of sampling variability associated with the descriptive statistical distributions used within the structure of the model. It is unreasonable to expect a model calibrated to 5 years of data to perform as well as a model calibrated to 100 years. In addition, sampling uncertainty associated with the probability descriptions that define the model can be a very real source of concern, particularly when comparing simulated and observed results. As the model uses a random number generator during simulation, two realisations of the model using identical model parameters can be significantly different if the sequence of random numbers is not the same. Providing a means to take this sampling uncertainty into account improves the validity of the reported simulation statistics for comparison, not only at sites which contain a long historical record for comparison, but perhaps more importantly when comparing model results after regionalisation.

In addition to describing the potential sources of model uncertainty, there is a need to understanding any influence of this uncertainty within the probabilistic model to ensure model results and limitations are also understood. In order to provide a more complete hydrological model and an accurate comparison between observed and historical data statistics, (particularly relevant when investigating model

regionalisation performance) an investigation into the influence of parameter uncertainty when simulating rainfall time series has been incorporated into the Heneker *et al.* (2001) model.

Parameter uncertainty is an important consideration in models of any type but is particularly important for stochastic models. Consider the calibration of a model to two data sets, one set containing 5 calibration data points and the other 100,000. It is reasonable to assume that 5 data points would be insufficient to describe the variation in the input data to the same degree of accuracy as 100,000 points. Extending this, a simulation model calibrated to 100,000 data points should result in better defined model parameters and consequently a model more likely to replicate the observed data statistics than one calibrated to 5 points. If parameter uncertainty is not treated explicitly, then care must be taken in assuming the same confidence in the two model outputs. To investigate this potential influence within the rainfall model, estimated posterior parameter distributions have been determined and used within the simulation routine. As these posterior distributions are influenced by the length and variability in the calibration data, the resultant simulation results are also affected. This provides an explicit measure of the confidence in the simulation outputs.

A multiple Monte Carlo simulation structure has been included in the simulation model to provide the ability to generate numerous realisations of the output data which take into account parameter uncertainty and sampling variability between model runs. The end result is a set of final results which can be presented with simulation limits, providing a superior comparison between observed and simulated statistics. This ability is a valuable improvement to the original model and enables improved inferences to be made about the quality of the resultant simulations and their ability to represent the observed data. When considered within the application of a regionalisation procedure, simulation bounds will provide an immediate indication of the certainty in the resultant simulations which have been calibrated within the regionalisation framework. This is a valuable tool for assessing the quality of a regional calibration.

4.2 Review of Existing Models to Capture Uncertainty

The importance of treating uncertainty explicitly within a given model structure was emphasised by Katergaris and Hadgraft (1994). Their work was based on a simple flood frequency analysis and was able to show that due to parameter uncertainty, the width of the confidence intervals after analysis increases by 15-30%. This lack of confidence in model output is important when model results are used in any hydrological or hydraulic system analysis. Their idea was further generalised by Chaubey *et al.* (1999) who suggest that the treatment of parameter uncertainty is a necessity in any complex model. During their study into a water quality model, they were able to show that model results were influenced by the variability in the observed data. Without capturing this uncertainty explicitly, the accuracy of the resultant simulation models is questionable.

In its original form the Heneker *et al.* (2001) rainfall model includes a treatment of data uncertainty associated with the binned nature of historical rainfall observations. As the historical data is recorded only every six minutes, the exact duration of inter-event times and storm durations are not known exactly. An adjustment to the likelihood calculation (Lambert and Kuczera (1998)) was included which took this into account during the calibration procedure, providing an improved estimate of the optimum parameter values. However, the extent of parameter uncertainty remaining after calibration was not explicitly defined or previously taken into account. In addition to this, simulation results from the model have typically been presented as a single realisation with the optimum parameter values (Heneker *et al.* (2001), Heneker (2002)). What these works have failed to consider is the issue of sampling variability when comparing observed statistics against model simulations.

A significant contribution in the area of parameter uncertainty was presented by Kuczera and Parent (1998). Their discussion centred on conceptual catchment models and the use of Monte Carlo Markov Chains and in particular the Metropolis algorithm to accurately estimate the parameter posterior distribution. Once this posterior distribution has been estimated, parameter uncertainty can be taken into account through numerous simulation replicates that use samples from the posterior

parameter distributions. These simulations can then be collected and provide simulation confidence limits on model outputs/statistics.

The use of MCMC's for estimating parameter uncertainty was also used by Guame *et al.* (1998) within an urban stormwater quality model and in a different application by Park *et al.* (2002) for a model estimating the location and amounts of chemical pollution sources.

While their work was focussed on catchment models, the simulation algorithm presented by Kuczera and Parent (1998) was adapted and applied to the Heneker *et al.* (2001) rainfall model. This allowed the generation of 90% simulation limits and model outputs which take into account the uncertainty in model parameters. The adapted algorithm is described below.

1. Randomly sample rainfall model parameters $\theta_1 \ldots \theta_n$ from their corresponding posterior distributions $p(\theta_j | y)$.
2. Undertake a simulation using this set of parameters $(\theta_1 \ldots \theta_n)$
3. If the number of simulations is less than the number of replicates N required, then return to step 1 to re-sample a parameter set and re-simulate an additional realisation of the model.
4. For each comparison statistic required (annual rainfall, inter-event times etc), rank the N simulated sequences and extract the $(100-\alpha)$ and α percentiles to obtain the $(100-2\alpha)$% simulation limits.

While this algorithm is sufficiently generic to take into account any model structure, efficient sampling of the posterior distributions is still required and in complex models this is not a trivial task. While sampling from the posterior distribution had previously been undertaken by Van Straten and Keesman (1991) (using Monte Carlo set-membership) and by Beven and Binley (1992) (using a GLUE approach), Kuczera and Parent (1998) note that these approaches belong to the family of importance sampling algorithms and are restricted in their application due to the massive computing resources required to characterise a highly dimensioned parameter space. This was particularly important for rainfall models (including this model) where a large number of parameters are used. As a result Kuczera and Parent (1998) presented the Metropolis algorithm as an alternative to provide efficient

sampling from $p(\theta_i|y)$. As discussed in Chapter 3, the Metropolis algorithm was incorporated into the rainfall model to investigate parameter correlations and required only slight adaptation for use during this section of the research.

4.3 Incorporating Parameter Uncertainty into the Rainfall Model

To incorporate parameter uncertainty within the rainfall model, an estimate of the model's posterior parameter distributions was required. This followed a similar process to that used in Chapter 3 where a normal model calibration is undertaken and determines the optimal parameter values for a given set of observed rainfall data. These optimal parameter values were then used as the initial seed locations within the parameter space for use with the Metropolis algorithm. The Metropolis algorithm then proceeds through a 'warming up' process which ensures that the sampling distribution within the algorithm approaches that of the "real" posterior parameter distribution. Once this warming up process is completed, subsequent samples from the sampling distribution can be assumed to be equivalent to samples from the posterior parameter distribution. At this stage samples can be taken from the posterior parameter distributions for each model parameter of interest and used to generate a simulated realisation of the model. For the purposes of this research the number of replicates required (and as a result the number of parameter samples) has been set to 100. Each parameter set then generates simulated model results with the complete sample of results analysed and used to generate simulation bounds (generally 90% simulation limits within which 90% of the simulation results lie).

A major advantage of using the Metropolis algorithm to provide efficient samples from the posterior model distributions is the ability to determine the influence on the resultant simulations from uncertainty associated with the dry spell (inter-event), storm duration and storm intensity parameters. By comparing simulations that include uncertainty associated with one model variable at a time (i.e. storm duration), the influence on the resultant simulation as a result of this uncertainty specifically can be investigated. If storm duration is taken as an example, it could be reasonably expected that the inclusion of parameter uncertainty associated with the storm

duration parameters would influence the resultant storm duration statistics. Intuitively it is less clear as to whether the uncertainty associated with the storm duration parameters would have a noticeable difference on the average storm depth or perhaps even on inter-event times? There is also the chance that parameter uncertainty has no influence on the resultant simulation confidence suggesting that the model could be calibrated to very small historical records without a noticeable decrease in accuracy.

In order to investigate the relative influence of parameter uncertainty, a number of 100 year simulations were undertaken while including the various uncertainty options. This provided the opportunity to compare the relative influence of each storm event parameter set on the resultant model outputs. In addition to investigating this parameter uncertainty associated with the storm event parameters, the influence of sampling variability associated with changes to the random number sequence used during the simulation were also included. This provided an understanding of the sampling uncertainty within the model and was determined through changes to the seed parameter within the random number generator. In Section 4.3.1 and 4.3.2 the influence of uncertainty associated with the bulk storm parameters is investigated for models calibrated to long historical records. This provides an insight into the relative influence of each bulk storm parameter on the resultant model simulations. Further work provided in Sections 4.4 and 4.5 investigates the influence of the length of available calibrating data on the resultant model simulations.

4.3.1 Influence of Intensity Parameter Uncertainty on Rainfall Model Simulations

The posterior parameter distributions estimated through the use of the Metropolis algorithm were used to determine the influence of including uncertainty on the rainfall intensity model parameters. These parameter distributions describe the distributions of storm intensity model parameters at each breakpoint within the piece wise linear structure which has been used to describe the conditional intensity – duration relationship. It was expected that the influence of uncertainty associated with the calibration of the storm intensity parameters would be minimal on the

aggregation statistics in the model results. This was due to the relative influence that the distribution of storm intensity has on the simulation statistics, particularly at a large time scale.

To understand why the uncertainty associated with the simulation of storm intensity should only have a minimal influence on the resultant model simulations; one has to compare the relative influences each section of the model has on the model outputs. For example it is understood that the simulation of inter-event times influences the simulation process by determining the number of rain events and their distribution throughout the year. Similarly the simulation of storm duration directly influences the length of rain events and provides a lesser influence on the number and distribution of these events throughout the year. (A lesser influence due to its relative average length in comparison to the average inter event times) In addition, the simulation of storm durations influences the storm intensity as a result of the conditional intensity - duration relationship within the model. In contrast, uncertainty associated with storm intensity only influences the rain intensity/depth for a given storm. As a result slight changes to the storm intensity parameters will have a small influence on aggregation statistics at short time scales with the influence proportionally decreasing as the aggregation level is increased. (i.e. a + or - 5% change to storm intensities throughout the year has a minimal impact on the resultant annual rainfall in contrast to a + or - 5% change to the inter event times which could result in a significant decrease in the number of storm events within a year)

To investigate the influence of uncertainty associated with the simulation of storm intensity, simulations were undertaken with two different configurations within the uncertainty model. The first configuration ignores the potential uncertainty associated with storm intensity. This provides a base simulation which is only influenced by the sampling variability associated with the use of the random number generator within the probabilistic model. A second configuration was also used which includes the uncertainty associated with storm intensity and is also still influenced by the sampling variability associated with the probabilistic model. If the influence of uncertainty associated with storm intensity is minimal (as expected), the difference between these two simulations should be minimal. Figure 4.1 presents the

comparison between the two model configurations for Perth data. The resultant difference between these model outputs is approximately 5mm per year, suggesting the influence of additional uncertainty associated with intensity parameters on the reproduction of annual rainfall is small.

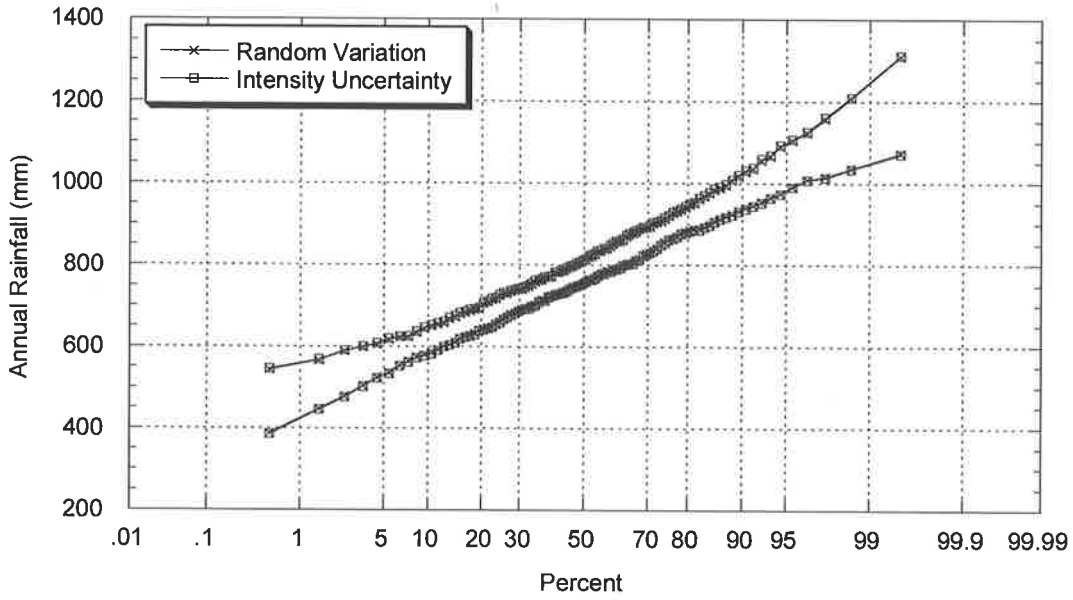

Figure 4.1: Comparison between Intensity Uncertainty and Random Variation, Annual Rainfall (Calibrated with Perth Rainfall Data)

At aggregation levels over a much shorter timescale a similar result is observed. Figure 4.2 displays the average monthly rainfall for data from Perth. This also indicates minimal influence on the model results as a consequence of intensity uncertainty.

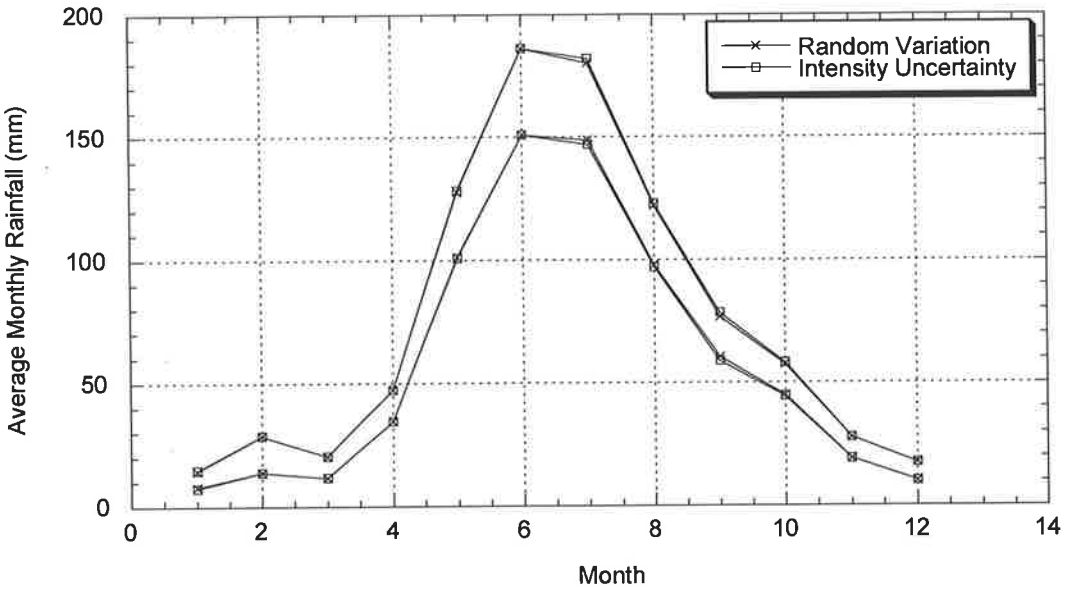

Figure 4.2: Comparison between Intensity Uncertainty and Random Variation, Monthly Rainfall (Calibrated with Perth Rainfall Data)

While intensity uncertainty has little effect on simulated rainfall values, it was expected that uncertainty associated with storm duration would have a far greater influence. This is due to the relative influence of storm duration (2-5 hours) in comparison to storm intensity (1-3 mm) which are used to generate storm depth. This influence is investigated in the following section.

4.3.2 Influence of Inter-Event Time and Storm Duration Uncertainty on Rainfall Model Simulations

To investigate the influence of uncertainty associated with the inter-event time and storm duration model parameters, the Metropolis algorithm was again used to estimate the posterior parameter distributions. These were then used in various combinations to assess the relative influence of each parameter set on the model output. Again, for each option 100 simulation realisations were run and the results used to generate simulation bounds describing the influence of each model section on the resultant output.

Figure 4.3 presents the 90% simulation bounds for the mean inter-event time for each month given the inclusion of various uncertainty options. As was the case

when investigating the uncertainty associated with storm intensity, the sampling variability resulting from the use of a random number sequence within the probabilistic model is included and named "Random Number Uncertainty". This relates to the simulation bounds generated as a result of the probabilistic structure of the model and will be observed in all simulation results. The "Inter-event uncertainty" relates to the simulation output when parameter uncertainty associated with the inter-event times is considered. Similarly "Storm Duration Uncertainty" presents the results when parameter uncertainty for storm duration parameters is included, and finally the "All Uncertainty" option is presented to display how these individual options combine to produce a complete simulation outcome.

It can be seen from Figure 4.3 that the upper bound "All Uncertainty" result for month 2 drops slightly below all other uncertainty configurations for the month of February. While initially seen as peculiar, this outcome is a result of the sampling uncertainty within the model. As each configuration run of the model results in a new random number sequence (which is used to generate the mean inter event times, storm durations and storm intensities), each run has the opportunity to be influenced slightly differently by any associated sampling uncertainty. In this case this results in a slightly lower mean inter event time for the "All Uncertainty" option in comparison to the individual uncertainty configuration runs.

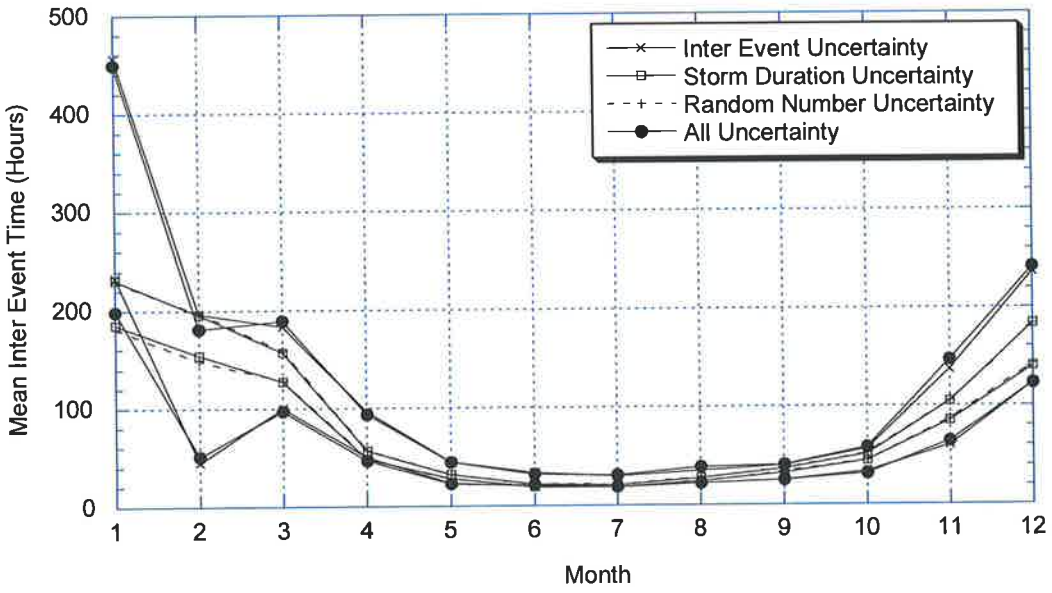

Figure 4.3: 90% Simulation limits for the mean Inter-Event time due to various Uncertainty Options (Calibrated with Perth Rainfall Data)

To provide an improved visualisation of the results, the distance between the simulation limits were calculated and extracted into column format for presentation below (Figure 4.4).

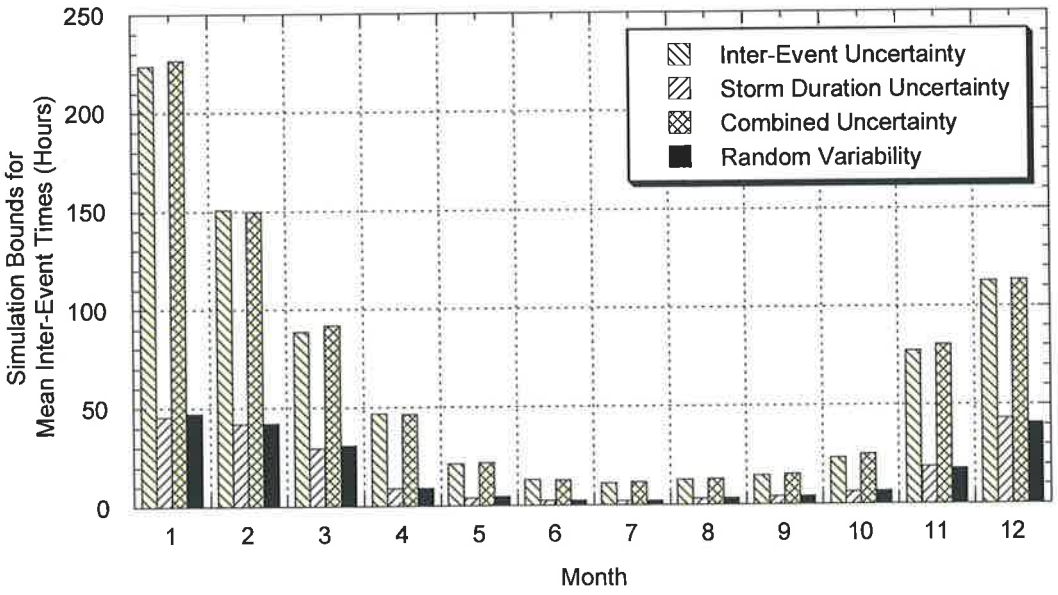

Figure 4.4: Comparison of 90% simulation limit ranges for mean Inter-Event time due to various Uncertainty Options (Calibrated with Perth Rainfall Data)

It is clear from Figure 4.4 that the inclusion of parameter uncertainty associated with inter-event times has a strong influence on the resultant simulation of inter-event times. As expected the uncertainty associated with inter-event times has influenced the simulation of inter-event times and hence the model output, but this influence is not consistent across the year. It is also not likely to be consistent between different groups of years, particularly in Australia which undergoes periods of drought associated with climatic events such as El Nino.

Another expected outcome from the results in Figure 4.4 is that the influence of uncertainty associated with storm duration on the simulation of inter-event times is negligible. The simulation bounds generated when simulating with storm duration uncertainty are the same as that for the model run when only sampling variability was included. This was an expected outcome because the certainty (or lack of) in the calibration of one model parameter (storm duration) has no influence on the calibration certainty of another (inter-event times) for models of this type which are calibrated to independent storm events.

It is also important to note that the influence on the resultant simulation limits generated by the sampling variability within the model provides a pseudo lower bound in terms of uncertainty. If the distance between simulation limits is close to that of the random sampling equivalent, then the influence on that statistic of including parameter uncertainty for that parameter is minimal.

A similar result can be observed when the influence of parameter uncertainty on the simulation of storm duration is presented below in Figure 4.5. Again the greatest influence on the 90% simulation limits was provided by uncertainty associated with the variable of interest, in this case storm duration. The inclusion of uncertainty associated with inter-event time has had no impact on the resultant simulation of storm duration.

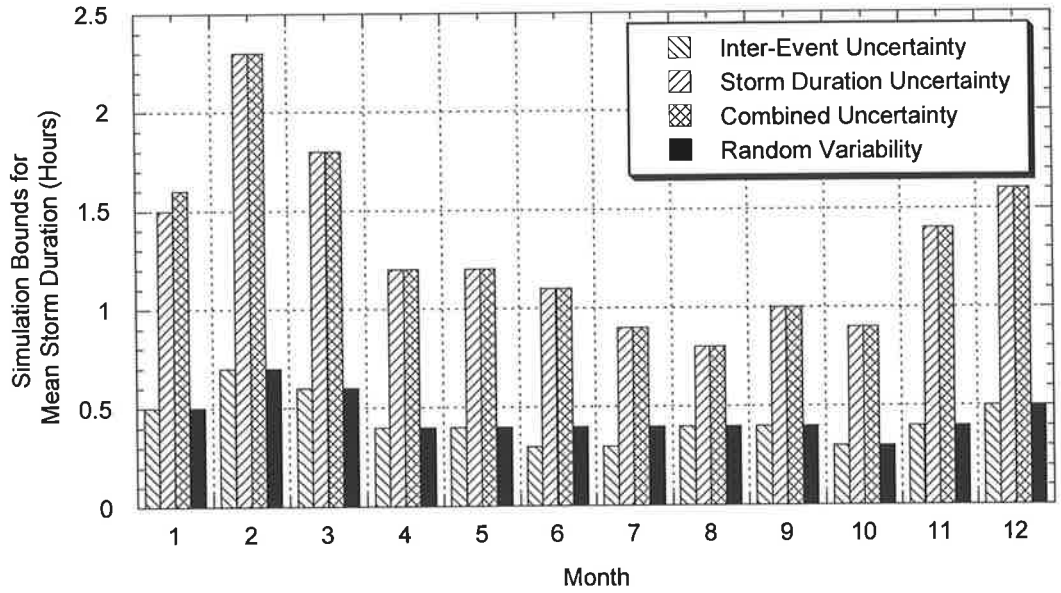

Figure 4.5: Comparison of 90% simulation limit ranges for mean Storm Durations due to various Uncertainty Options (Calibrated with Perth Rainfall Data)

The results presented above give further evidence that without including parameter uncertainty explicitly, model results can be misleading. Even with the use of a long historical record, there is an element of uncertainty associated with the stochastic parameters in the model. The incorporation of the metropolis algorithm and the ability to produce multiple realisations for the purposes of generating 90% simulation limits ensures this uncertainty is quantifiable in the model.

4.4 Influence of Record Length on Posterior Parameter Distributions

Having established that parameter uncertainty does indeed have an influence on model simulations, the question then becomes what influence does the length of the historical record have on the accuracy of the model and the resultant 90% simulation limits? Section 4.2 described how an increase in uncertainty should result as the calibration length is decreased. It would be reasonable to expect simulations incorporating uncertainty resulting from a calibration to shorter historical records would produce larger 90% simulation limits due to the increased variability in the calibrated parameters. To investigate the issue of record length, the model was

calibrated to three different lengths of historical record namely the full length, half length and a short 5 year record. A comparison could now be made between the simulation results from each length of calibration record. With a large number of sites across Australia containing 5 years of historical record, if it could be shown that calibration to this short 5 year record did not increase the model uncertainty significantly, then there was the possibility that a regionalisation technique would not be required.

The incorporation of the Metropolis Algorithm into the rainfall model provides the ability for efficient sampling from the posterior parameter distributions $p(\theta_j|y)$. Once a large number of samples have been generated, a histogram can be formed indicating the shape of this posterior distribution. If the algorithm has been successfully integrated into the model and is working well, then as the length of the historical data available for calibration is increased, the associated certainty in the model parameters should also increase. When analysing the shape of the resultant posterior distribution, an increase in certainty is evidenced by a posterior histogram that has a smaller spread and a much higher peak than a corresponding histogram from the same model calibrated to a shorter historical record. In extreme cases, increasing uncertainty will result in a significant decrease in the confidence of the estimation of the optimum parameter value, and as a consequence the accuracy of model output is questionable.

An investigation into the influence of record length on the resultant posterior parameter distribution has been undertaken with data from Perth. Figure 4.6 shows the resultant parameter distribution given a calibration to the full available record length (42 years) for Perth while Figure 4.7 presents the result from a 5 year subset of the full length record.

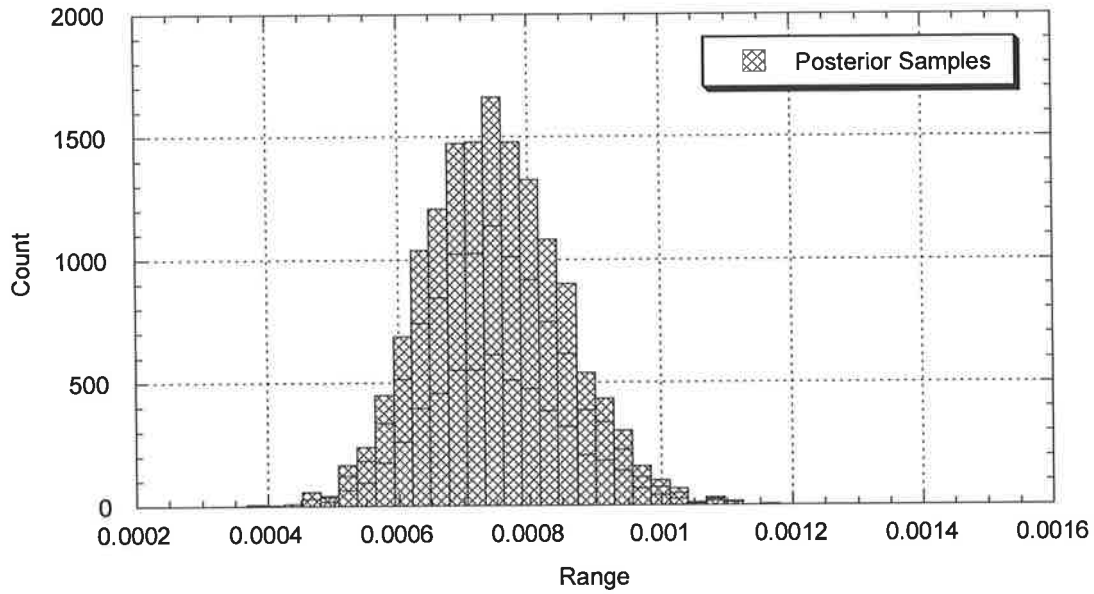

Figure 4.6: Posterior Distribution Dry Spell Constant Parameter θ_1, Full Length Calibration Data (Data from Perth, January)

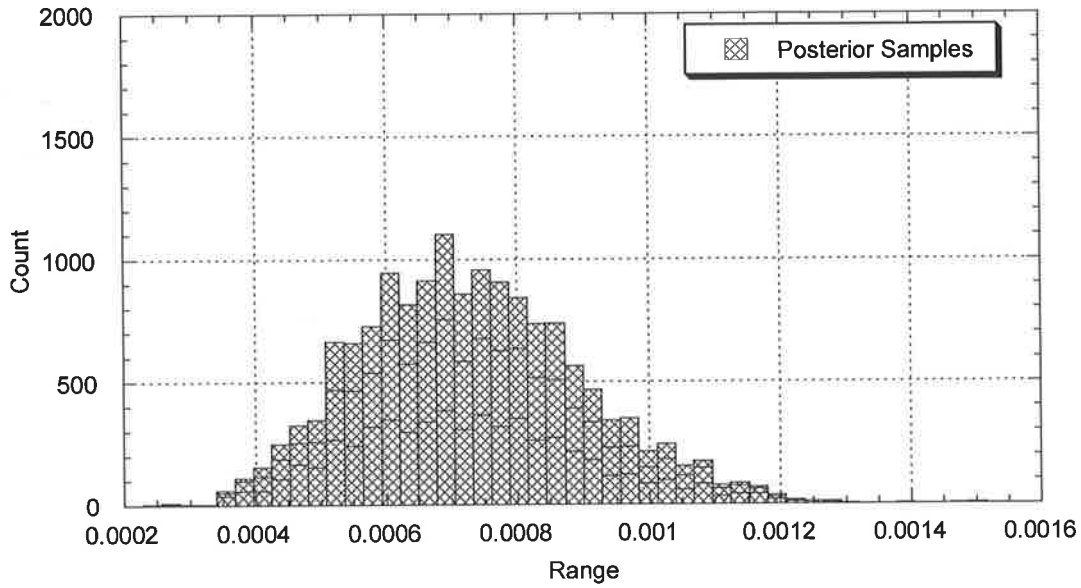

Figure 4.7: Posterior Distribution Dry Spell Constant Parameter θ_1, Short Length Calibration Data (Data from Perth, January)

It is clear that the uncertainty associated with this particular parameter (the 'constant' parameter in the description of inter event times with an optimum value of 0.00075) increases as the calibration data length decreases. This is evidenced by the increase in the breadth of the histogram and a smaller peak value, a typical result for all model

parameters. The optimum parameter value of 0.00075 has successfully been located with the full length record and there is a regular histogram shape and a definite peak at this optimum value. For the model calibrated with 5 years of data, the histogram is flat, with little or no definition at the peak. The optimum value determined via the search routine was estimated at 0.0007 but it is evident from the histogram that a range of possible values between 0.0006 – 0.000775 provide almost the same level of fit. The parameter distribution is not well defined and as a result the model calibrated to only 5 years of record will produce vastly different results to that model calibrated to the full length record. Not only is the histogram flat increasing the uncertainty in the calibration which in turn will increase the resultant simulation bounds, but the optimal parameter value estimated via a maximum likelihood approach is actually a different value when comparing the two models. In this case the optimum parameter is known to be 0.00075 and as the model calibrated to the short record has not been able to successfully determine this value, the simulated distribution of inter-event times generated by the short calibration model will not replicate the observed statistics of the full length historical record.

Approximately 20 years or half of the available historical rainfall data is sufficient to obtain the optimal inter-event parameters for Perth as can be seen in Figure 4.8. Even in this case however where the optimal parameter value has been identified, there is significant additional uncertainty associated with this calibration as evidenced by the flatter histogram in comparison to the parameters calibrated to the full length record.

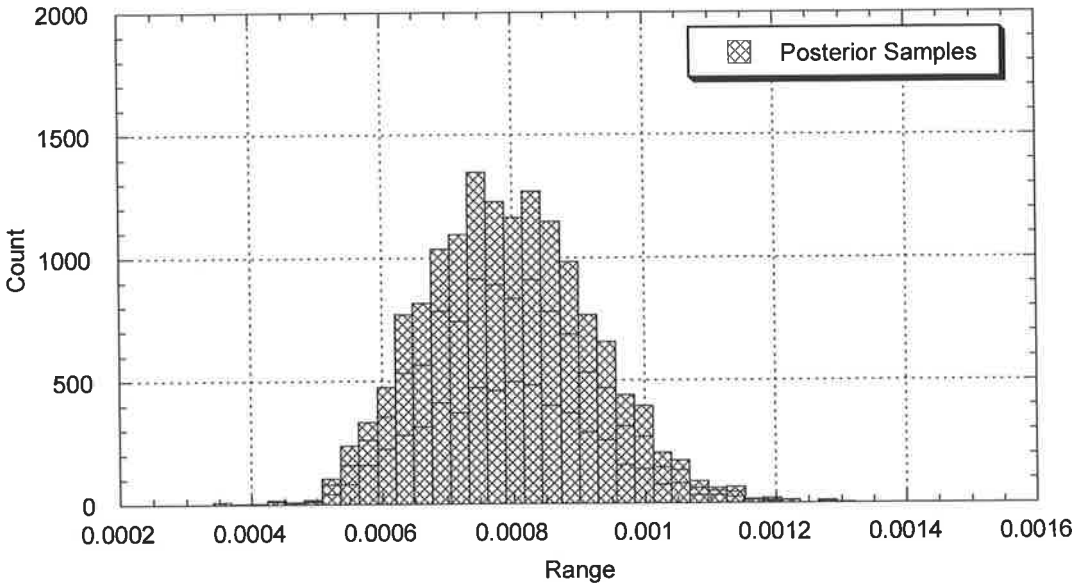

Figure 4.8: Posterior Distribution Dry Spell Constant Parameter θ_1, Half Length Calibration Data (Data from Perth, January)

The implication of this result can be understood if we consider the application of the model to Australian rainfall data. As discussed previously, there are a significant number of sites with short term historical records. The parameter distributions above show that direct calibration to these short term data sets is problematic at best. It would be reasonable to assume given these results that a direct model calibration would require a data record of approximately 30 years. Given this is the case, in order to apply the model to shorter historical records the development of a modified calibration or robust regionalisation technique is required to allow the use of these short data sets for model calibration. While regionalisation approaches are discussed in detail in Chapters 6 and 7, it is useful to understand these restrictions on applying the model directly to shorter record lengths. The influence this has on the resultant model simulations is discussed in the next section.

4.5 Influence of Record Length on Resultant Model Simulations

In order to understand the implications of calibrating with short historical records and the resultant impact on the accuracy of the model, the rainfall model was

calibrated to both the full length record and the 5 year subset. Once calibrated, simulation results were then generated to gain an insight into the resultant influence of record length on the model outputs.

Figure 4.9 compares the simulation limits for mean inter-event times between the model calibrated to the full-length record and that calibrated to the 5-year subset. It is obvious that there is an increase in uncertainty associated with calibration to the short 5 year record. This is particularly the case for the Summer-Autumn period where the number of storms and hence data points from the data are at a minimum. The significant size of the resultant simulation bounds suggests the accuracy of the resultant simulation must be under question.

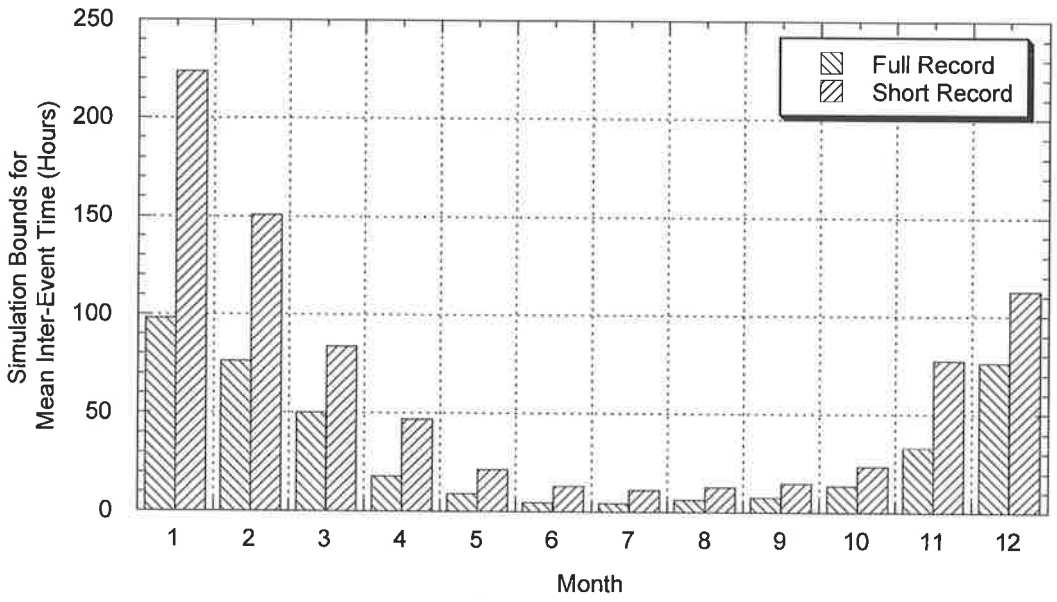

Figure 4.9: Comparison of simulation limit ranges for mean Inter-Event times due to different Calibration Record Lengths (Calibrated with Perth Rainfall Data)

A similar result can be seen in Figure 4.10 when applied to mean storm durations. Again the lack of storm events available for calibration is evidenced by the large differences through Summer-Autumn.

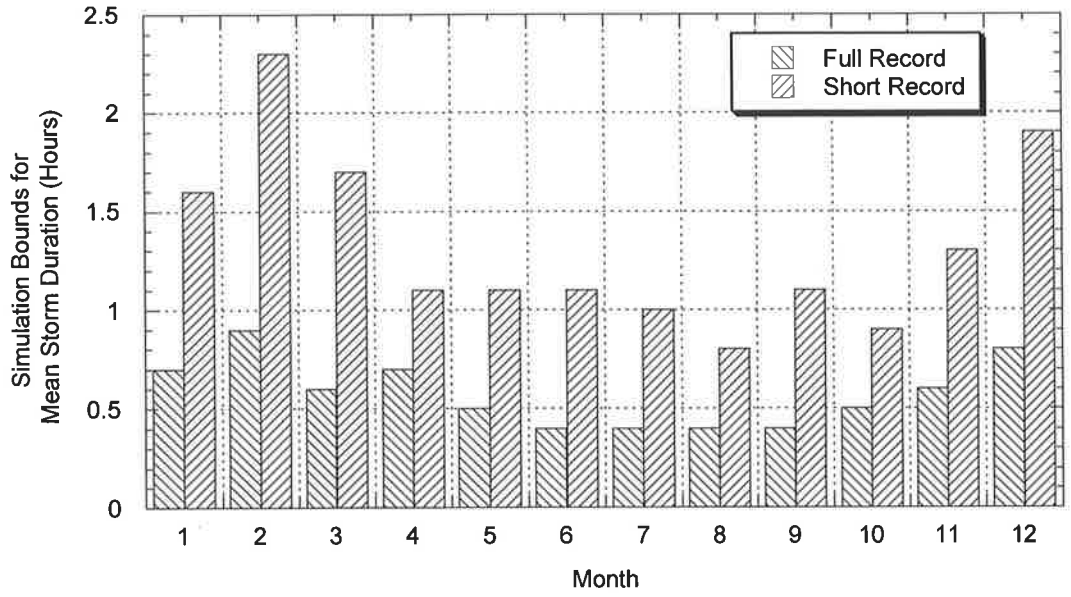

Figure 4.10: Comparison of simulation limit ranges for mean Storm Durations due to different Calibration Record Lengths (Calibrated with Perth Rainfall Data)

The decrease in calibration data was also expected to increase the 90% simulation limits on aggregated statistics and result in a less accurate simulation. Figure 4.11 presents the aggregated annual rainfall simulation comparison between a model calibrated to the full length data record and one calibrated to 5 years of data. As expected the model calibrated to the short record has not reproduced the observed data with all observed data points residing outside of these 90% limits. In comparison, the model calibrated to the full record length has successfully captured the annual rainfall distribution.

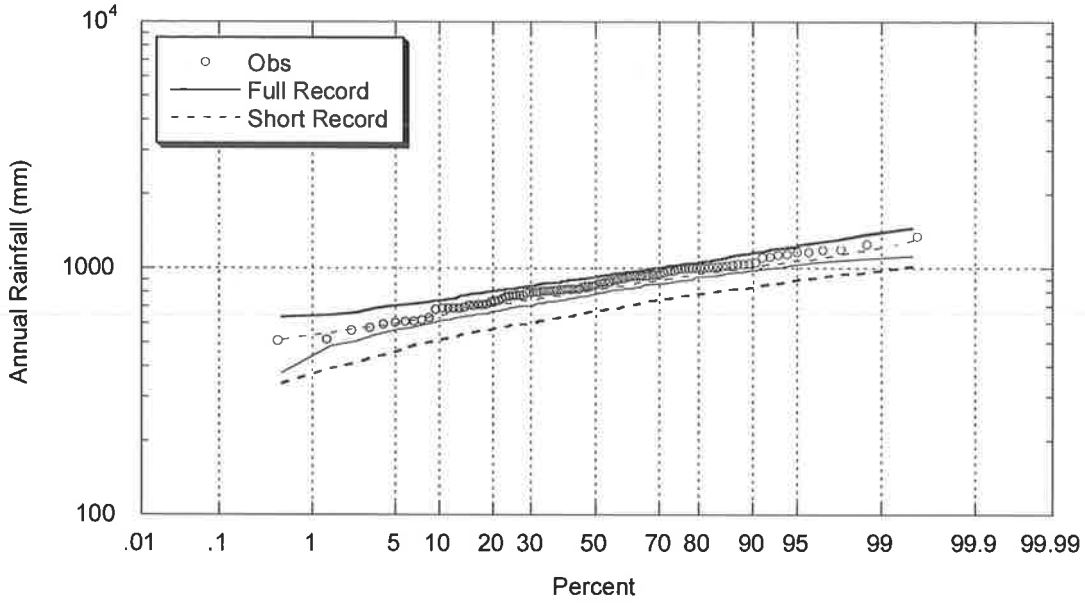

Figure 4.11: Comparison of Annual Rainfall limits due to Full and Short Record Lengths and Incorporating All Uncertainty Options (Calibrated with Perth Rainfall Data)

The inability of the model to successfully reproduce the observed statistics with a short historical record is a clear indication that the model in its original form required significant data records to ensure accurate calibration and as a result limits its application. It follows that a technique to enable calibration to small data sets would be of great benefit and was a major incentive to develop a regionalisation model.

In order to graphically represent the additional influence of record length on the uncertainty 90% simulation limits, annual simulation results were adjusted so that the results from the different length calibration records (full, half and short record lengths) were similar in respect to their simulation medians. In this way the aggregated rainfall from each simulation could be plotted on the same graph and a direct comparison of the spread of these 90% simulation limits undertaken.

Figure 4.12 and Figure 4.13 present the influence of data length on the resultant simulation of adjusted monthly and annual rainfall respectively. Again included with the annual rainfall is the sampling variability associated with the probabilistic nature of the rainfall model. This random sampling variability was produced by setting the model parameters constant at their optimal values ensuring the only source of variation was a result of this sampling variation. As was the case when investigating

the inter-event time and storm duration uncertainty, this sampling variability result can be thought of as the lower limit of uncertainty and is a result of the stochastic structure of the model. As expected, decreasing the length of calibration record has introduced more uncertainty and therefore greater simulation bounds.

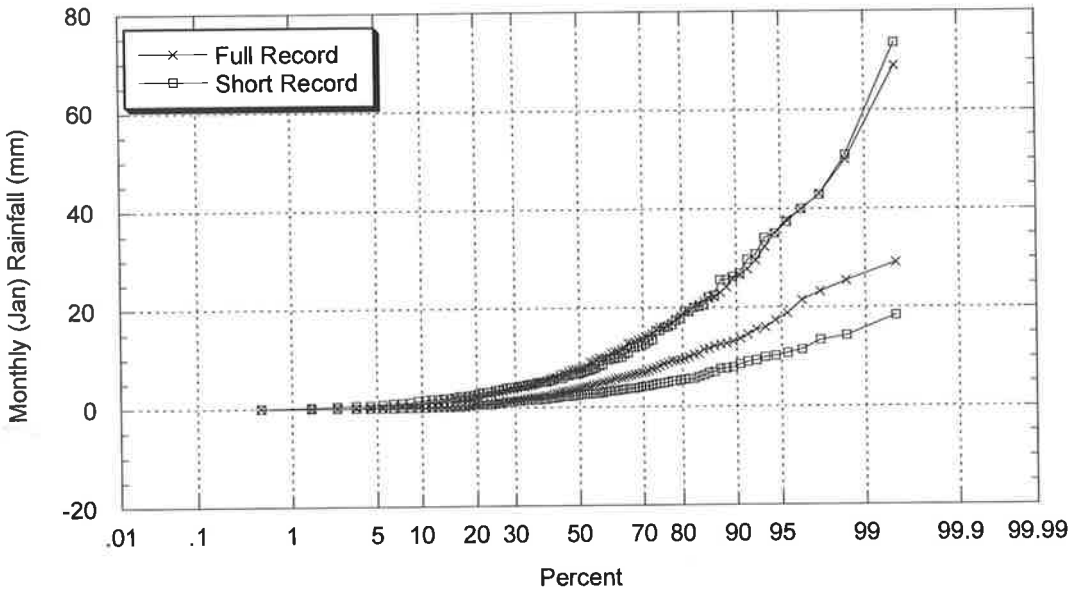

Figure 4.12: Comparison of January Rainfall limits due to different Calibration Record Lengths and Incorporating All Uncertainty Options (Calibrated with Perth Rainfall Data)

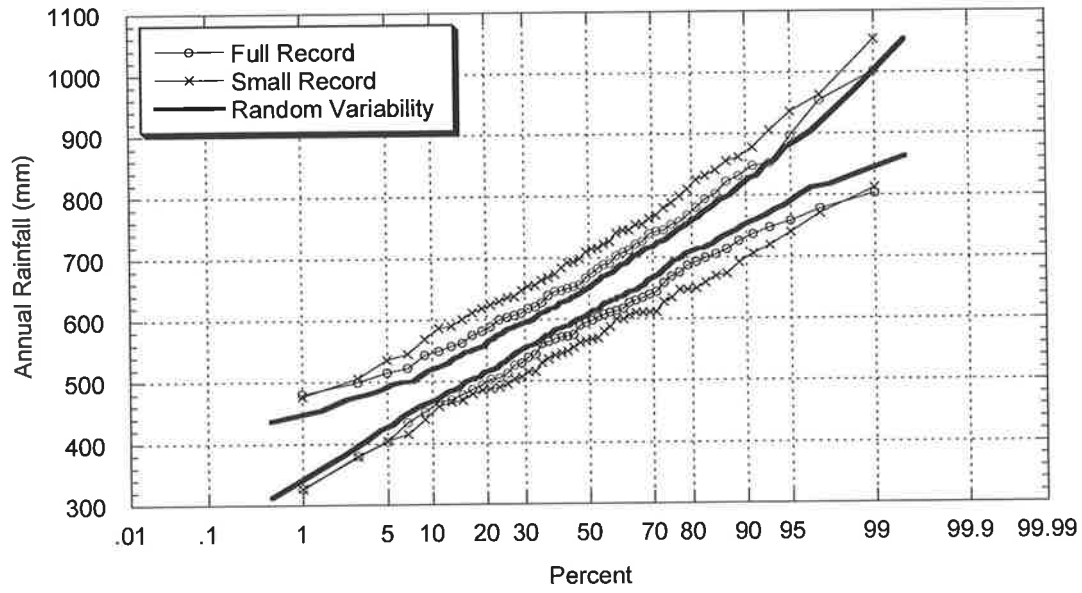

Figure 4.13: Comparison of Annual Rainfall limits due to different Calibration Record Lengths and Incorporating All Uncertainty Options (Calibrated with Perth Rainfall Data)

To gain insight into the influence of specific parameter uncertainty on annual rainfall statistics with decreasing record length, further results were produced with simulations considering only inter-event uncertainty (Figure 4.14) or storm duration uncertainty (Figure 4.15).

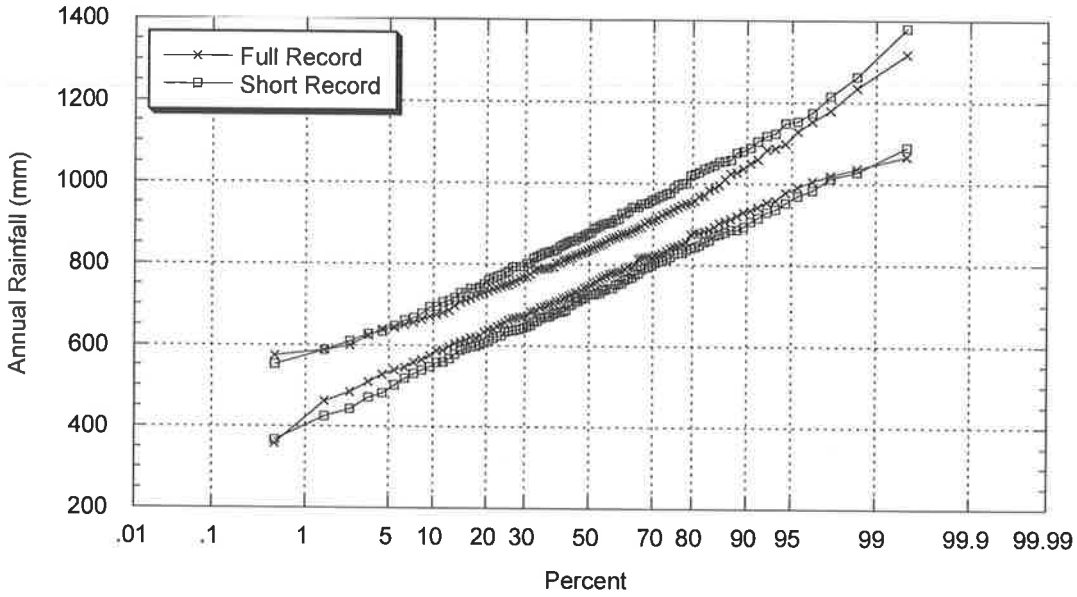

Figure 4.14: Comparison of Annual Rainfall limits due to different Calibration Record Lengths and Incorporating Inter-Event Uncertainty (Calibrated with Perth Rainfall Data)

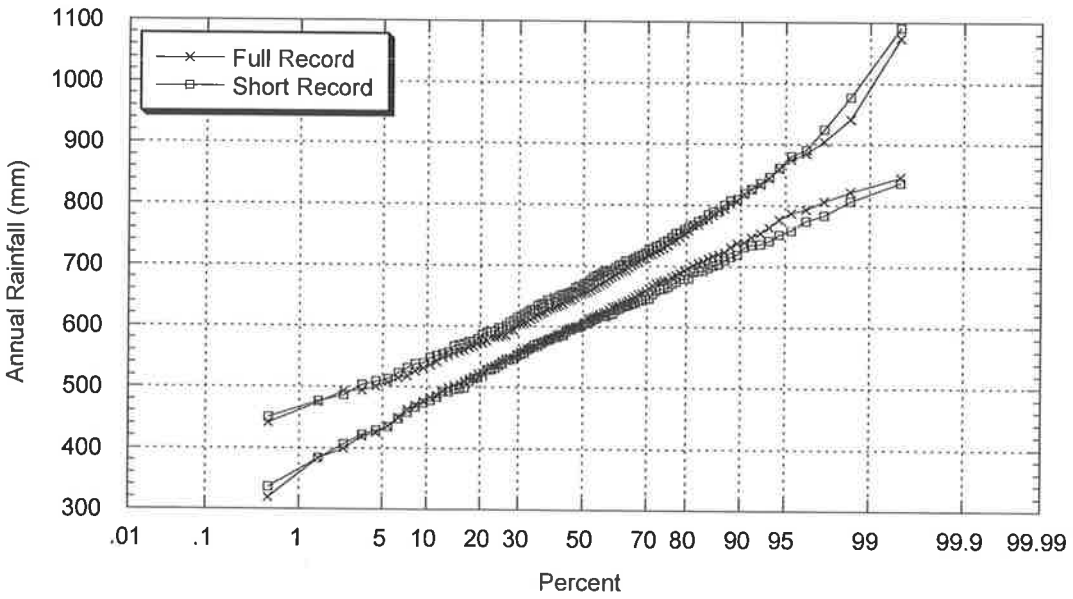

Figure 4.15: Comparison of Annual Rainfall limits due to various Calibration Record Lengths and Incorporating Storm Duration Uncertainty (Calibrated with Perth Rainfall Data)

A comparison of Figure 4.14 and Figure 4.15 shows a greater increase in the 90% simulation limits for inter-event times than for storm durations with a decrease in calibration record length. This result suggests uncertainty associated with the simulation of inter-event times has a greater influence on the annual rainfall than that for storm duration. To understand why this is the case, the relative average length of these events and the associated uncertainty must be considered. The uncertainty range associated with inter-event times is almost 100 times that of the corresponding value for storm duration. For example in January, the uncertainty associated with the inter-event times is in the order of 200 hours, whereas this drops to 1.6 hours for the corresponding January storm durations. Clearly there is a greater scope for a small change in the inter-event times to influence the rainfall totals for a given period and as a result this larger variability has a greater impact on the monthly and subsequently annual rainfall variability, increasing the annual rainfall limits. Importantly this suggests that in order to accurately model the aggregated rainfall distributions, the inter-event parameters must be well founded and clearly defined. As a result any regionalisation work must provide a model which captures the inter-event times to an adequate level of accuracy to enable reproduction of these aggregated statistics.

4.6 Summary

The treatment of parameter uncertainty is a key ingredient to accurate stochastic rainfall models. Variability in the length and quality of existing rainfall data ensures simulation results from models which do not explicitly incorporate uncertainty either through calibration or simulation must be viewed with extreme caution. The Metropolis algorithm has been incorporated into the existing rainfall model presented by Heneker *et al.* (2001), enabling the identification of calibrated parameter distributions. These calibrated distributions are influenced by the length and quality of the historic data set providing a direct treatment of parameter uncertainty. Including these parameter distributions and introducing a Monte Carlo structure to the simulation process has resulted in significantly improved simulation outputs. These outputs allow improved comparison between the extended model and

historical rainfall statistics, providing additional confidence in the ability of the model to reproduce calibrated and non-calibrated statistics.

Uncertainty associated with the calibration of inter-event times and storm durations provides the greatest influence on the resultant simulations within the model. In comparison only a minor influence is observed due to uncertainty associated with storm intensity. As expected, the simulation bounds expand with a decrease in calibration data, and the large uncertainty associated with very short calibration records introduces some doubt onto the validity of simulations that do not attempt to describe uncertainty.

The performance of the model in its current form when calibrated to short historical records is under question. The level of confidence in model calibration and the poor comparison to observed rainfall statistics suggest the model is incapable of direct calibration with short records. In order to use these numerous short pluviograph records available across Australia and provide confidence in the resultant model structure and simulations, a regionalisation approach is required. Before developing a regionalisation approach, complete validation of the model after the incorporation of the improvements discussed in the previous chapters was undertaken and is presented in the next chapter.

CHAPTER 5

IMPROVED MODEL VALIDATION

5.1 Introduction

This chapter presents a performance examination of the rainfall model originally developed by Heneker *et al.* (2001) and then improved upon during this study for its suitability for use in developing a regionalisation procedure to enable application with short pluviograph or daily calibration data. It is important to validate the performance of the model at this time to not only instill confidence in its potential use as a synthetic rainfall generator but also to ensure additional works undertaken as part of this study have not decreased the capabilities of the original model. Without a robust working model, there is little point in developing a regionalisation process to increase its application. Conversely, if the model is robust and able to reproduce observed rainfall statistics, then this provides an ideal starting point for increasing its application and usefulness through regionalisation.

Changes to the distributions describing inter-event times and storm durations in Chapter 3 and the new intensity calibration model presented in Chapter 4 have been incorporated into the stochastic rainfall model and require validation through simulation. Where possible the simulated results are presented with associated simulation limits, enabling a more rigorous comparison to observed values than has been possible previously.

An examination of the observed and simulated probability distributions for inter-event time, storm duration and average event depth are presented. While these probability distributions are used during the calibration and subsequent simulation process, this comparison provides a level of confirmation that the calibration was

successful and the observed distributions are being effectively reproduced by the simulation.

Intensity-Frequency-Duration (IFD) curves, aggregated statistics and the probability distribution of annual totals are also examined. These statistics were not used during the calibration process and provide further evidence of the credibility and structure of the model. In particular the reproduction of aggregated statistics in comparison to their daily counterparts (i.e. daily mean depth, daily dry probability) is investigated as these statistics could then be used during regionalisation with daily data. The influence of record length on the model calibration and resultant simulation was also examined.

5.1.1 Observed Data Records

To thoroughly test the improved rainfall model, significant historical records across a range of climates within Australia were used. The sites chosen for specific comparison and their historic data record length were Melbourne (95 years), Sydney (78 years), Brisbane (83 years), Perth (45 years) and Adelaide (30 years). The development of successful simulation models at these sites provide evidence that the model works well across a range of climates within Australia (given adequate calibration data), providing a basis for future regionalisation work.

To provide adequate data for comparison, each model was simulated at an equivalent length to that of the historical record and repeated for 100 simulation replicates. This provided the ability to produce simulation limits and provide an accurate comparison to the observed data records.

5.2 Calibrated Event Probability Distributions

A necessity of any model calibrated to historical storm event data is its ability to reproduce the distributions and statistics used during the calibration process. If this is not the case, the model can be assumed to be poorly formulated and serious questions must be asked of its validity as a hydrological tool. This section presents

comparisons between observed and predicted event probability distributions for inter-event times and storm duration as well as comparisons between simulated and observed monthly statistics for inter-event times, storm durations and event storm depth. As these distributions were used during the calibration process, it was expected that they would compare favourably. Monthly parameters have been used for all calibrations in this study with three (3) parameters calibrated for the distributions of inter-event time and storm duration events each month while nine (9) breakpoints at 0.2, 0.3, 1.0, 1.8, 2.4, 3.0, 9.0, 12.0 and 24.0 hours were used for the piece wise linear intensity model.

Figure 5.1 (Brisbane in February), Figure 5.2 (Sydney in May) and Figure 5.3 (Melbourne in September) show that the inter-event time distribution is well represented for data at various sites and months. (These plots are presented as exponential probability plots). There is good agreement between observed and predicted event times indicating a good calibration with historical data. (This is not a simulation output; it is a comparison between the observed event distribution and the predicted distribution from the calibrated model).

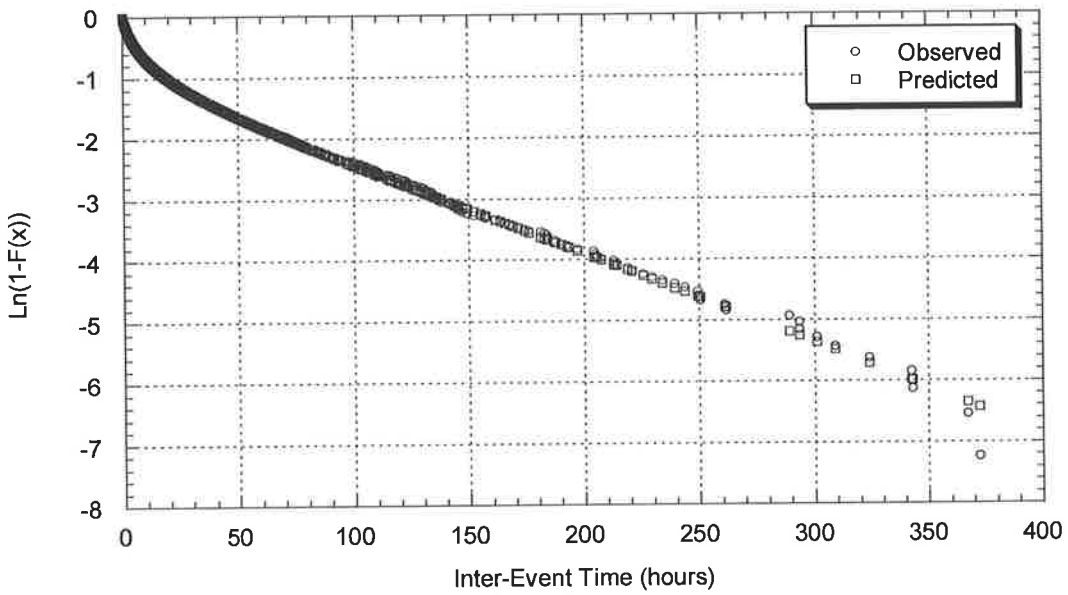

Figure 5.1: Observed and Predicted Inter-Event Distribution, Brisbane, February

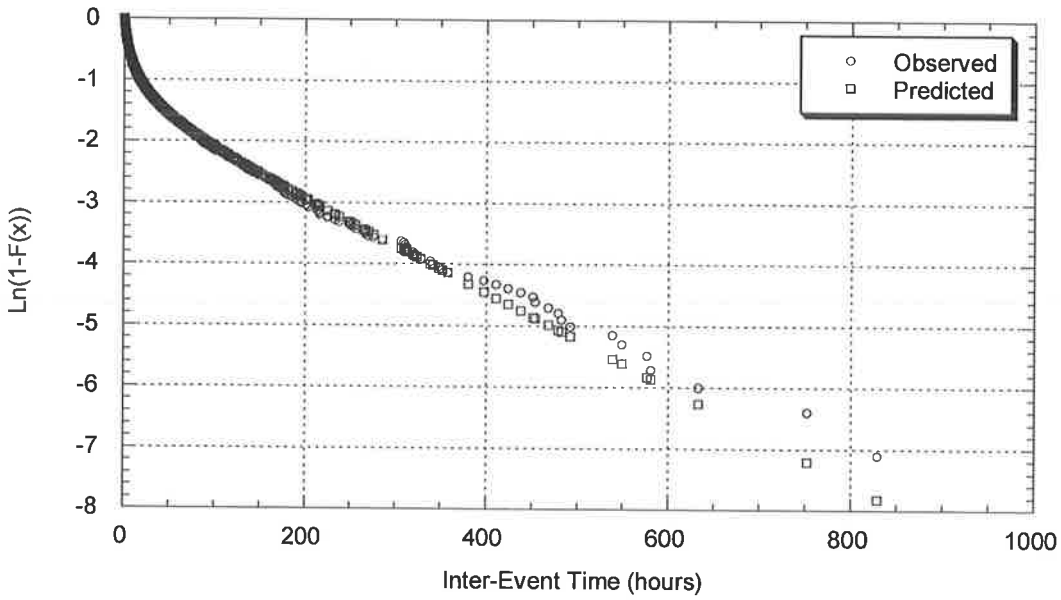

Figure 5.2: Observed and Predicted Inter-Event Distribution, Sydney, May

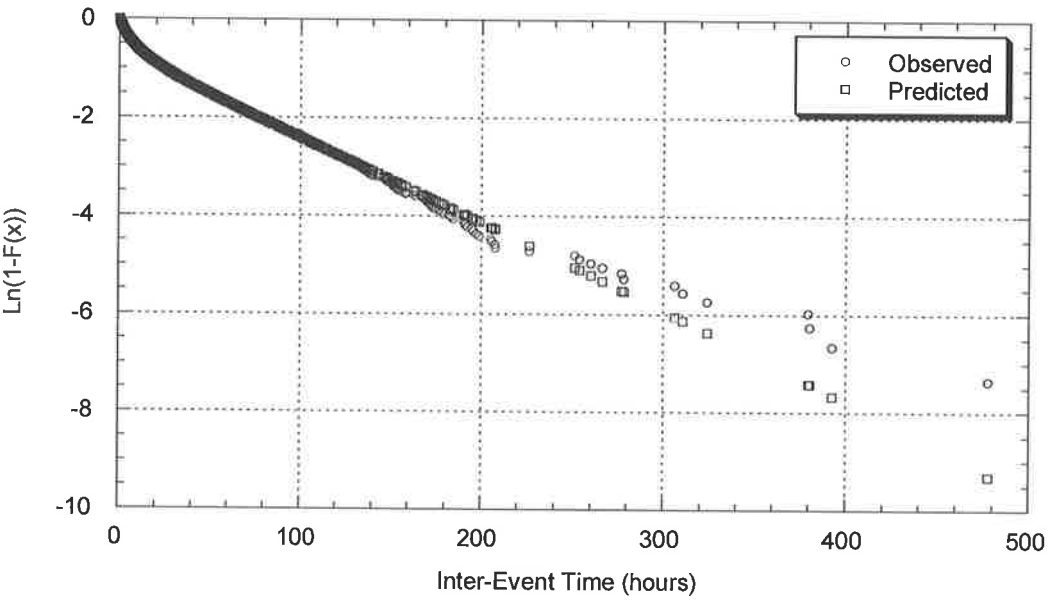

Figure 5.3: Observed and Predicted Inter-Event Distribution, Melbourne, September

A similar result was observed for the representation of storm duration event distributions. Figure 5.4 (Brisbane in July), Figure 5.5 (Sydney in January) and Figure 5.6 (Melbourne in October) show good agreement between observed and predicted values after calibration.

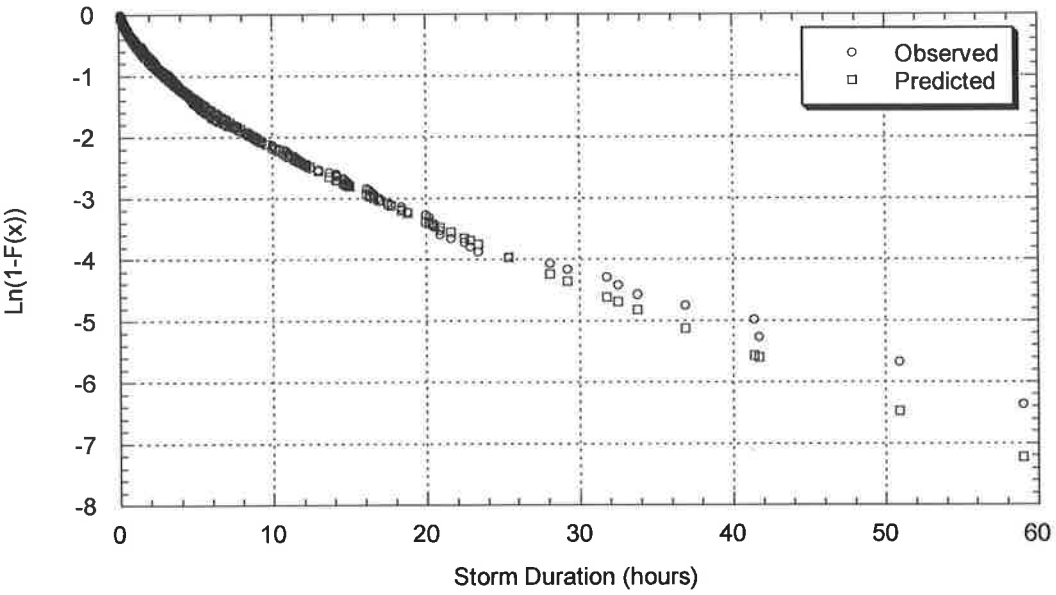

Figure 5.4: Observed and Predicted Storm Duration Distribution, Brisbane, July

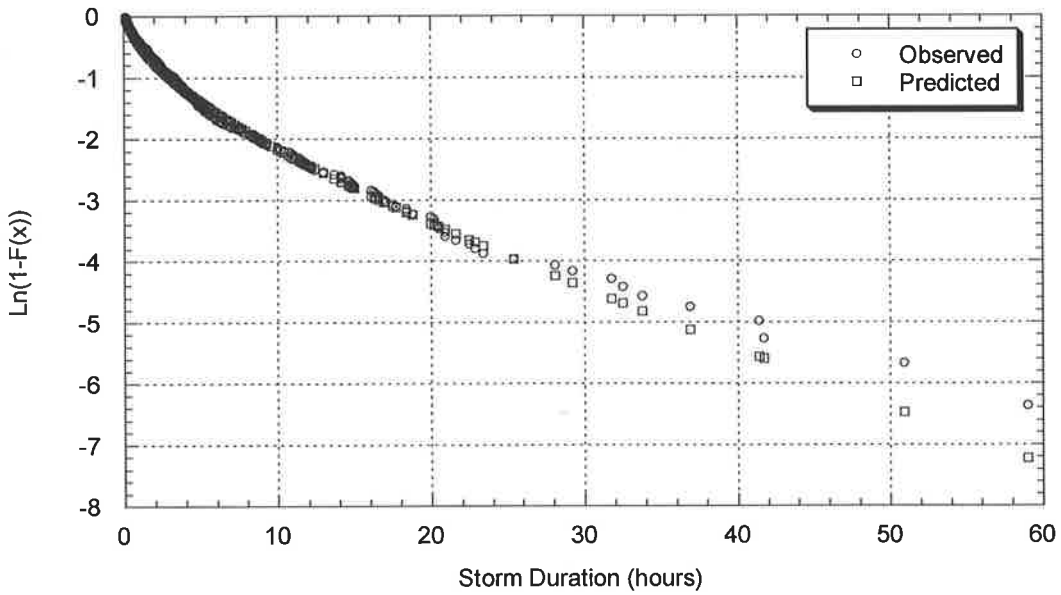

Figure 5.5: Observed and Predicted Storm Duration Distribution, Sydney, January

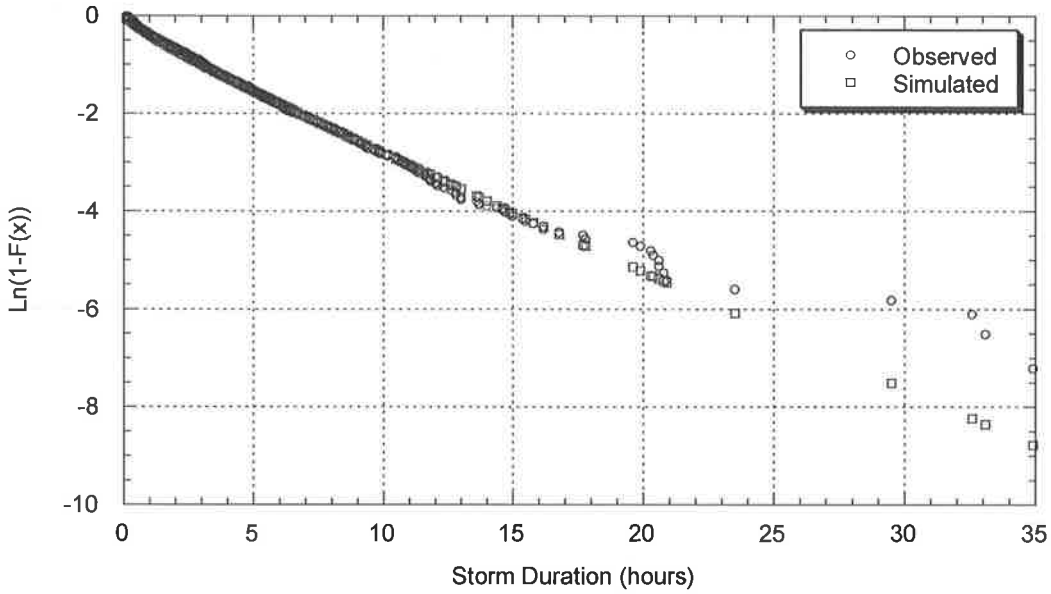

Figure 5.6: Observed and Predicted Storm Duration Distribution, Melbourne, October

Once the distributions of storm event parameters are satisfactorily replicated, a comparison can be made between observed and simulated monthly averages for these storm event variables. Figure 5.7 (Brisbane) and Figure 5.8 (Sydney) present the mean and standard deviation of inter-event times for each month. The results display excellent agreement between the observed data and simulated results with all observed points sitting well within the 90% simulation limits.

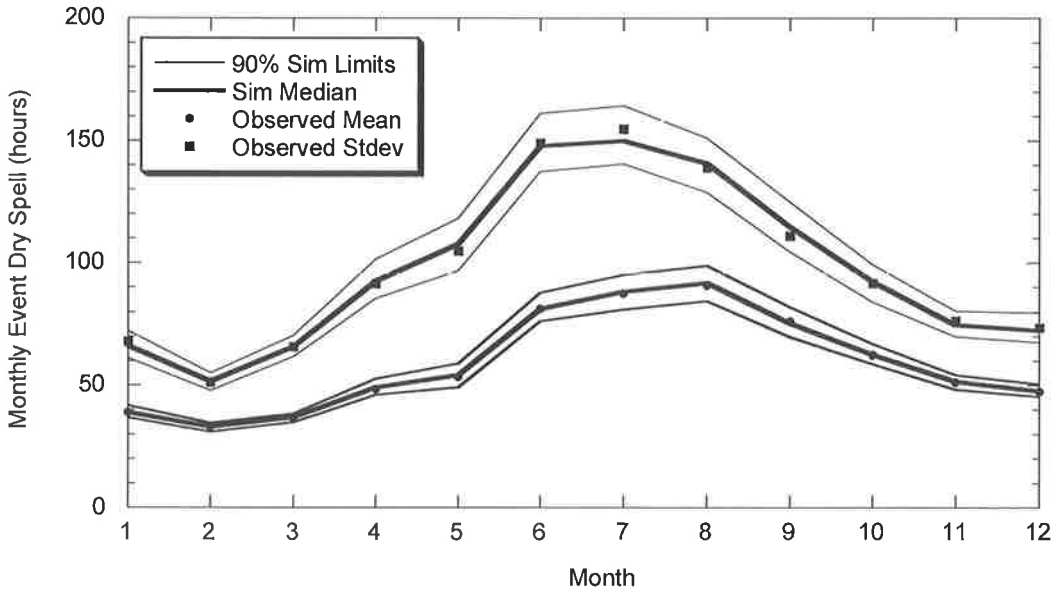

Figure 5.7: Mean and Standard Deviation of Event Dry Spells for Each Month, Brisbane

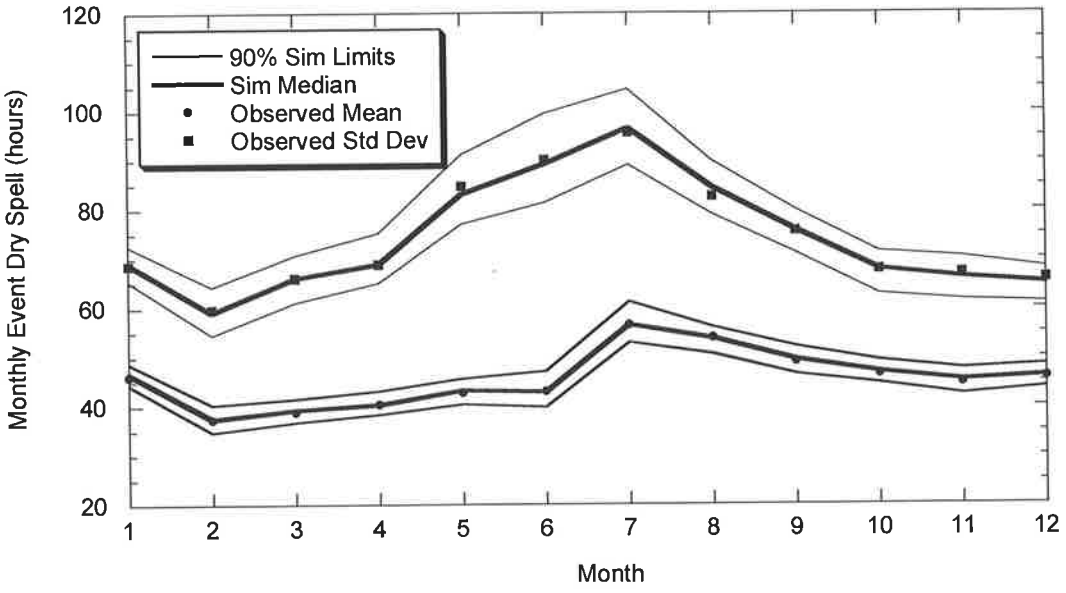

Figure 5.8: Mean and Standard Deviation of Event Dry Spells for Each Month, Sydney

Again, a similar result was observed for the reproduction of storm durations. Figure 5.9 (Brisbane) and Figure 5.10 (Sydney) present these results. In almost all cases the observed data lies within the 90% simulation limits. The mean storm duration for June in Brisbane has been simulated at 7.3 hours just outside the 7.2 hour simulation limit. Overall, this is a satisfactory representation of event storm duration.

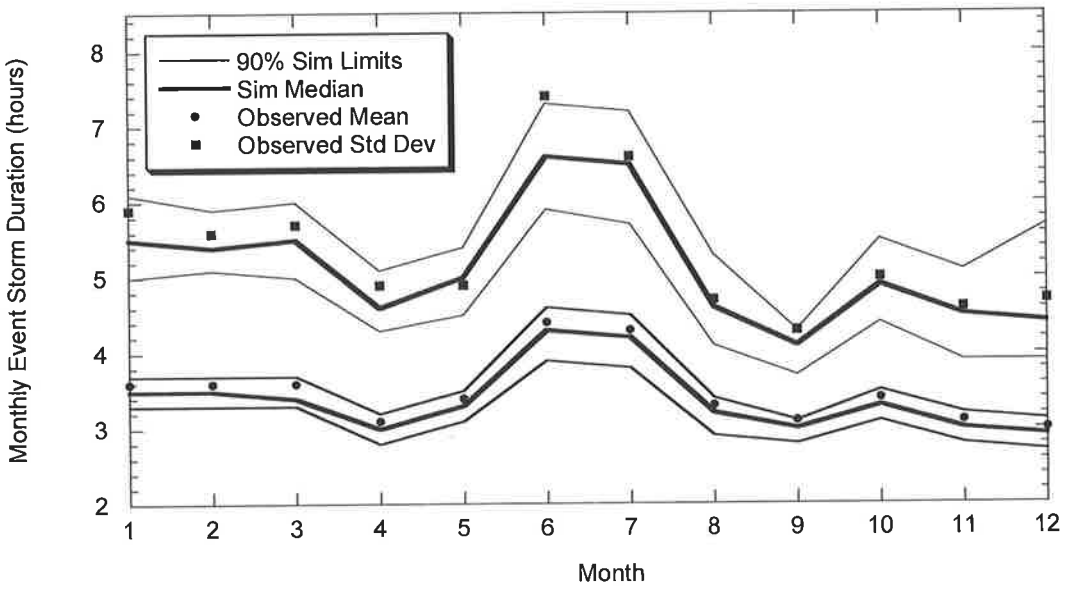

Figure 5.9: Mean and Standard Deviation of Event Storm Durations for Each Month, Brisbane

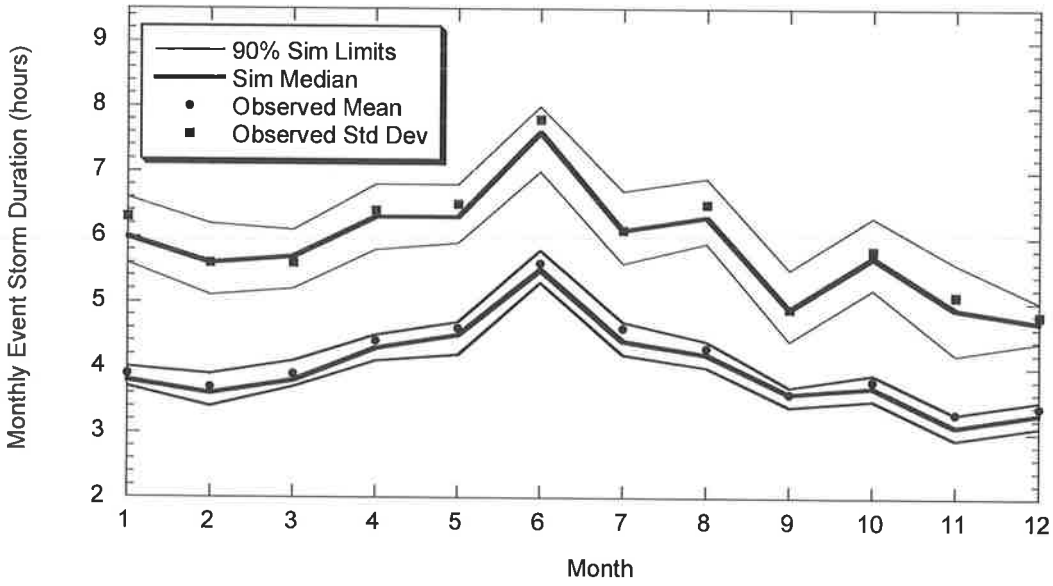

Figure 5.10: Mean and Standard Deviation of Event Storm Durations for Each Month, Sydney

The reproduction of event storm depth is extremely important to the overall success of the rainfall model for both the generation of synthetic pluviograph data and to capture aggregated rainfall statistics such as monthly and annual totals. Figure 5.11 (Brisbane) and Figure 5.12 (Sydney) present the mean and standard deviation of event storm depths for each month. These results display excellent agreement between the observed data and simulated results with all observed points sitting within the 90% simulation limits.

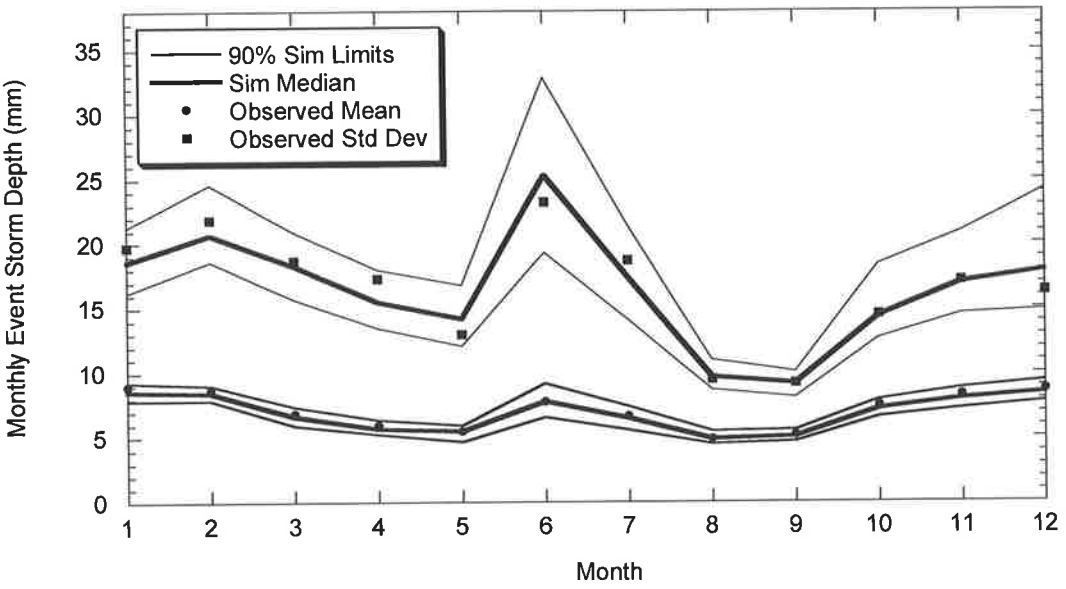

Figure 5.11: Mean and Standard Deviation of Event Depths for Each Month, Brisbane

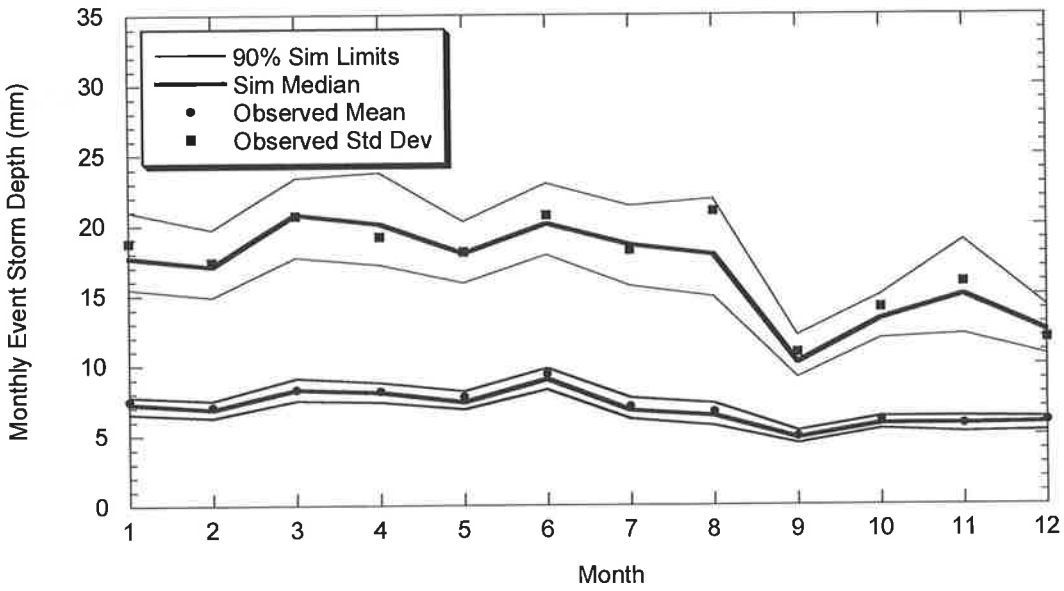

Figure 5.12: Mean and Standard Deviation of Event Depths for Each Month, Sydney

5.3 Intensity-Frequency-Duration

A comparison of extreme rainfall between observed and simulated data provides an insight into how well the rainfall model is able to reproduce the internal storm characteristics or the temporal pattern. In particular, the successful generation of short duration IFD curves (such as 1 and 3 hours) is dependent on the rainfall model accurately reproducing bursts of rainfall that occur during storm events. It is conceivable that the annual maximum 1-hour rainfall value used to produce the IFD curve may result from an hour burst inside a longer duration storm. This point is important as it provides an opportunity to validate the temporal pattern generator incorporated into the model. If the model is able to reproduce these IFD curves, then the temporal pattern generator has been successful. Conceptually longer durations such as those ranging from 12 to 72 hours are more likely to be a result of individual storm events, which in turn provide a validation of the original intensity-duration relationship rather than the temporal pattern.

IFD curves were obtained by moving windows of a fixed duration incrementally through each year and determining the annual maximum rainfall depths for each of these windows. A frequency analysis was then undertaken on these annual maxima for the various durations to produce the IFD curves. This is a standard method to determine the statistics of extreme rainfall over different durations for use in engineering design.

Figure 5.13 compares the observed and simulated IFD curves for 1, 12 and 72 hours for Brisbane. The simulated and observed IFD curves are similar with almost all of the observed data points sitting between the simulation limits suggesting that the model is able to reproduce the random bursts associated with short duration rainfall as well as the intensity-duration relationship. Figure 5.14 shows a similar result for data from Sydney.

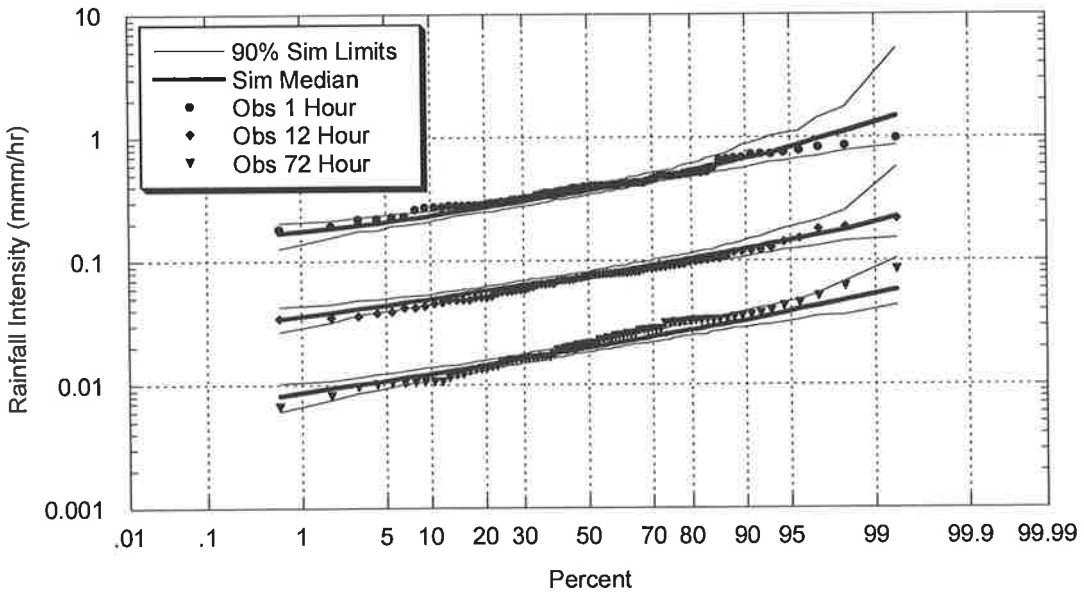

Figure 5.13: Simulated and Observed IFD Curves for Brisbane

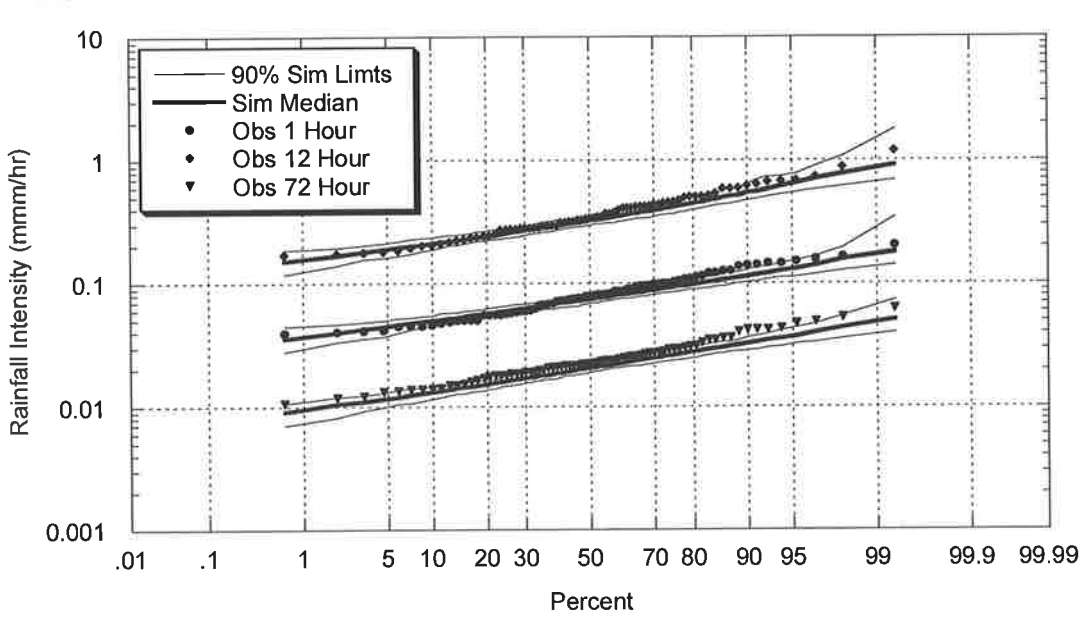

Figure 5.14: Simulated and Observed IFD Curves for Sydney

In conjunction with the IFD results presented above, these plots confirm the ability of the model to simulate the internal storm characteristics, which are important when simulating rainfall at a fine time scale.

5.4 Aggregated Depth Statistics

The reproduction of aggregated rainfall depth statistics at time scales such as hourly or daily is an important attribute of any model under consideration for use in volume based scenarios or water balance calculations. Reproduction of the 24-hour or daily aggregation values was also important if the model was to be adapted further for calibration with daily data using these statistics as part of a new regionalisation process.

The use of both a pluviograph and a daily data record provide the observed statistics for comparison. A concern when using pluviograph records in an application such as this is the possibility of missing data. It is common to see sections of missing data, which spans a few weeks or months within a pluviograph record. At smaller time scales it is acceptable to ignore these sections of missing data given the remaining large sample set that provides an adequate estimate of the aggregated statistics. When the time scale is increased however (i.e. monthly or annual rainfall totals), the pluviograph data tends to underestimate the actual observed rain totals. As a result, the observed daily data is aggregated for all comparison statistics at time scales greater than 24 hours. These daily records are generally longer, free of missing data and hence provide an improved estimate of these statistics.

Comparisons between the 1 hour aggregation statistics for observed and simulated rainfall depth are presented in Figure 5.15 for Brisbane and Figure 5.16 for Sydney. In almost all cases, the observed data sits within the 90% simulation limits. Only a slight deviation is observed for mean 1-hour April rainfall in Brisbane and September and November mean 1-hour rainfall in Sydney. This is considered a satisfactory result as these statistics were not part of the calibration process.

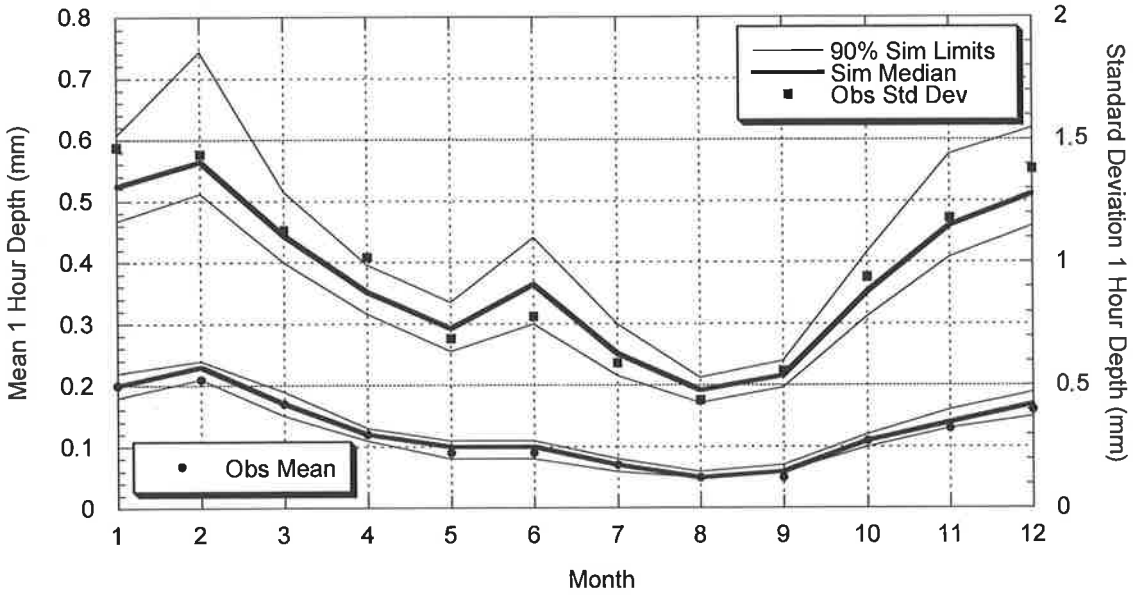

Figure 5.15: Mean and standard deviation of the aggregated 1-hour rainfall depth for Brisbane (mm).

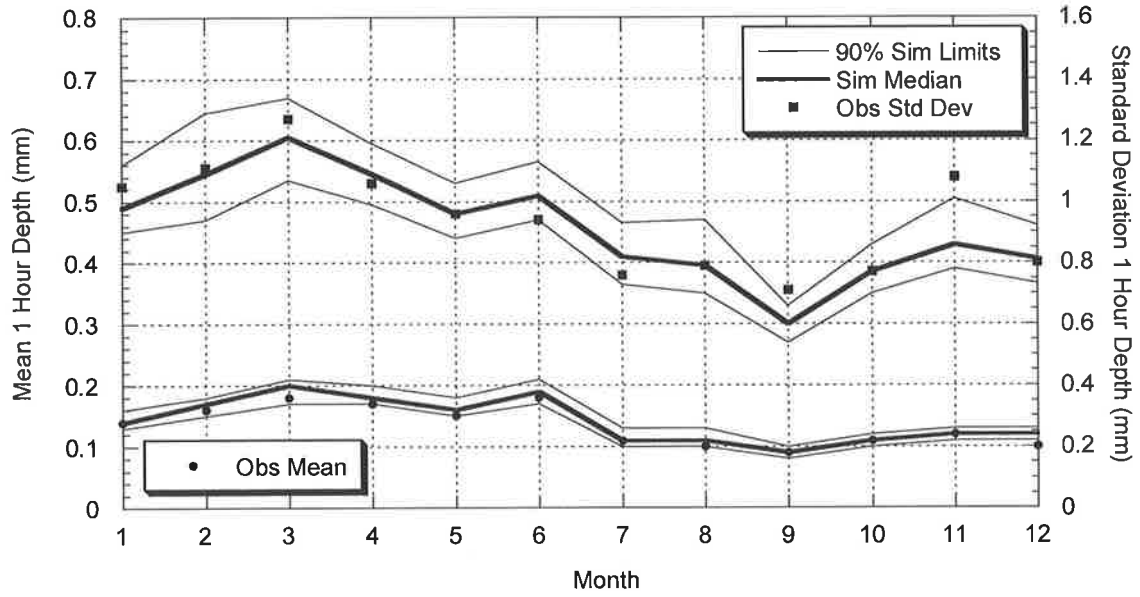

Figure 5.16: Mean and standard deviation of the aggregated 1-hour rainfall depth for Sydney (mm).

The reproduction of daily aggregated statistics was a major factor when considering a model for further development and specifically for regionalisation to locations where only daily data is available for calibration. Any model developed for the purposes of regionalisation should estimate a statistic for comparison to observed data (in some

form) in order to assess its performance during development (and potentially ongoing). In this case, the ability of the model to replicate 24 hour or daily statistics successfully provides an opportunity to compare the model to daily data after regionalisation and more importantly provides confidence in using the available daily data as part of the model calibration process. Without this ability, there would be no way to either develop a model calibrated to daily data nor compare the calibration or performance of the model at sites with daily data even if a regionalisation/calibration approach was developed.

Figure 5.17 (Brisbane) and Figure 5.18 (Sydney) present the mean and standard deviation of 24 hour aggregated rain depth. Importantly the improved model has demonstrated its ability to reproduce these statistics with good agreement between the observed points and the model simulations.

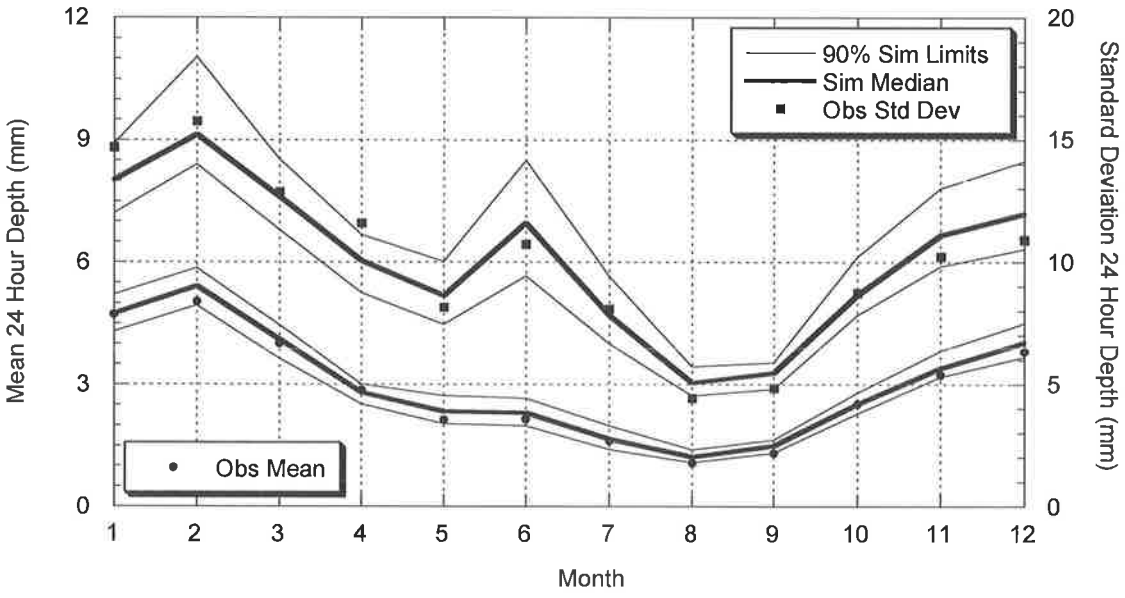

Figure 5.17: Mean and standard deviation of the aggregated 24-hour rainfall depth for Brisbane (mm).

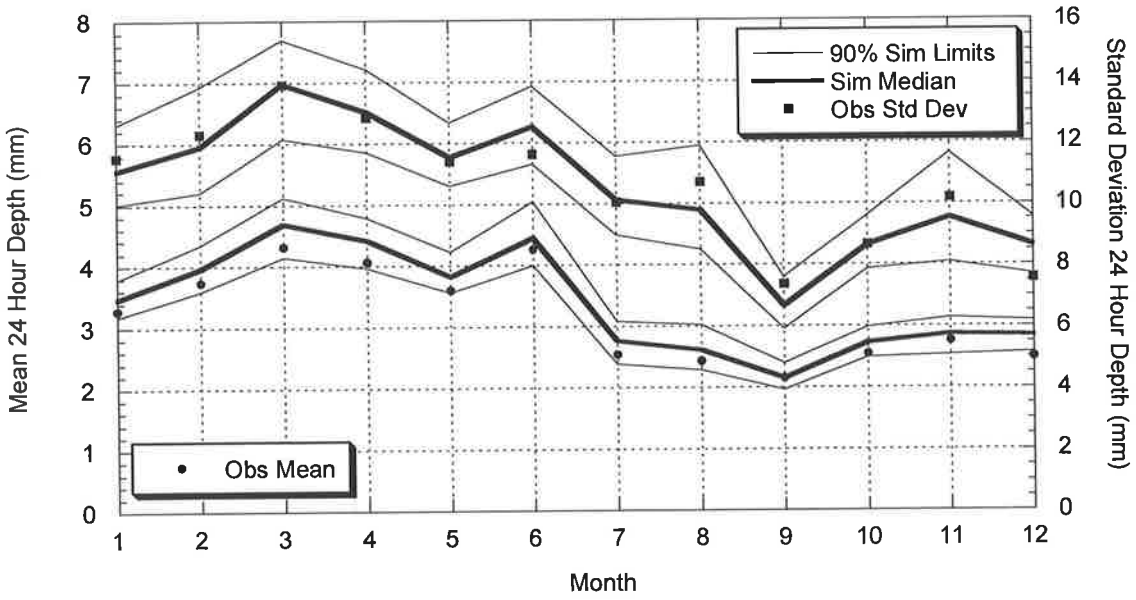

Figure 5.18: Mean and standard deviation of the aggregated 24-hour rainfall depth for Sydney (mm).

5.5 Annual Rainfall

Annual rainfall is an important statistic in engineering analysis particularly for longer-term planning and option evaluation. Figure 5.19 presents the annual rainfall results for the model calibrated to Brisbane. This result suggests a successful reproduction of annual mean rainfall as indicated by the agreement between the observed and simulated values at 50%. However, the angle or slope of the plot gives an indication of the standard deviation and as the slope of the observed data is steeper than that of the simulated values, the standard deviation produced by model simulation is presently underestimated. Figure 5.20 presents a similar outcome for simulated and observed data in Sydney.

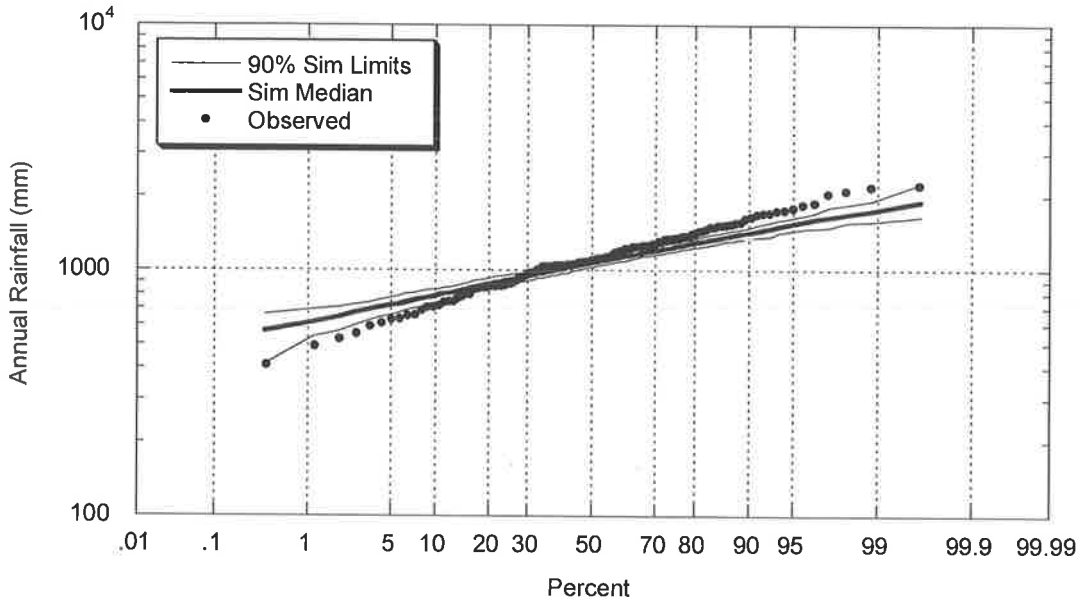

Figure 5.19: Simulated and observed annual rainfall distributions for Brisbane

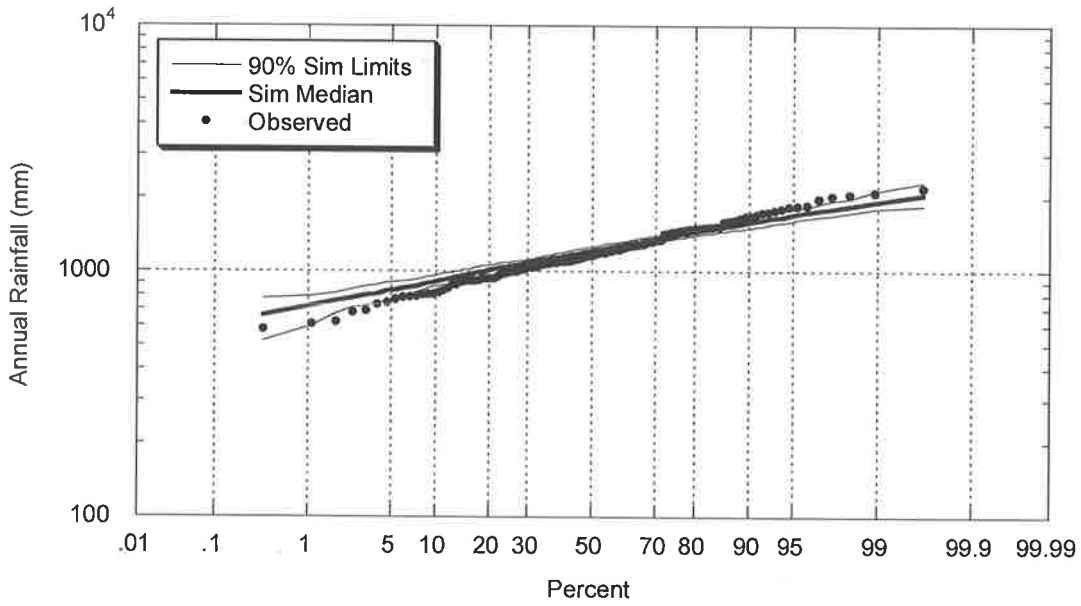

Figure 5.20: Simulated and observed annual rainfall distributions for Sydney

This result is typical of event-based models that attempt to describe the underlying rainfall process with independent wet and dry storm events. Historically the major consequence of this independence assumption has been an inability to take into account any inter-annual persistence such as El Nino or the Southern Oscillation. The result of this phenomenon is inter-annual persistence in the form of consecutive

very dry or wet periods in the rainfall record. The independent event-based model structure does not generally incorporate an inter-annual persistence characteristic, which in turn leads to an underestimation of the variability of annual rainfall totals. While research has begun to develop an understanding of exactly how El Nino affects rainfall in Australia (Chiew *et al.* (1998)), presently more work needs to be done before such relationships could be explicitly incorporated into the model structure.

As a potential solution to the problem, significant work has begun (Frost *et al.* (2000)) on introducing annual rainfall states into the event-based model structure. These states mimic the wet and dry years observed within the historical data set with each state requiring its own parameter set, calibrated against the observed rainfall data. Controlled switching between wet and dry model states during calibration provides a persistence structure capable of replicating the inter-annual persistence of El Nino. More information on the defining of these states can be found in the Ph.D. thesis of Andrew Frost (Frost (2004)) from the University of Newcastle.

5.6 Record Length

One of the issues with stochastic models calibrated to historical recorded data sets is the available locations and data available for calibration. As discussed in Chapter 1, this is particularly relevant for models such as this one which are calibrated to pluviograph data, a resource that is constantly under threat due to economic and political pressures. Caretakers of pluviograph stations can no longer continue to record data for little or no purpose. As a result more pluviograph stations are set up for a short (5-10 years) lifespan to serve a specific project or objective. All this puts additional emphasis on the ability of stochastic models to use alternate data sources for calibration.

The issue of appropriate record length for model calibration is one that is complex and cannot be answered definitively. Rainfall trends and observations are different at every site and as a consequence the data requirements at each site for a model such as this are not identical. This can be illustrated using an extreme comparison between the moderate climate in Adelaide and the tropics of Cairns. An Adelaide

record of 40 years should encompass numerous storms scattered throughout the year and as a consequence over a 40 year record provides a significant monthly data set for model calibration. In comparison a tropical location such as Cairns which observes defined wet and dry seasons may not encompass a significant number of storm events throughout the dry months within a 40 year record. As a consequence the model calibrated to 40 years in Cairns may not be as accurate as the one calibrated to the same 40 years in Adelaide.

While it is difficult to define an absolute record length requirement, it is clear that the model does require a significant historical pluviograph data for calibration in its original form. Results presented in Chapter 4 suggest the model calibrated to Perth data requires approximately 30 years of record for accurate calibration and that a short 10 years of data was not adequate to successfully capture the mean annual rainfall or the annual rainfall distribution. Experience suggests that 30 years is generally adequate to calibrate the original model. Figure 5.21 reinforces this idea and shows that the model calibrated to the full historic record has successfully reproduced the annual rainfall distribution, while the model calibrated to the short record has underestimated the annual rainfall totals.

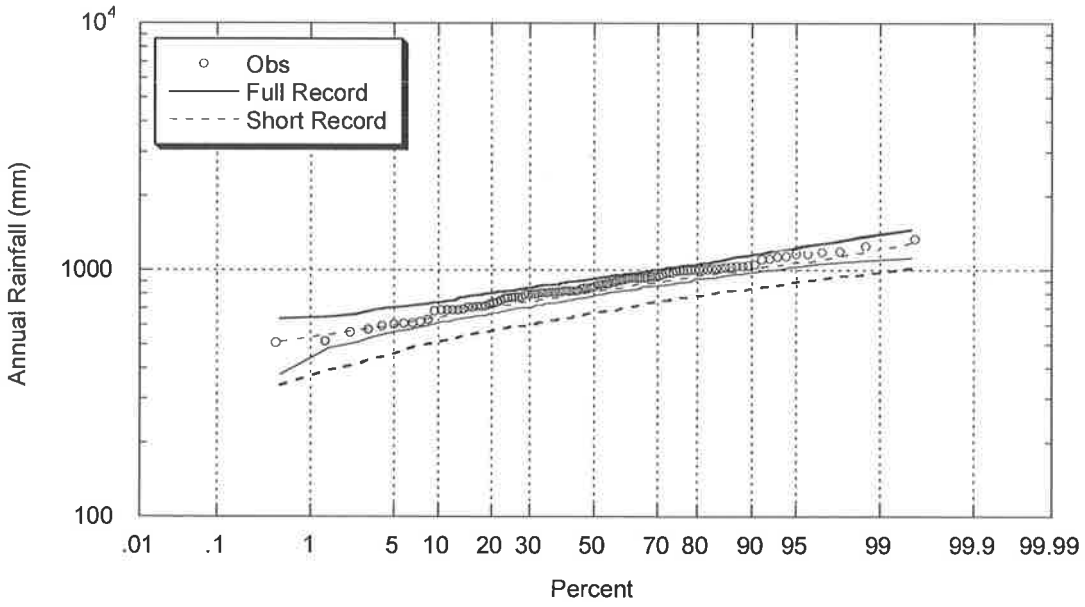

Figure 5.21: Comparison of Annual Rainfall limits due to Full and Short Record Lengths and Incorporating All Uncertainty Options, Perth

The inability to define with certainty the required length of calibration data does not reduce the effectiveness of the model. As with all models, uneducated use of resultant simulations and outputs is not advised. Model calibration and results require checks and possibly "engineering judgement" prior to successful application. The setup and structure of this model enables these checks to be undertaken easily and efficiently. If a model is not able to be calibrated directly with the available data, then an alternative calibration technique is required and can be provided using a regionalisation approach.

5.7 Summary

The generation of synthetic rainfall using the improved model adapted in this study was a success. The ideas of parameter uncertainty and sampling variability presented in Chapter 4 were used to improve the simulation and enable the generation of simulation limits for all model results. This in turn has provided an improved comparison between observed and simulated values. While a minor number of observed values fell immediately outside of the 90% simulation limits, overall model simulations compare favourably with observed statistics across all sites.

The distributions of inter-event time and storm duration were able to successfully describe the observed distributions using the new three parameter model. This model developed in Chapter 3 and adapted from the work of Heneker *et al.* (2001) has provided a robust and efficient description of these distributions. Both observed and predicted event distributions and comparisons between observed and simulated event statistics showed good agreement validating the new model setup. Observed and simulated storm event depth also compared favourably suggesting the selection of set breakpoints and using an initial continuous calibration function as described in Chapter 3 worked well.

The reproduction of aggregated rainfall depth statistics at the 24-hour or daily aggregation level was also important if the model was to be adapted further for calibration with daily data using these statistics as part of a new regionalisation process. The daily mean rainfall was well reproduced suggesting this statistic could

be used as part of any future regionalisation work. In addition sub-daily 1 hour aggregated rainfall was also well reproduced.

The rainfall model was able to replicate the mean annual rainfall however was not able to successfully reproduce the distribution of annual rain, slightly underestimating the standard deviation. This result is typical of event-based models that attempt to describe the underlying rainfall process with independent wet and dry storm events. As a result, they do not take into account inter-annual persistence due to long-term events such as El Nino. Work is currently underway to develop an explicit model of this underlying structure (see Thyer & Kuczera (1999) and Frost (2004)). It is important when developing a regionalisation structure that either this inter-annual persistence is taken into account or similar periods within the historical time line are compared when developing the regionalisation model (i.e. comparisons between sites over similar wet years).

The improved stochastic model of rainfall developed in this study was now very suitable for further development for potential regionalisation/calibration with short pluviograph or daily data records.

CHAPTER 6

REGIONALISATION WITH A SHORT PLUVIOGRAPH RECORD

6.1 Introduction

Rainfall models in Australia (and indeed internationally) are often restricted in their application as a result of inadequate data available for calibration. Models calibrated to storm events including the model at the centre of this research are particularly difficult to calibrate and require significant historical pluviograph data records. Unfortunately, data of this resolution is difficult and expensive to record and as a result long accurate pluviograph records in Australia are not abundant.

Of the more than 900 pluviograph sites in Australia managed by the Bureau of Meteorology, the combined average length is only approximately 15 years. Even if this value is slightly biased by a number of sites that are relatively new (or were recorded for a specific purpose and contain only a few years of record), more alarming is the fact that of all sites that are still active, only a few (35-40) have a record length greater than 40 years. However, the network of over 900 pluviograph sites which contain a short historical record is extensive and would provide an excellent source of calibration data if a procedure to take advantage of the availability of short pluviograph records was developed. If models such as the one used in this research are to be seriously considered for use as an engineering tool, a method to calibrate these model's with less observed data for calibration was required.

While the number of long term pluviograph records across Australia is small, it was the distribution of these records across the country which led to the development of a new regional calibration technique. With the spread of long term records scattered

throughout the major climatic regions of the country (i.e. Temperate climate of Adelaide/Melbourne/Sydney, Sub-Tropics of Brisbane etc) a process was developed which used the information at a long pluviograph site and then adjusted the model based on comparisons to a smaller data set at the site of interest. The adopted calibration framework uses a master to target (or slave) relationship which enables the successful updating of model parameters from one site to another. Model parameters are first calibrated to the long record at a master site and then translated to the target site using the available short pluviograph data set. The process uses the expectation that parameter distributions within the model should exhibit a similar shape when calibrated to data within a similar climatic region (i.e. sites surrounding Melbourne).

An intermediate step is also introduced which is able to manage issues that arise when data sets of non-concurrent time periods are compared. By developing an additional step in the process, the model is able to consider the real parameter changes in the model between the master and target sites and is not influenced by the differences in data length or recording periods.

The introduction of the master – target relationship and the regional calibration process produces a model that is well calibrated at the target site and able to synthetically extend the short historical pluviograph record.

6.2 Regionalisation Model Structure

Investigations outlined in the previous chapters into the various components of the rainfall model and their variability and predictability between sites led to the development of the final regionalisation model structure. The initial problem itself is well founded; a model simulation is required at a particular site which has a short historical pluviograph record and either cannot be calibrated directly due to the lack of data or the resultant calibration produces a model which does not adequately reflect the rainfall pattern at the site of interest. In the first instance, the extreme lack of data ensures the model calibration process itself cannot be completed (i.e. parameter values cannot be found) while the second situation relates to the fact that while it may be possible to complete the calibration process with limited available

data, this does not necessarily ensure that the model will perform adequately and represent the required statistics and patterns. The challenge then becomes one of how to utilise the available data set which is not sufficient on its own to produce a successful calibration to provide a model which does accurately reproduce the observed long term rainfall statistics.

The improved rainfall model consists of four major components which work together during simulation to replicate the observed record. Initially a dry spell or inter-event time is generated. This is followed subsequently by a wet spell or storm duration, with these two event parameters defining the rainfall time series. Associated with this storm duration is the sampling of a storm intensity which provides the total storm depth and finally the process is completed by disaggregating the total storm depth via the temporal pattern model to produce 6 minute rainfall data. To provide an accurate simulation, these components must all be successfully calibrated at the site of interest. The model structure dictates that each component is independent, so for the purposes of calibration these components can be investigated independently and a process developed for each which enables these components to be successfully applied to a particular site with only a small amount of historical data.

If direct calibration at the site of interest is not available due to a lack of data, a regionalised calibration could be achieved through using this short historical record to update an existing calibration at a nearby pluviograph site with a long historical record. This structure should ensure that any major rainfall processes and events over time have been described within the calibration at the nearby long pluviograph data site, while minor differences between the sites are captured via the updating regionalisation procedure. During simulation, the model simulates storm events at the master site which are then translated in a reverse of the process to equivalent storm events at the target site by using the same model setup. The master – target framework is displayed schematically in Figure 6.1 below.

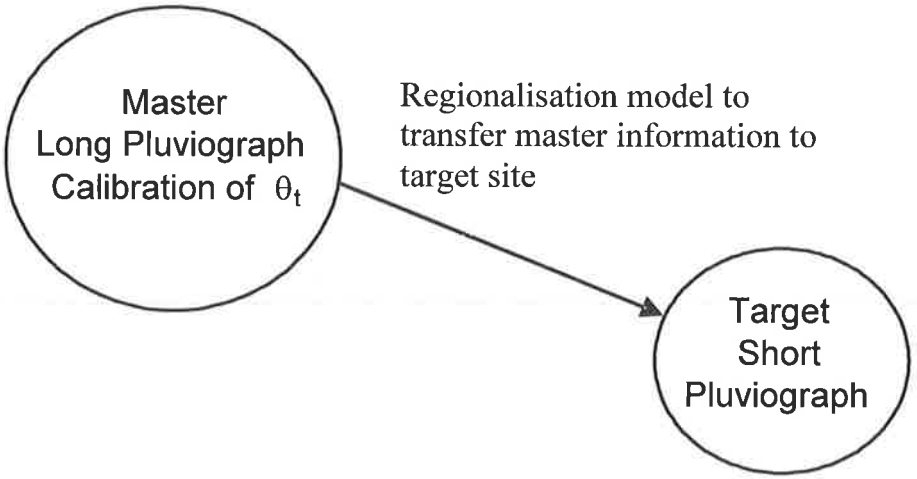

Figure 6.1: Schematic of Regionalisation Structure

For this structure to work, sites must be situated close enough to have experienced similar climatic patterns (i.e. the process does not work when shifting from the temperate climate of Adelaide to the tropics of Brisbane). Unfortunately as of the time of writing, there has been no definitive formula or basis found which can determine whether or not a parameter set can be shifted from one site to another. Experience to date indicates that sites anywhere from 10 to 580 km away from the corresponding master site have been successfully calibrated and simulate accurately the observed pluviograph statistics. It is anticipated in the future that as the number of significant data sets increase, and more data becomes available that work can be done to either formulate such a basis or possibly in the long term develop parameter contours or a similar generic mapping of model parameters across the country. Until this occurs, the simulation accuracy of the model should be tested at each individual site with the available pluviograph and/or daily data as a comparison to ensure that the regionalisation process has been successful.

6.2.1 Preliminary Investigations

To understand the possible adjustments that would need to be made during the regionalisation process, changes in parameter distributions were investigated by calibrating the model to numerous long historical data records across Australia. These calibrations were then compared across neighbouring sites influenced by similar climatic conditions. As a basis, pairs of sites surrounding a master site were selected to ensure they had a similar climate classification by the Australian Bureau of

Meteorology. It was expected that the parameter distributions for each model component would display a similar shape for sites within a similar climatic region (i.e. inter-event times in a particular region would exhibit the same general shape). This was because sites located in a similar region should be influenced by the same major climatic fronts and forces.

In addition to this over-arching consistent shape, it was anticipated that micro-climatic influences which influence each individual rainfall site would alter these distributions and could be described in some form to enable a regionalisation updating process to be incorporated into the model. These local site adjustments may be a result of micro climatic factors (i.e. altitude, aspect, and distance from coast etc) which do not alter the climatic driving forces influencing the major storm systems, but they do alter the frequency and volume of rainfall recorded between sites within the same climatic region. This minor influence is postulated to change the parameterisation of the distributions in the model, but not the overall distribution shape. This would allow the parameter distributions to remain the same while still enabling the use of a scaling shift between sites to capture this minor influence.

To gain an understanding of potential adjustments required between sites, a number of calibrations at long pluviograph sites were compared. To ensure an accurate comparison, concurrent data periods were selected and used in the calibration process. As an example to understand the model development process, work on the inter-event times is presented below. Figure 6.2 displays the distributions of inter-event times for Melbourne and East Sale in Victoria. These sites are 190 km apart but both are influenced by the Southern Australia weather patterns and experience similar climatic conditions.

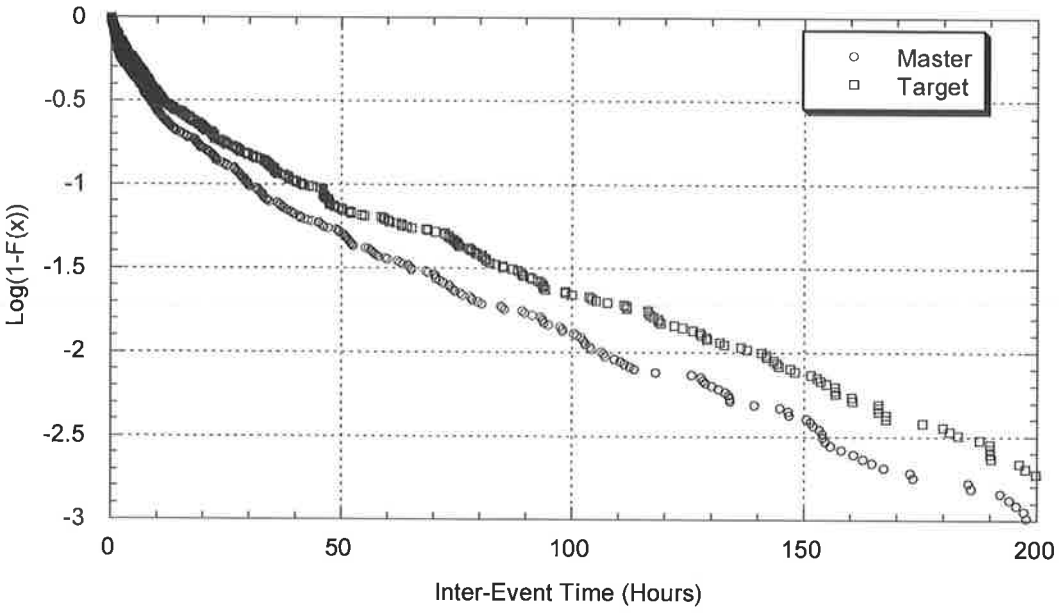

Figure 6.2: Comparison between Melbourne 86071 (Master) and East Sale 85072 (Target) Inter-Event Distributions for the month of April.

This plot suggests that while the two sites have distinctly different parameter distributions, the underlying shape (i.e. initial curvature and slope) is similar. Additional pairs of plots across varying climatic conditions provide a similar result as evidenced by Figure 6.3 below (Comparison between a calibration at Sydney and Richmond RAAF)

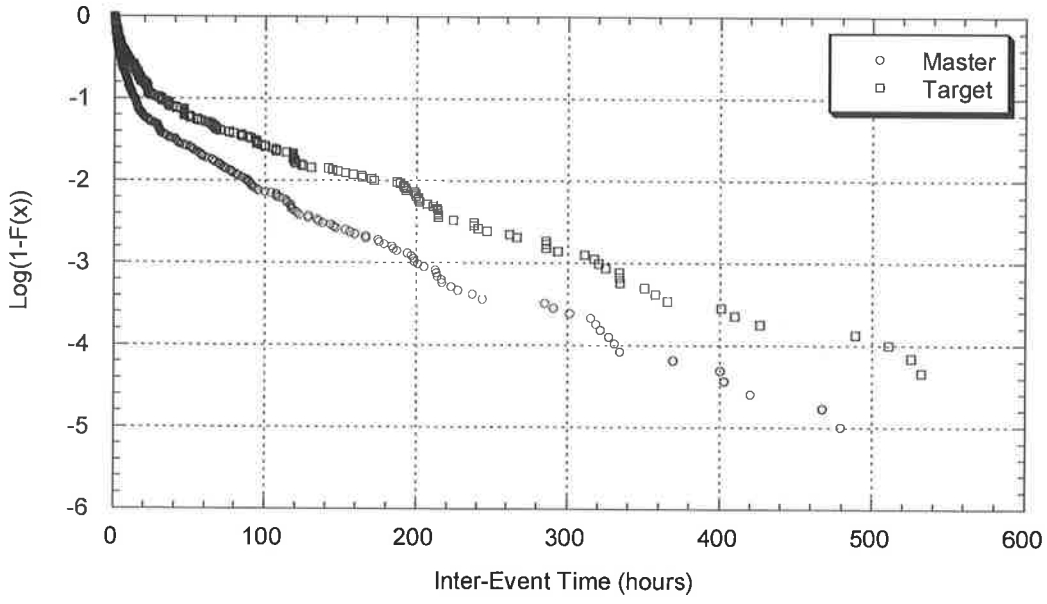

Figure 6.3: Comparison between Sydney 66062 (Master) and Richmond 67033 (Target) Inter-Event Distributions for the month of June

From these and comparisons completed at numerous other pairs of sites, it is evident that the differences between parameter distributions within a region are adequately consistent for the development of a regionalisation model. There are no sudden changes in the shape of the distribution from site to site and not only that, the difference between distributions is relatively consistent (i.e. the differences increase as the length of inter-event time increases). This was an important result as it suggested that differences between inter-event time distributions could be described through the introduction of a linear scaling relationship.

To incorporate the linear scaling relationship into the model, a new relationship was applied between inter-event times at the longer master site and those at the shorter target site. This scalar relationship was dependent on the length of the observed storm event and translates master site events into a corresponding storm event at the target site. If we consider the original calibration equation for inter-event time at the master site as presented in Chapter 3, the relationship is as described by

$$F(x \mid \theta_t) = P(X \le x \mid \theta_t) = 1 - \exp[-g(x, \theta_t)] \qquad x > 0 \qquad (6.1)$$

Introducing a new linear scaling relationship and variable λ into this relationship to provide the required shift between the master and target distributions yields a new equation (6.2) for calibration of a regionalised inter-event time

$$F_Y(y \mid \theta_t, \lambda) = F_X(\lambda y \mid \theta_t) = 1 - \exp[-g(\lambda y, \theta_t)] \qquad (6.2)$$

F_Y denotes the distribution function at the target site while F_X denotes the distribution function at the master site. The new scaling value λ acts as a multiplier on the target pluviograph data (y), to provide a best fit between the master calibration and the target data. In this manner the minor differences between the two distributions can be described during the calibration. This process is displayed in graphical form below

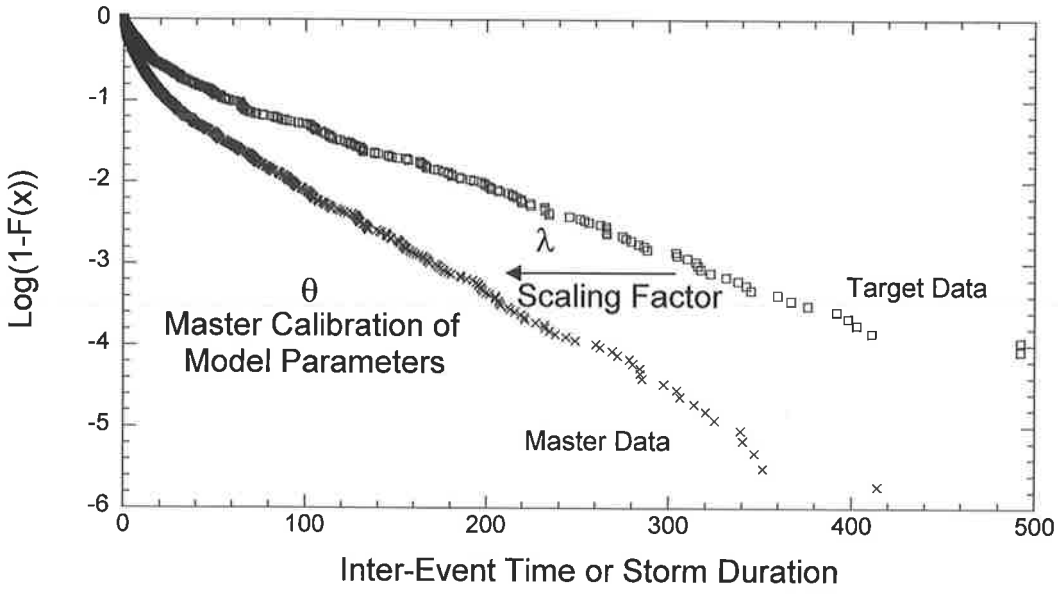

Figure 6.4: Schematic of Master Target Event Scaling

While this model worked when comparing data sets of identical length and concurrent time periods, the process in this form is not adequate for describing distribution changes between sites which contain different record lengths or are recorded over different periods of time. To understand why, we need to consider the sampling variability at the annual scale, an issue which was discussed briefly in the work on parameter uncertainty. In general terms, when comparing statistical distributions of significant data sets the removal of a small section of data from one record should not significantly alter the distribution or the comparison. However, rainfall records can be short (small in data terms) and can also incorporate long term persistence within sections of the record relating to a particularly unusual dry or wet period which may continue over a multi-year period. If we consider a shorter subset of a long historical rainfall record, it is evident that the subset cannot be influenced by all of the wet or dry periods (sampled in the longer record) which in turn can produce a different data distribution between the long record and its own subset. Expanding this, a similar result is observed when comparing a long pluviograph record and a neighbouring short record in that the short record has not been influenced to the same degree by any long term influences and may have missed significant periods of rainfall information. As a result, calibration of the linear regionalisation model between a long and short record (as described in Figure 6.1

above) would not only be influenced by small climatic differences at the sites, but also by any differences between the recorded time periods of the records. This introduces a major source of error and makes it impossible to describe the event distribution changes which arise only from the slight climatic differences between sites. Before the regionalisation model structure could be adopted and further developed to incorporate storm duration and storm depth parameters, this issue of sampling variability at multi-year scales and how to deal with non-concurrent data periods required careful consideration.

6.2.2 Treatment of Sampling Variability at the Annual Scale

To demonstrate the potential sampling variability issue, a pseudo calibration was undertaken by shifting from a calibrated master parameter site (using inter-event parameters) to a smaller subset of the same master record which was acting as the target site. As the master and target data are from the same site, the regionalisation model should return a scaling factor of one. (i.e. there are no adjustments to be made to the calibrated model parameters as the data is from the same site and therefore the master parameters are a true reflection of the required parameter set) Any sampling variability as a result of the smaller data set not observing the same rainfall periods as its longer counterpart would be evidenced by a deviation from this expected value of one. In this case using data from Melbourne (Full length 1900-1995, subset 1985-1995), Figure 6.5, it is clear that this is indeed the case and the resultant scaling factors did not equal one.

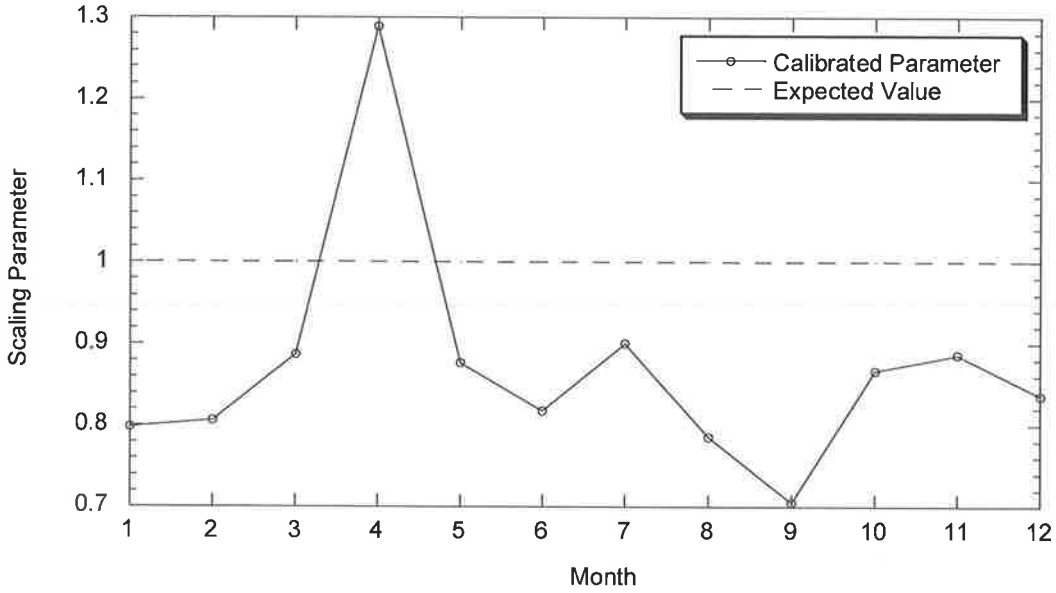

Figure 6.5: Calibration of Scaling Parameter without adjustment due to Different Record Periods (Data from Melbourne; Full length 1900-1995, subset 1985-1995)

If the time span of the shorter record coincided with a predominately wet or dry period within the full record, then the short record will be influenced by these events. As a result the distributions describing the storm events for the shorter record will differ (to those of the full record) and the resultant scaling factors will not equal one. Clearly this problem also translates when considering regionalisation between sites. Direct comparison between the full length master calibration and the shorter target site data will lead to a potential bias in the resultant model parameters as a result of regional wet or dry periods coinciding (or not) with the time span of the shorter record. To accurately calibrate the scaling factor between sites, the model needs to compare the differences between the distributions which relate to local climatic conditions and should not be influenced by issues arising from differences with the time span and length of the two data sets under comparison.

Additional understanding of this potential bias can be seen with a simple graph comparing the inter-event distributions of a full length record to a subset of that record for data from Brisbane and Perth (Figure 6.6 and Figure 6.7).

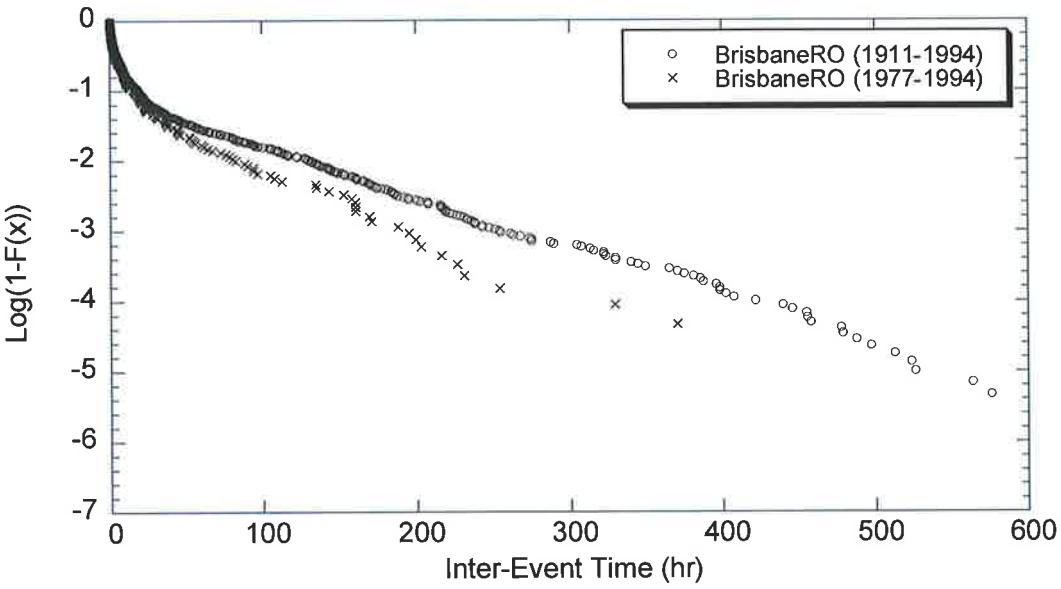

Figure 6.6: Inter Event Distributions, Brisbane Different Record Length

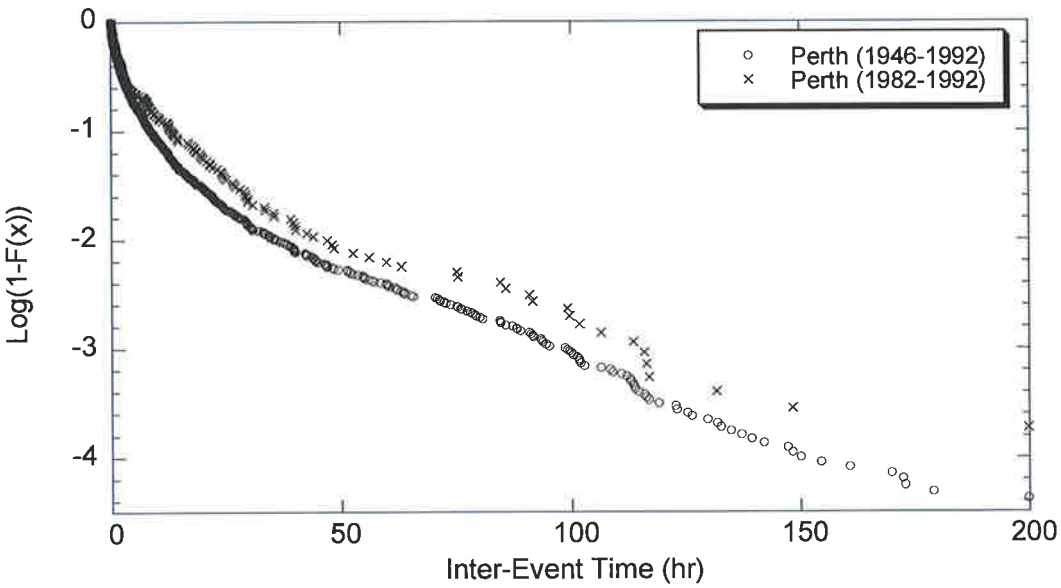

Figure 6.7: Inter-Event Distributions, Perth Different Record Lengths

It is clear from these plots that while the shape of the distribution has remained consistent, the actual parameterisation of the distributions would differ. Considering Figure 6.7, a model calibrated to Perth data from 1946-1992 would have a slightly

different set of parameters to the model calibrated to the subset of Perth data from 1982-1992.

To remove any sampling variability and ensure an accurate calibration of the required scaling factor for regionalisation, the target data should be compared to master model parameters which are calibrated to a master data set equivalent in length and concurrent in years. The use of concurrent data would ensure the calculated scaling factors are a true indication of the local distribution adjustments and are not a residual result of the different record lengths. However, in order to use a master – target relationship with concurrent data periods directly, the model would first need to be calibrated to a short sub-set of the master record equivalent in length to that of the target site. It is the inability of the model to be successfully calibrated to such a short pluviograph record that has lead to the development of a master-target relationship to circumvent the problem. Indeed, if the model was capable of being directly calibrated to a sub-set of the master record, then it would be reasonably safe to assume the model could also be calibrated directly to the target data set and a regionalisation process would not be required. This is not the case, as calibration of the original model to short pluviograph records is problematic, often resulting in model parameters that are not well identified and do not describe the required distributions due to the lack of available calibration data. A process which firstly allows the model to be calibrated to the complete record at the master site, secondly takes into account the differences in data lengths and time periods between the two data sets, and finally is still able to accurately adjust these model parameters based on the limited information from a short pluviograph record was required.

Given the target pluviograph data set will always be shorter and contain less information than the master record and that the master parameters should always be calibrated to the full length of the master record to ensure accuracy, the resultant problem is how to set up the regionalisation model to enable comparison between data sets which are not concurrent and/or equivalent. To solve this problem, an additional step in the regionalisation process has been introduced. This additional intermediate calibration step results in the calculation of an intermediate pre-scale parameter denoted λ_p. This λ_p parameter describes the differences between data distributions relating to non-concurrent time periods and by using it in the

calibration process it ensures comparisons between distributions at master and target sites are not influenced by non-concurrent data.

The amended master – target process to calculate the scaling parameter λ and the intermediate parameter λ_p is shown graphically below in Figure 6.8. The first step remains the same as the original master – target relationship whereby the model parameters are calibrated to the full length record at the master site. This parametric description of the master storm event distributions is denoted θ. The second step involves an interrogation of the target data set which allows the model to extract a subset of the master data which is concurrent and equivalent in length to the target data. Once this data is extracted, the model uses the master – target relationship to compare the original master calibration (θ) and the new subset of the master data (now acting as the target data) to calculate an intermediate scaling factor λ_p.

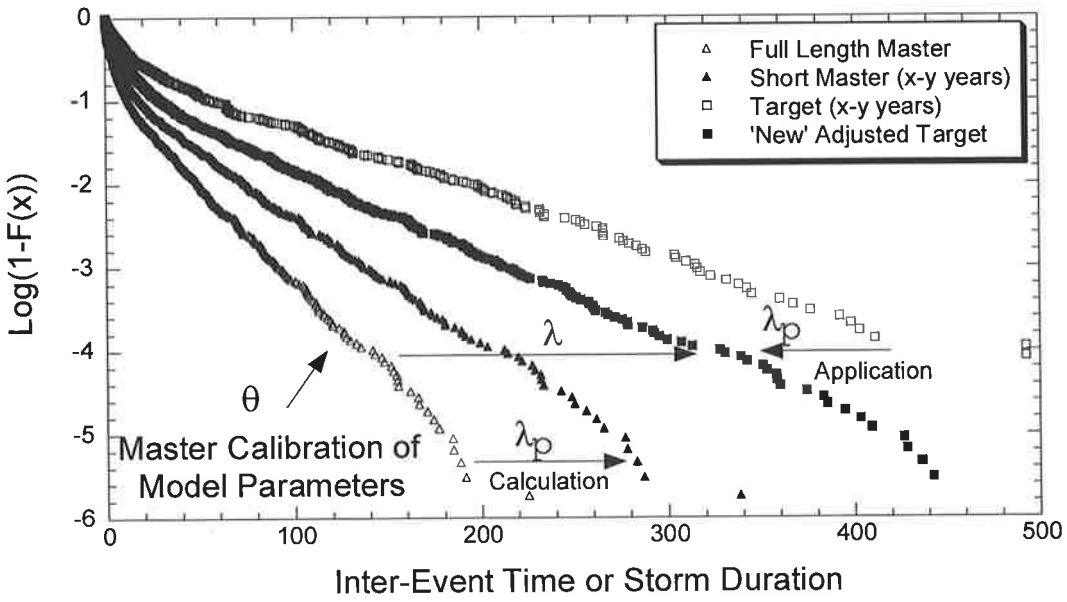

Figure 6.8: Schematic of Pre-Scaling Calibration Step

Once the intermediate factor λ_p has been determined, it is then applied as a pre-multiplication factor to the target data. By pre-multiplying by λ_p, the target data is adjusted to take into account any differences in the storm event distributions that may have occurred as a result of the differences in data lengths and time periods. Once this transformation of the target data is complete, the master - target scaling

process is applied a second time between the master parameters and the 'new' target data to determine the 'real' regionalisation scaling factor λ. The introduction of this intermediate step allows the regionalisation model to more easily determine the changes in storm event distributions and model parameters as a result of the differences between sites alone and not to be influenced by the differences in data lengths or time periods.

More explicitly, if we again consider the distribution of inter-event times, then as before (in equation 6.2) we know the introduction of the regionalisation model is described by

$$F_Y(y \mid \theta_t, \lambda) = F_X(\lambda y \mid \theta_t) = 1 - \exp[-g(\lambda y, \theta_t)] \qquad x>0,\ y>0 \qquad (6.3)$$

If we extend this and denoting λ_p as the intermediate pre-scaling factor to take into account the differences between data lengths and periods, the calculation of λ_p at the master site is described by

$$F_X(X \mid \theta_t, \lambda_p) = F_X\left(\left(\frac{1}{\lambda_p} X_{t\,arg}\right) \mid \theta_t\right) = 1 - \exp\left[-g\left(\lambda\left(\frac{1}{\lambda_p} x_{t\,arg}\right), \theta_t\right)\right] \qquad (6.4)$$

where F_X is the distribution at the master site and x_{targ} is the subset of the master data equivalent in years to the target data. This provides the calculation of the pre-scaling factor λ_p. Applying this as a multiplication to the target data and re-applying the master – target scaling relationship (6.3) gives the overall calibration of the scaling factor λ

$$F_Y(y \mid \theta_t, \lambda, \lambda_p) = F_X(\lambda(\lambda_p y) \mid \theta_t) = 1 - \exp[-g(\lambda(\lambda_p y), \theta_t)] \qquad y>0 \qquad (6.5)$$

As explained previously the target data y has been initially multiplied by λ_p to remove any bias due to the different time periods covered by the full length master pluviograph and the shorter target site. This allows a direct comparison between the calibrated master parameters and target distribution to calculate the real λ.

To test the introduction of the intermediate step, the technique was again used to calculate λ between the full length and a subset of the Melbourne data set but this time including the intermediate calibration step. Again, the expectation was the calculation of a scaling factor of one (in this case for inter-event times). Figure 6.9 shows that the inclusion of this intermediate step has corrected any bias due to the different record lengths and as expected returned a scaling parameter value of 1.0.

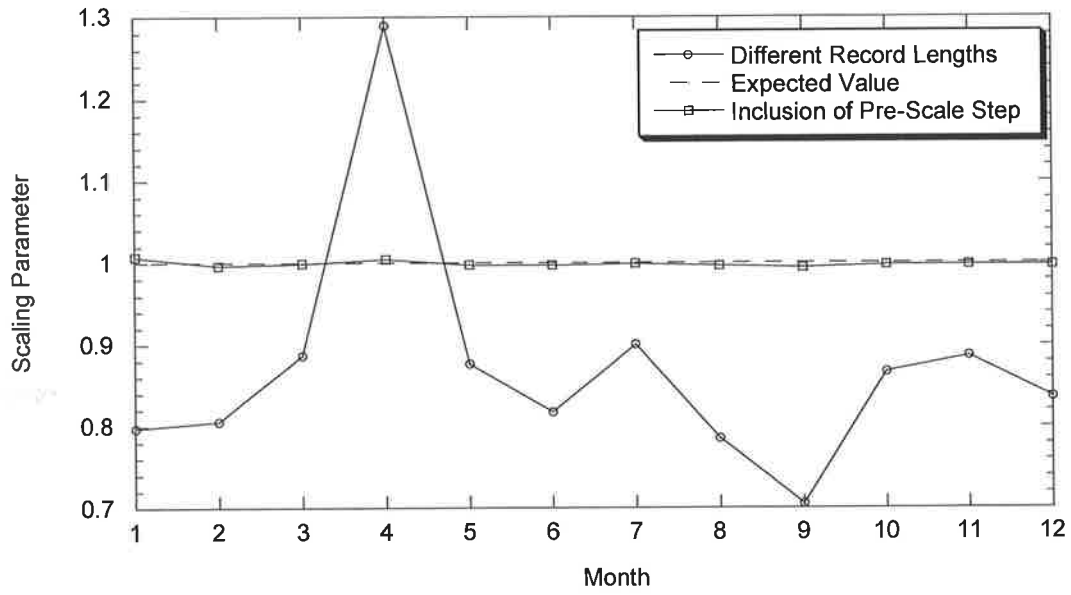

Figure 6.9: Calibration of Scaling Parameter with and without adjustment due to Different Record Periods (Data from Melbourne)

The development of a master – target framework and the introduction of an intermediate calibration step to remove the sampling variability issues associated with the non-concurrent data periods enable the master – target relationship to be adopted for each model parameter set of the event based rainfall model. The application and development of these relationships for each section of the model and the subsequent testing and simulation results are described below.

6.3 Regional Model Application to Inter-Event Times and Storm Duration Parameters

The ability to successfully reproduce the observed distributions of inter-event time and storm durations at sites with a short historical record ensures the simulation of

an accurate storm time series and provides a significant step towards successful calibration at sites with minimal historical data. The development and performance of a master – target relationship to calibrate these sections of the model is now described in detail.

6.3.1 Model Development for Inter-Event Times and Storm Durations

As previously shown the distributions of storm events and dry periods are similar and are both described by a generalised exponential distribution given by

$$F(x \mid \theta_t) = P(X \le x \mid \theta_t) = 1 - \exp[-g(\lambda y, \theta_t)] \qquad x > 0 \;\; y > 0 \;\; \lambda > 0 \qquad (6.6)$$

Introducing a master – target relationship and the aforementioned scaling parameter (λ) and intermediate scaling parameter (λ_p) into this event distribution calibration process results in

$$F(x \mid \theta_t) = 1 - \exp[-g(\lambda(\lambda_p y), \theta_t)] \qquad y > 0 \qquad (6.7)$$

As the distributions of inter-event times and storm duration are calibrated and simulated independently, separate scaling parameters (λ) and intermediate scaling parameters (λ_p) are required to be calculated for each event distribution. The development of a linear scaling factor was important as it increases the likelihood that it can be identified successfully given the limited data available with only a short pluviograph record. As model parameters are calculated on a monthly basis, the master – target relationship is also developed for each month between the corresponding event distributions. Using only a linear scaling factor ensures only a single parameter is required to be calculated for each month in comparison to the 3 parameters which are used to describe the event parameters at the master calibration site. Maximum likelihood techniques similar to those employed to fit the parameters at the master site are used to determine the pre-scale parameter λ_p and the optimum scaling parameter λ.

During simulation, a master storm event is simulated using the master calibration parameter set and is then multiplied by the appropriate scaling factor (i.e. storm duration or inter event time) to transform the simulated master event into an equivalent length of storm event at the target site.

Before the model could be applied and tested at sites around Australia, careful consideration of the regionalisation model setup is required in order to adopt appropriate methods and statistics to check model performance. In the case of a typical model calibration, a long historical data set is used during calibration. As an example, the Melbourne record contains 95 years of pluviograph storm data. To assess model performance, the model can be simulated for 95 year replicates and a direct comparison between observed and simulated data undertaken. This provides an accurate comparison between observed and simulated data because the calibration data length is long and provides a reasonable description of the rainfall processes at the master site. Additional testing using daily data or similar can be used to compare annual statistics and long term trends.

In contrast to a typical calibration, the regionalisation model uses a short pluviograph record at the target site to update the master calibration from a site with a long historical data record. Let us consider Melbourne as the master and a surrounding site y as the target with only a short 10 year record. If the regionalisation model has been successfully applied, the resultant model should be capable of simulating the rainfall events and long term rainfall statistics at site y. Extending this further, the resultant model would be equivalent to a model calibrated to 95 years of historic data (equal to that of the Melbourne master) at site y if this data was available. As a result, simulated statistics reflect a 95 year period of rainfall records at site y instead of the observed 10 years. Direct comparison between this 10 year observed record and the model 'calibrated' to 95 years of record is problematic due to the aforementioned issues relating to sampling variability at the annual scale in short historical data sets. The regionalised model calibrated to effectively 95 years of data and the observed 10 year record now encompass different periods of time.

For these reasons when applying the regionalisation technique in practice, comparisons should be made between the regionalised model and an alternate data source, usually a long daily rainfall record. This will enable checks to occur between the aggregated statistics of the simulated model and the daily record over the same time period as was used for the calibration of the master model parameters. This provides the best comparison between observed and simulated values as equivalent periods of time are compared.

For the purposes of this study and model verification, all target sites have been chosen with significant periods of data records available. This was important to ensure that adequate high resolution rainfall data was available for comparison to validate model simulations and statistics. To ensure the model tests were as 'real' as possible, the length of data at the master and target sites were adjusted to ensure the master calibration was undertaken to a subset of master data equivalent in length and concurrent with the full length target data set. Regionalisation was then tested by utilising a small sub-set of the target data for calibration while model simulations and results were compared to the full length target record. This results in some sections of the available master data being ignored during the testing process but should produce a regionalised model that is equivalent to a model calibrated to the full length target record. This also ensures that the final calibrated target model should reproduce the observed target data distributions without bias.

As an example, consider Melbourne as the master site with its record length from 1890 to 1995, and East Sale as the required target with data from 1953 – 1992. For testing purposes the Melbourne record was clipped to coincide with the East Sale data period of between 1953 and 1992. This was then treated as a full length master record for the purposes of calibrating a master parameter set. In addition, the target site at East Sale was then clipped to only a 10 year subset of the 40 year record to formulate a realistic short pluviograph record as a target data set. These two records were then used as the master & target sites during the calibration of the scaling factor (λ). Successful implementation of the master – target relationship should then produce target scaling factors which, during simulation, generate results that compare favorably to the original full length (1953 – 1992) data set at East Sale.

In order to show that the distribution of inter-event time could be successfully translated from master to target sites, a number of master sites were selected from various locations across Australia. These were then used as a basis for shifting to other shorter target sites in the area. The sites selected for master calibrations and the associated target sites are shown in Table 6.1. While the table shows the available data at the target site, for the purposes of this study only the final 10 years of target data was actually used during the calibration process. The remainder of the record was only used for comparison to model simulation results.

Table 6.1: Master and Target Rainfall Record Details (Pluviograph Model)

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	Melbourne	86071	1900	1995	0
Targets	East Sale	85072	1953	1992	190
	Ellinbank	85240	1961	1992	95
	Laverton	87031	1965	1992	20

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	Sydney	66062	1913	1991	0
Targets	Richmond	67033	1953	1993	45
	Chichester	61151	1960	1980	185

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	Adelaide	23034	1967	1997	0
Targets	Williamstown	23763	1971	1997	40
	Stirling	23785	1964	1981	15

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	Perth	9034	1946	1992	0
Targets	Esperance	9631	1963	1991	580

	Name	BOM #	Start Year	Finish Year	
Master	Brisbane RO	40214	1908	1991	0
Targets	Brisbane AMO	40223	1949	1992	10
	Kirkleagh	40318	1959	1990	70

All master calibration results are presented in Appendix B, while the complete set of target results are shown in Appendix C. For verification purposes, selected target results for Melbourne (master) – East Sale (target), Sydney (master) – Richmond (target) and Adelaide (master) – Williamstown (target) are shown and discussed throughout this Chapter.

6.3.2 Simulated Inter-Event Time Results at Selected Target Sites

The successful scaling of inter-event time to the target site has two benefits. It provides the necessary accurate description of dry periods of the target site, but less obviously, the distribution of inter-event times also has a major influence on the number of storms and therefore the resultant rainfall for a given month. In order to verify that these distributions were successfully transferred from master to target site, two comparisons were made. The distributions of inter-event time for each month for the master, scaled and target data sets were compared to ensure the distributions were replicated. These scaling factors were then used in a simulation and the mean and standard deviation of dry events for each month were compared. Undertaking these two checks ensures that the calibration of the scaling parameter was successful and that this was implemented correctly during simulation. Results are shown in Figure 6.10 for the reproduction of inter-event distribution for a calibration shifted from Melbourne to East Sale.

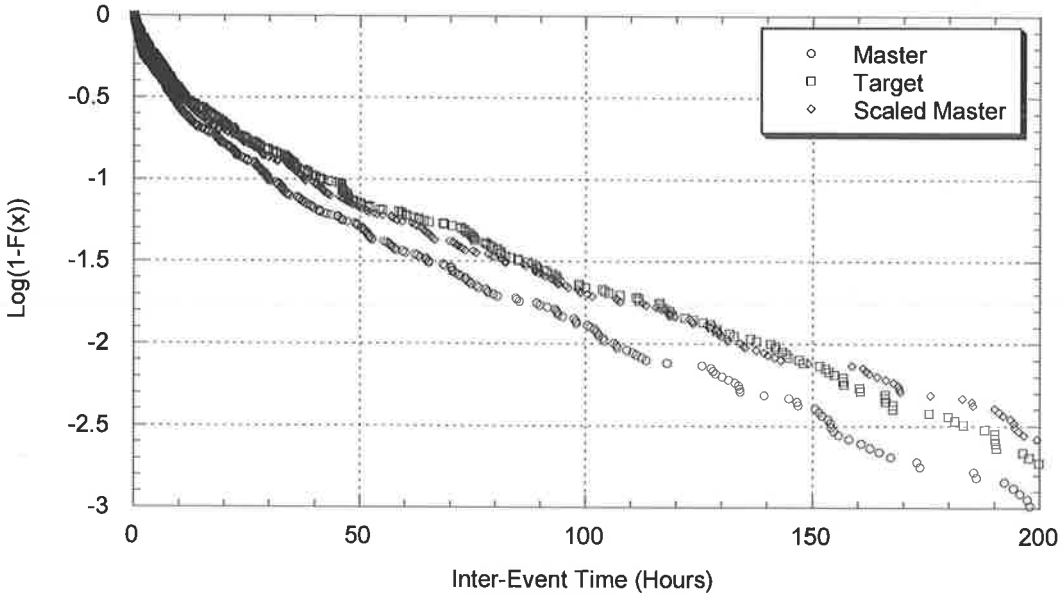

Figure 6.10: Comparison between Master, Target and Shifted Master Inter-Event Distributions (Data from Melbourne (Master) East Sale (Target), April)

As can be seen, the inter-event distributions for this month have been well represented with the introduction of the linear scaling parameter. To further test this application, simulation results for the monthly mean and standard deviation of inter-event times have been calculated and are presented in Figure 6.11 and 6.12. In these plots the observed data has been calculated from the full length target site pluviograph record. The master statistics have been identified with a dashed line linking discrete statistics to improve the visual representation of the regional scaling. The master statistics were also calculated from the full length master record.

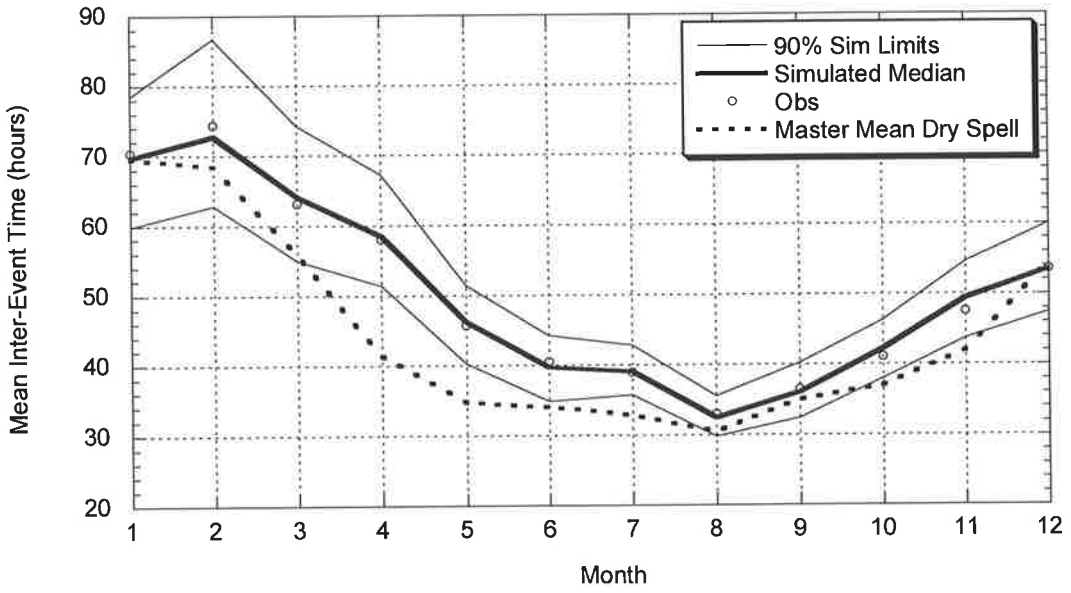

Figure 6.11: Comparison between Observed and Target Simulated Mean Inter-Event Times (Master – Melbourne; Target – East Sale)

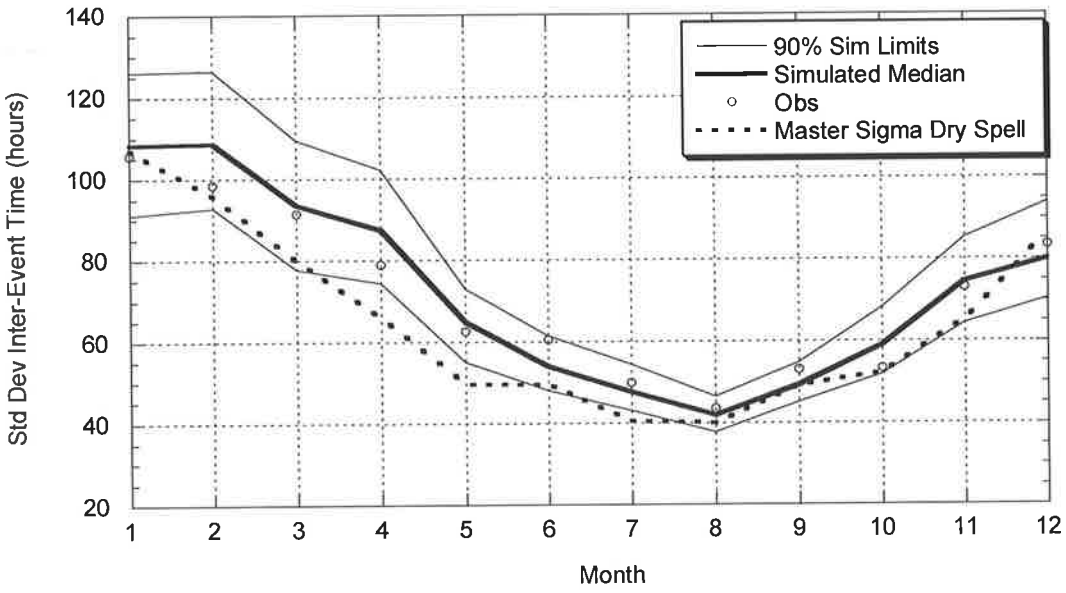

Figure 6.12: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Melbourne; Target – East Sale)

All observed statistics sit within the 90% simulation limits providing evidence that the mean and standard deviation of inter-event times have been well reproduced.

From these plots it can be concluded that the calibration and application of the regional scaling parameter have resulted in a successful inter-event shift from Melbourne to the target site of East Sale.

Further evidence is supplied by application of the model to Sydney rainfall data. In this example, data from Sydney is used as the master calibration site with Richmond RAAF base introduced as the target site. Again a simple plot of the resultant regionalised parameter distribution (Figure 6.13) suggests the model has been successfully shifted from Sydney to Richmond.

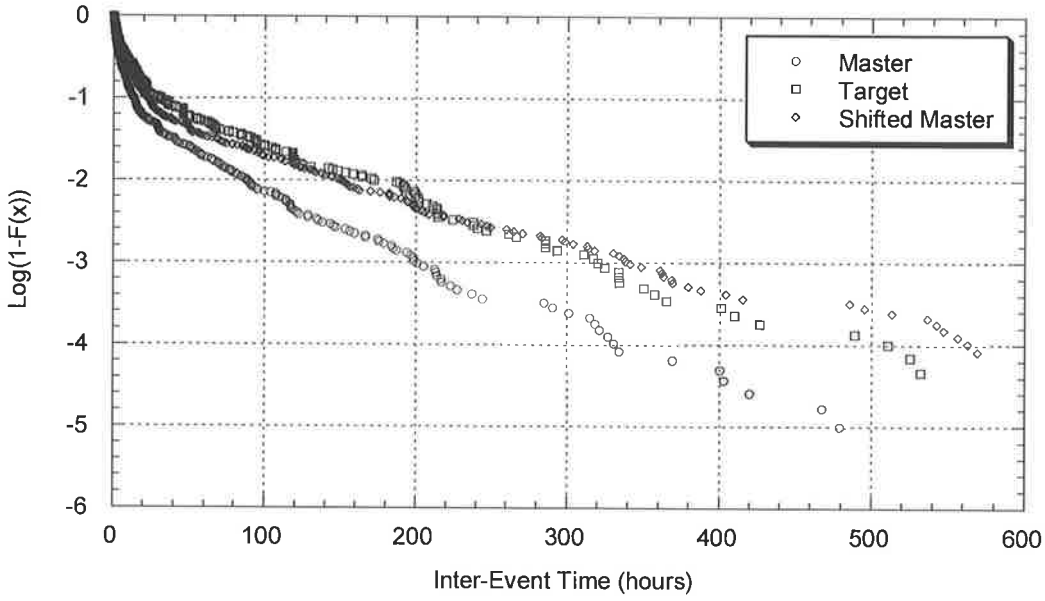

Figure 6.13: Comparison between Master, Target and Shifted Master Inter-Event Distributions (Data from Sydney (Master) Richmond (Target), June)

Resultant simulation statistics (Figure 6.14 and Figure 6.15) provide additional evidence that the model has been able to describe the required changes between the master and target site using the linear scaling parameter.

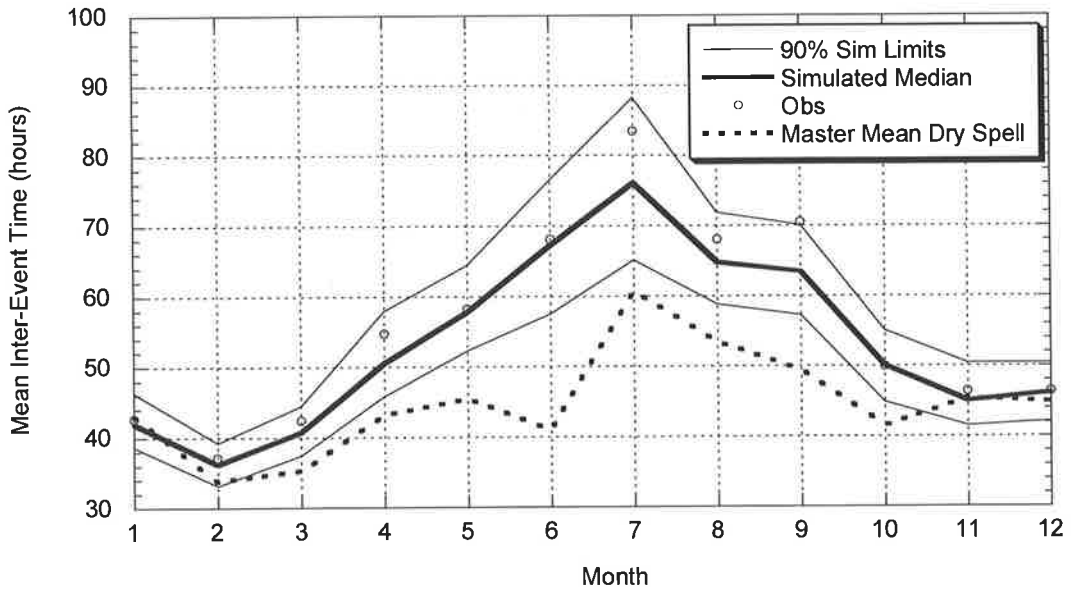

Figure 6.14: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Sydney; Target – Richmond)

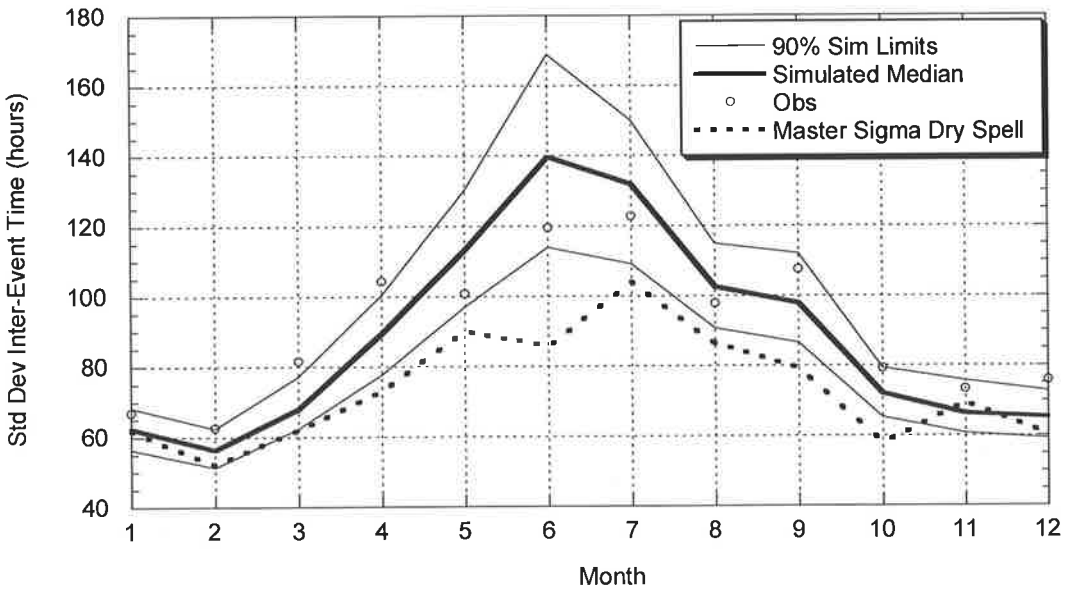

Figure 6.15: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Sydney; Target – Richmond)

The data comparisons between Melbourne-East Sale and Sydney-Richmond show a significant change between observed statistics at the master and target sites which has

been well captured by the regionalisation model. As an example the mean inter-event time for Sydney for the month of June is approximately 42 hours, with a corresponding value of 69 hours at Richmond. The model has been able to capture this significant change in inter-event time statistics while still adequately describing the distribution itself. A further test of the model arises at sites which do not have such a clear differential between master and target sites. For the regional model to work successfully it must still calibrate accurately when the variance between observed statistics at the master and target sites is minimal.

A good example of this situation is the comparison between Adelaide (master) and Williamstown (target). With the exception of February and November, the difference between the monthly mean inter-event times is minimal. Figure 6.16 presents the simulation results after a successful regionalisation shift. Not only has the model still been able to describe the required changes for February and November which required a large change, all other months have also been successfully reproduced where only a small adjustment was required.

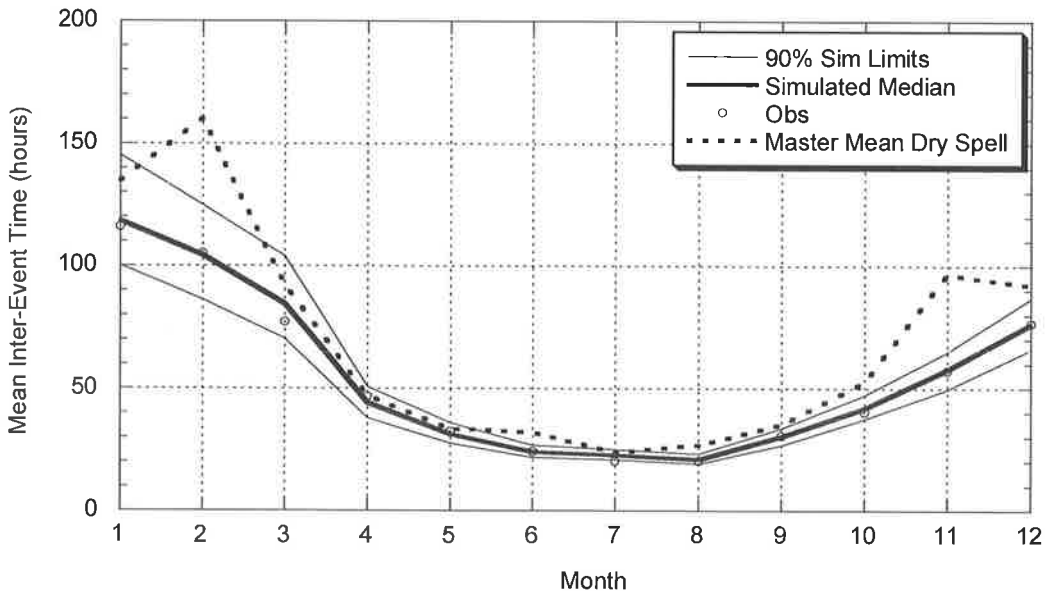

Figure 6.16: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Adelaide; Target – Williamstown)

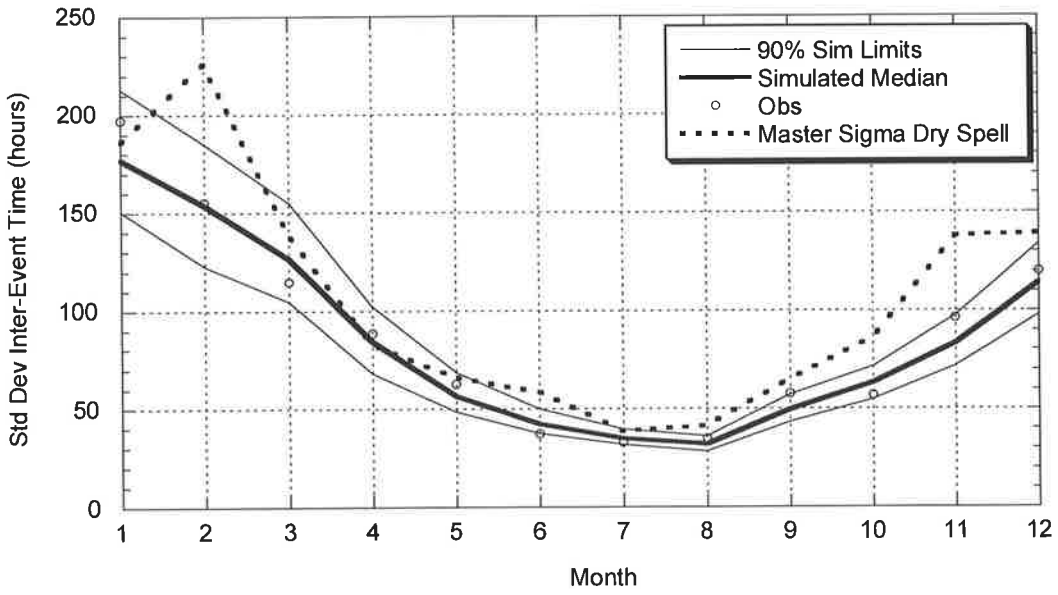

Figure 6.17: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Adelaide; Target – Williamstown)

6.3.3 Simulated Storm Duration Results at Selected Target Sites

The successful scaling of storm durations provides the necessary accurate description of the length of storm events but again in similar fashion to inter-event times, also has an influence on the number of generated storm events. The distribution of storm duration is also linked via the conditional intensity – duration relationship to storm depth and therefore it is important that this distribution is accurately reproduced by the regional model.

As the probability distributions and description of storm durations are identical to that of inter-event times, the regionalisation structure used for inter-event times could be adopted. A pre-scale factor was again introduced to remove sampling variability issues with a resultant linear scaling shift employed to describe the changes in the storm duration distributions between sites.

For consistency, results are presented for the same pairs of sites as that for inter-event times with further results available in Appendix C. Again if the observed data

statistics fell within the simulation limits this was considered a successful regionalisation of the model for storm durations.

Figure 6.18 & Figure 6.19 present simulation results from a successful shift between Melbourne and East Sale. There is significantly less adjustment required for storm duration parameters in comparison to inter-event times as can be seen with most of the master mean average storm duration statistics sitting within the simulation limits. As a result only minimal scaling was required from Melbourne to East Sale. However the model was able to adequately describe the required adjustments and successfully shift the model between these sites.

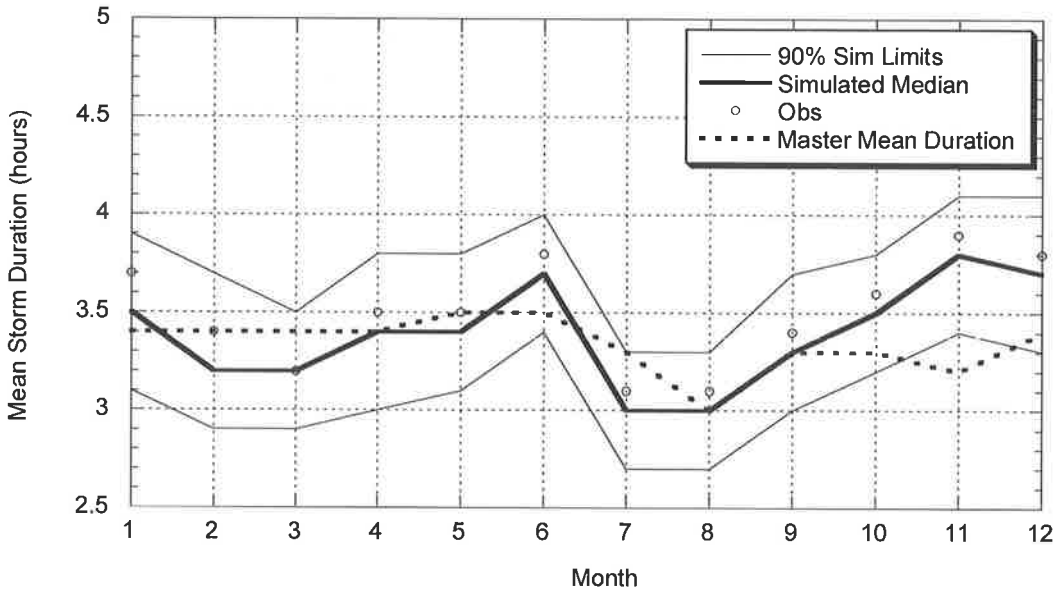

Figure 6.18: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Melbourne; Target – East Sale)

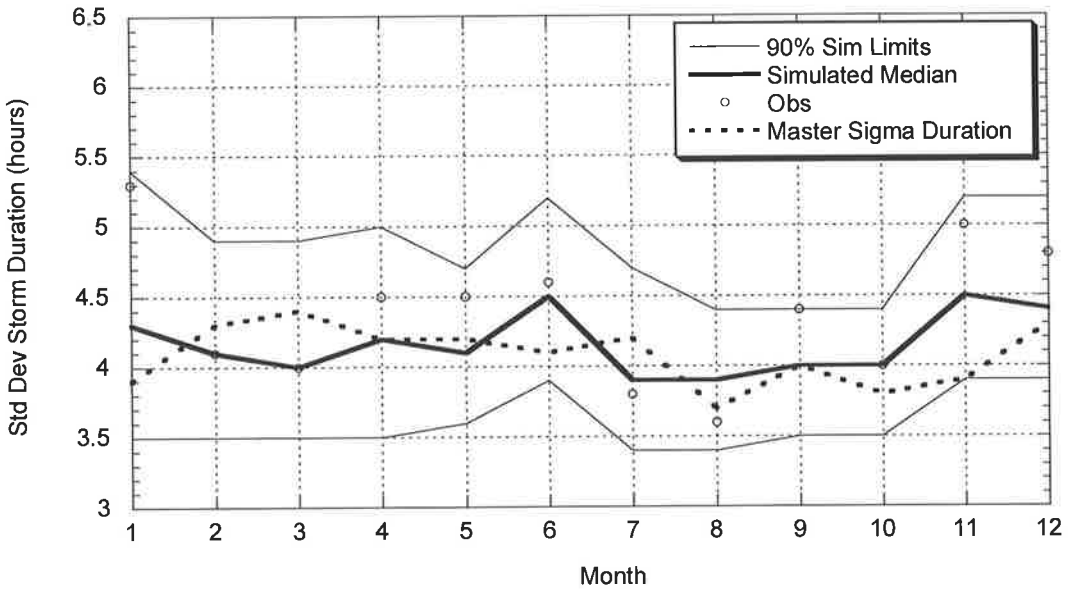

Figure 6.19: Comparison between Observed and Target Simulated Standard Deviations of Event Storm Durations (Master – Melbourne; Target – East Sale)

Similar results are evident from regional storm duration calibrations between Sydney and Richmond. The minimal adjustments have been successfully described by the incorporation of the linear scaling model.

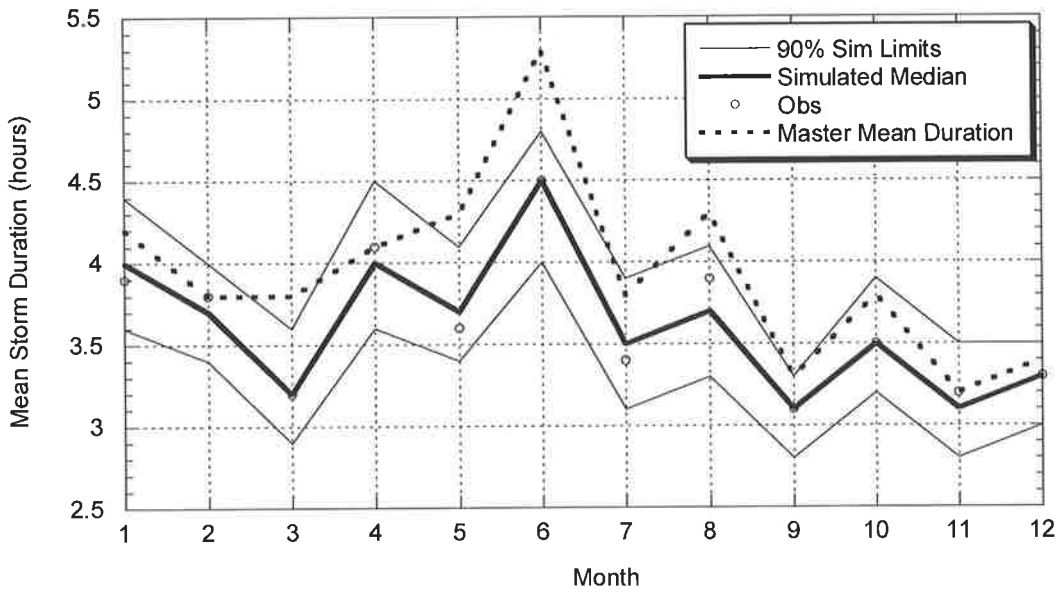

Figure 6.20: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Sydney; Target – Richmond)

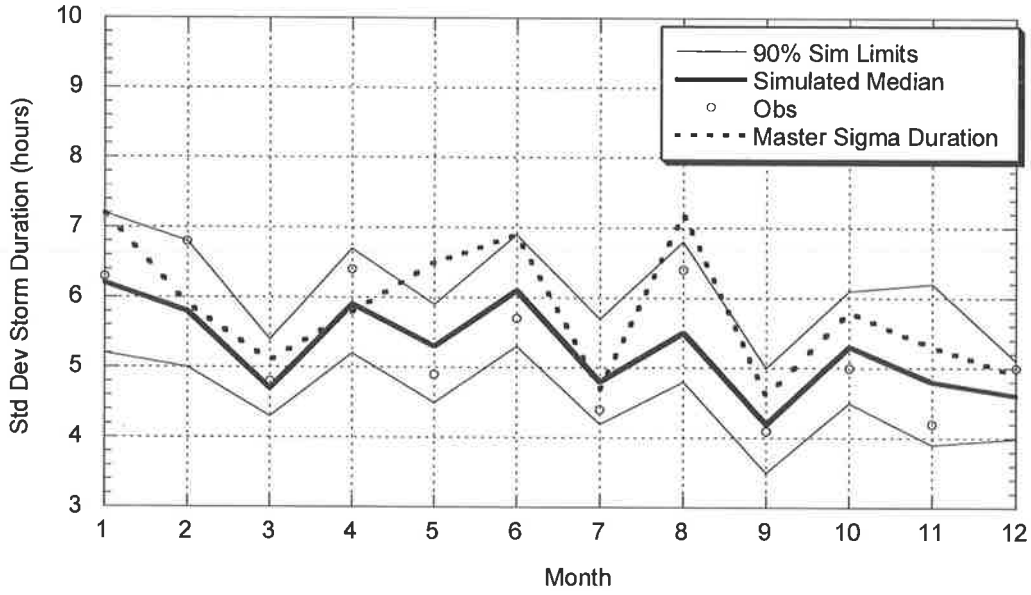

Figure 6.21: Comparison between Observed and Target Simulated Standard Deviations of Event Storm Durations (Master – Sydney; Target – Richmond)

In contrast to the inter-event times, the differences between storm duration statistics for Adelaide and Williamstown are significant. For the month of June, the mean storm duration from the full length Adelaide record is approximately 2.5 hours. In comparison the mean for the full record at Williamstown is close to 4.5 hours.

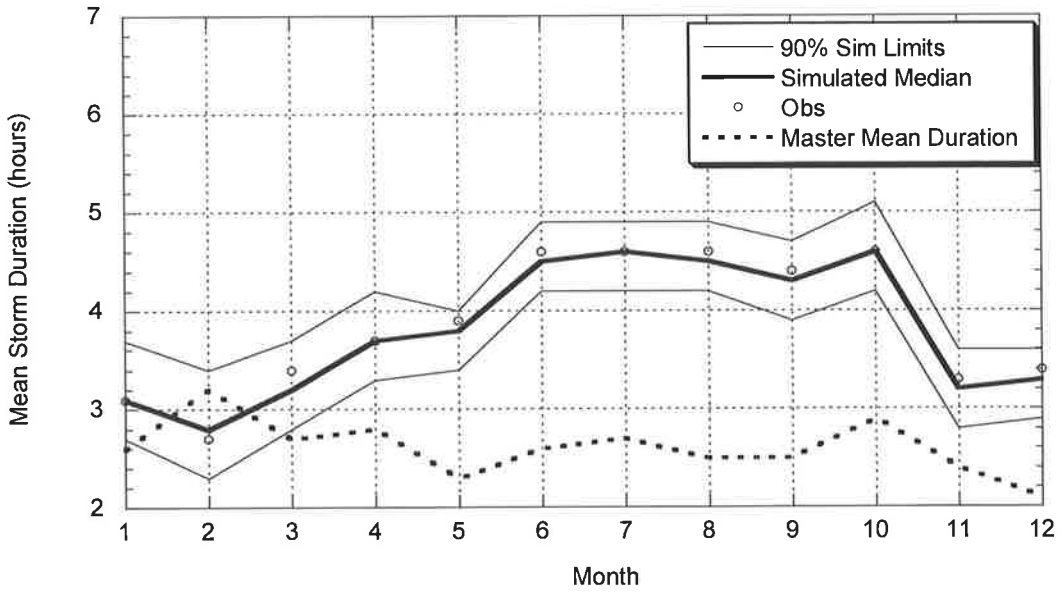

Figure 6.22: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Adelaide; Target – Williamstown)

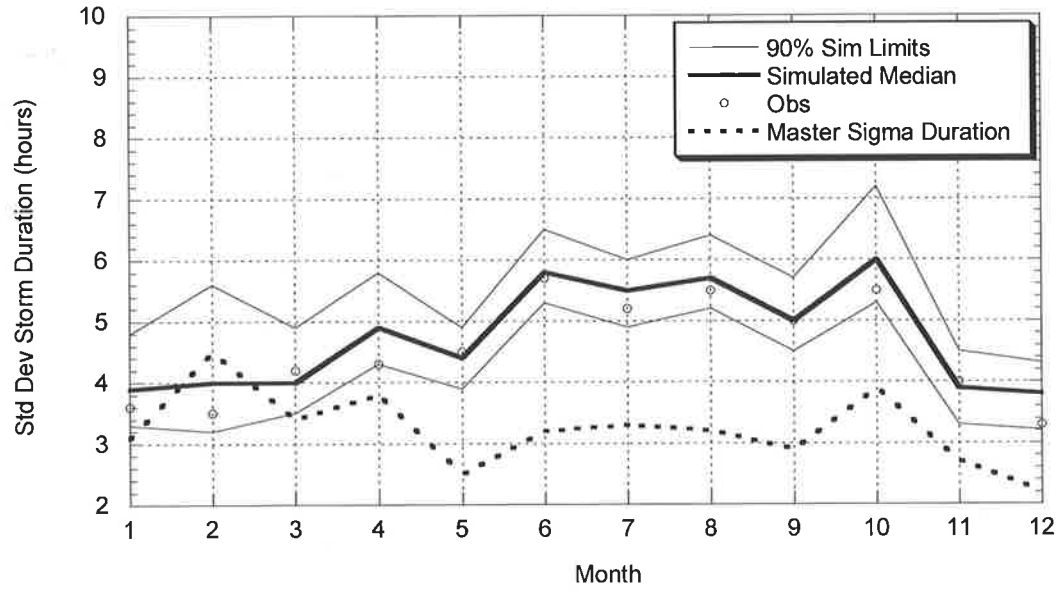

Figure 6.23: Comparison between Observed and Target Simulated Standard Deviations of Event Storm Durations (Master – Adelaide; Target – Williamstown)

It is important to remember that the two record lengths are not identical and therefore it is likely that the Adelaide record has observed significantly more short

storms which have effectively decreased the mean storm duration over time. The success of the regional model between these sites again indicates the importance of the pre-scaling step and further emphasises the ability of the regional model to capture the true adjustments required between parameter distributions at the master and target sites.

Results presented to date indicate the model has been successful in reproducing the required distribution changes between master and target sites for inter-event times and storm durations. However results presented thus far have been comparisons to calibrated statistics (i.e. the distribution of inter event times was used in the regional calibration and therefore should be well reproduced in the model simulations). To test the model further, it is important to compare the reproduction of non-calibrated statistics potentially from an alternate data source. This would also test the performance of the model when inter-event time and storm duration shifts were combined.

As the inter-event times and storm duration parameters define the number and distribution of storm events over time, it is appropriate to use the probability of observing or not observing a storm event over a given time frame as a relevant test of the models performance.

6.3.4 Comparison between Simulated and Observed Daily Dry Probabilities at the Target Site

The inter-event times and storm duration simulation within an event based model produce the storm event time series and therefore the probability of observing wet and dry events over a given aggregation period. To investigate the success of the regional model when considering inter-event times and storm durations it is therefore appropriate to compare the observed and simulated dry probability for a certain aggregation level. In addition, a major issue when using a regionalisation technique to simulate rainfall data is the ability to verify the success of the regionalised model with very little data available at the time scale of the model simulation. In general the regional model will be used at sites which do not have a long historical record for comparison. Without this comparative long 6 minute

pluviograph record at the target site, an alternative data source at a different time scale must be used to provide verification of the model output and an indication of the success of the regionalised calibration. For inter-event time and storm duration the reproduction of the daily dry probability (the probability of observing no rain in a given day) can be used and compared to the observed statistics from a long daily record at the target site to provide further model verification.

Extending this idea further, if it can be shown that the regionalised model is capable of reproducing this daily statistic after successful calibration, then there is potential for this statistic to be used as a check for a regionalised model at a site where minimal pluviograph data is available for verification. It is reasonable to assume that an event model which can successfully simulate the probability of observing a dry day is also adequately simulating the number and length of storm events and inter-event times. Therefore the reproduction of the daily dry probabilities is an important indication of the performance of the model and as it is a non-calibrated statistic provides further evidence of the models structure.

Figure 6.24 presents the probability of observing a dry day for both the observed data at the target site (East Sale), the simulated limits form the regionalised model and a dashed line representing the master site statistics (Melbourne). The master statistics are in fact the same as that of the target site data in that they are discrete points of one value per month, however they have been represented as a dashed line for display purposes as it was difficult to distinguish between the master and target data points.

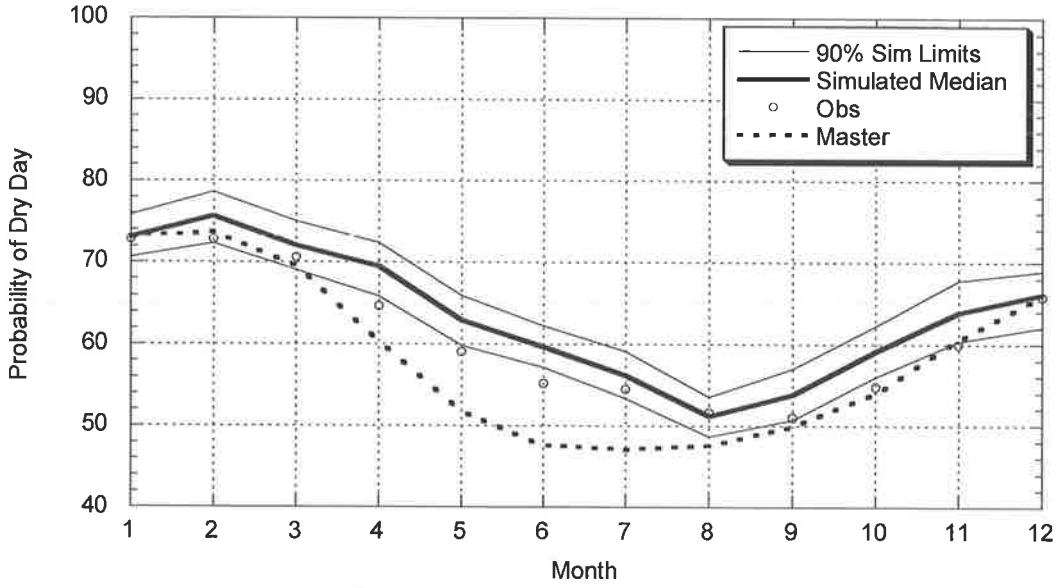

Figure 6.24: Comparison between Observed and Target Simulated Probability of a Dry Day (Master – Melbourne; Target – East Sale)

Figure 6.24 indicates that the model has been able to improve the reproduction of the daily dry probability for all months. The majority of points sits within or just outside the simulation limits suggesting that the model has been able to adequately reproduce the daily dry probability at the target site of East Sale. Figure 6.25 and Figure 6.26 present similar results from Sydney and Adelaide respectively. The model regionalised from Sydney to Richmond shows a substantial shift between the master and the resultant simulation at the target site and the target values have been well reproduced.

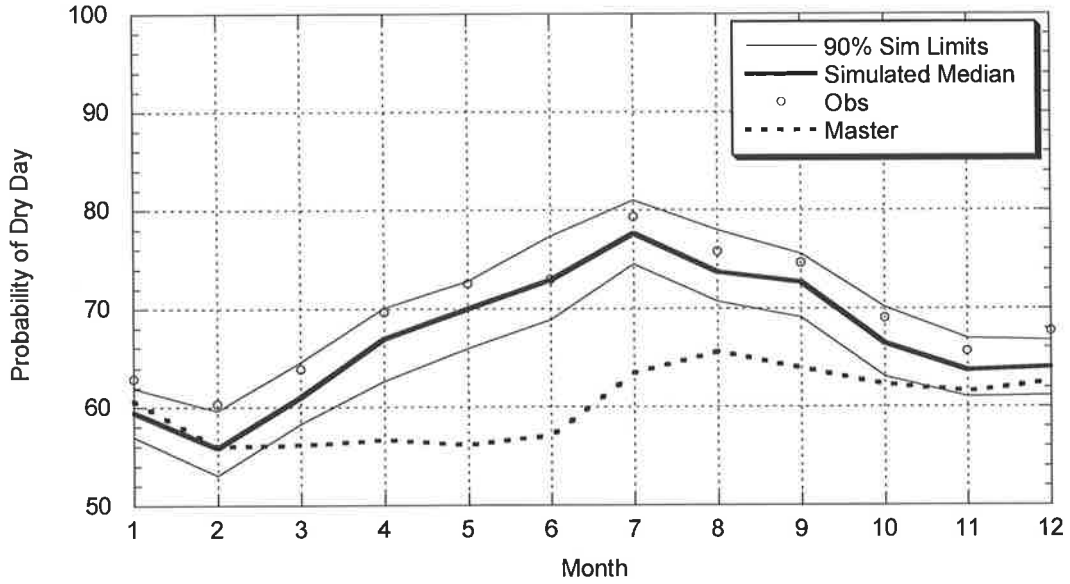

Figure 6.25: Comparison between Observed and Target Simulated Probability of a Dry Day (Master – Sydney; Target – Richmond)

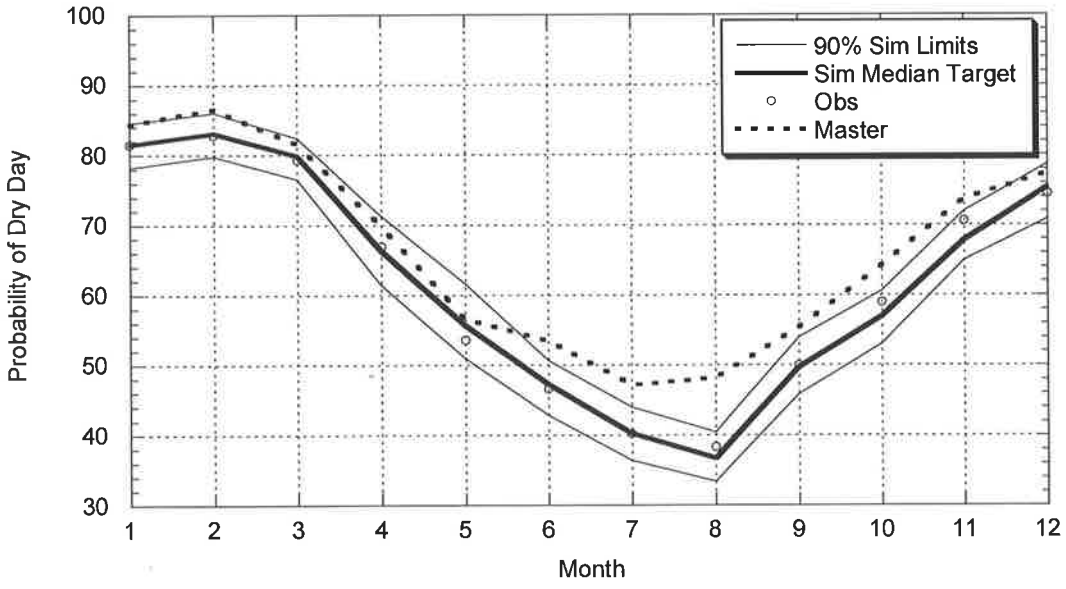

Figure 6.26: Comparison between Observed and Target Simulated Probability of a Dry Day (Master – Adelaide; Target – Williamstown)

The data from Adelaide required a less significant change, however again the model was able to capture the required changes when shifting from master to target. Combining these results with those presented in Appendix C, the regionalisation of inter-event times and storm durations using a distribution scaling factor and a master – target relationship has been successfully introduced into the rainfall model and can be applied with confidence to sites with short pluviograph records.

This was also an important result for the further development of the regionalisation model as can be seen in the later chapters, as it enabled the development of a technique using only daily rainfall data further generalizing the technique and improving the application of the model across Australia.

6.4 Regional Model Application for Storm Event Depths and Temporal Pattern Parameters

The successful simulation of bulk rainfall amounts and their subsequent disaggregation into the required time step is the final step in the rainfall model process. With a successful process developed and tested to calibrate the distribution of storm durations and inter event times at the target site, a similar process was required to ensure the reproduction of rainfall distributions and aggregated totals was accurate. As was the case with the distribution of Inter-Event times and Storm Durations, the calibration of storm event depths and temporal pattern disaggregation parameters are undertaken independently within the model and as a result the potential links and relationships between these parameterisations at the master and target sites were also investigated separately and are described below.

6.4.1 Model Development for Storm Event Depths

An important outcome from the earlier work presented in Chapter 3 was the investigation into the conditional intensity – duration relationship. Chapter 3 demonstrated the similar shape of the conditional intensity – duration relationships between adjacent rainfall sites. For example sites around Sydney all display the same basic intensity – duration shape and parameter characteristics. This result was important for the work in Chapter 3 as it provided the basis for a new calibration

procedure for the storm event depth model parameters. For the purposes of this Chapter it is also important as it provides the base for the development of a regionalisation process for these rainfall amounts.

While the results of Chapter 3 indicated that the intensity – duration parameter characteristics are similar and consistent between sites, the relationship is not so identical as to require little or no adjustment. Some sites receive consistently more intense rain events than others and as a result a regionalisation adjustment for differences in the intensity of rain events between sites is required. It was initially hoped that a simple scaling of the marginal (as opposed to the duration conditional) intensity distribution between the master and target sites could successfully describe any required adjustments. Not only would this have provided a simple approach, it would also tie in nicely with the adopted process for inter event times and storm durations. Unfortunately this was not the case.

An investigation into the potential of a scaling factor applied directly to the marginal intensity distribution provides an insight into why such a simple approach did not work. In the original model the use of the conditional intensity-duration relationship enables the parameters which govern the distribution of storm intensity to be influenced by the storm duration. As a result a short duration storm will have a different mean intensity in comparison to a longer duration storm. This is to be expected and was a requirement of the original model to be able to replicate observed statistics at various time scales. If the marginal distribution of intensity is investigated directly for the purposes of calibrating a scaling factor, this conditional link to storm duration is ignored and the accuracy of the regionalisation model suffers as a result.

The application of a scaling relationship directly between both conditional intensity – duration distributions at the master and target sites was also problematic. The relationship would be required to take into account changes to both storm durations and storm intensities between sites explicitly in the calibration of event depths in order to keep the conditional relationship intact. This resulted in scaling parameters at each of the breakpoints in the model, introducing a large number of parameters

requiring calibration each month at the target site, which was not appropriate given the lack of calibration data available.

The adopted regionalisation model considers the conditional intensity – duration relationship implicitly by comparing the distribution of storm depths rather than storm intensities in order to calibrate the regionalisation requirements between a master and a target site. By developing the model to compare storm depths instead of storm intensity, the conditional intensity – duration relationship is taken into consideration implicitly as the data pairs of intensity and duration are considered together. Almost as importantly, because the process uses the distribution of event depths rather than storm intensity, the process could be (and subsequently was) developed further to be used with daily rainfall data with the view of further generalising the model. This additional work to regionalise the model with only daily data at the target site is presented in Chapter 7.

The regionalisation model for event depths works in a similar manner to that for storm durations and inter-event times. Using the master site calibration the model initially simulates a distribution of storm event depths at the master site. (These are simply a product of simulated storm duration and corresponding simulated storm intensities) The model then calculates a scaling relationship (λ_{Depth}) between the master (D_m) and target depth (D_t) distributions, i.e.

$$P(D_m) = \lambda_{Depth} P(D_t) \tag{6.8}$$

In order to calculate the required regionalisation for the storm intensity distribution specifically, the model can use the previously described storm duration regionalisation relationship (λ_d) (which can be calculated first independently) to pre-translate the storm durations from master to target sites. If we let i = intensity and d = duration then

$P(D_m) = \lambda_{Depth} P(D_t)$; now $P(D) = P(i \mid d) * P(d)$ so

$$P(i_m \mid d_m) P(d_m) = \lambda_{Depth} P(i_t \mid d_t) P(d_t) \tag{6.9}$$

If storm duration regionalisation is pre-calculated, then

$P(d_t) = \lambda_d P(d_m)$ so substituting into (6.9) gives

$$P(i_m \mid d_m)P(d_m) = \lambda_{Depth}P(i_t \mid d_t)\,\lambda_d P(d_m) \tag{6.10}$$

And simplifying

$$P(i_m \mid d_m) = \lambda_d \lambda_{Depth}P(i_t \mid d_t) \tag{6.11}$$

Any resultant regionalisation relationship determined when comparing the event depth distributions becomes solely a factor of the storm intensity.

Unlike the scaling factors for inter-event times and storm duration, a simple constant linear factor was not adequate to describe the required distribution changes between master and target sites. Investigation into the storm depth distributions indicated that a more complex relationship was required. Figure 6.27 displays the comparison between storm event depth distributions (plotted on a log scale) at Melbourne (Master) and East Sale (Target). It is clear that there is a more pronounced difference between sites for storm depths around the 0.9mm mark as opposed to storm events of greater and lesser depths.

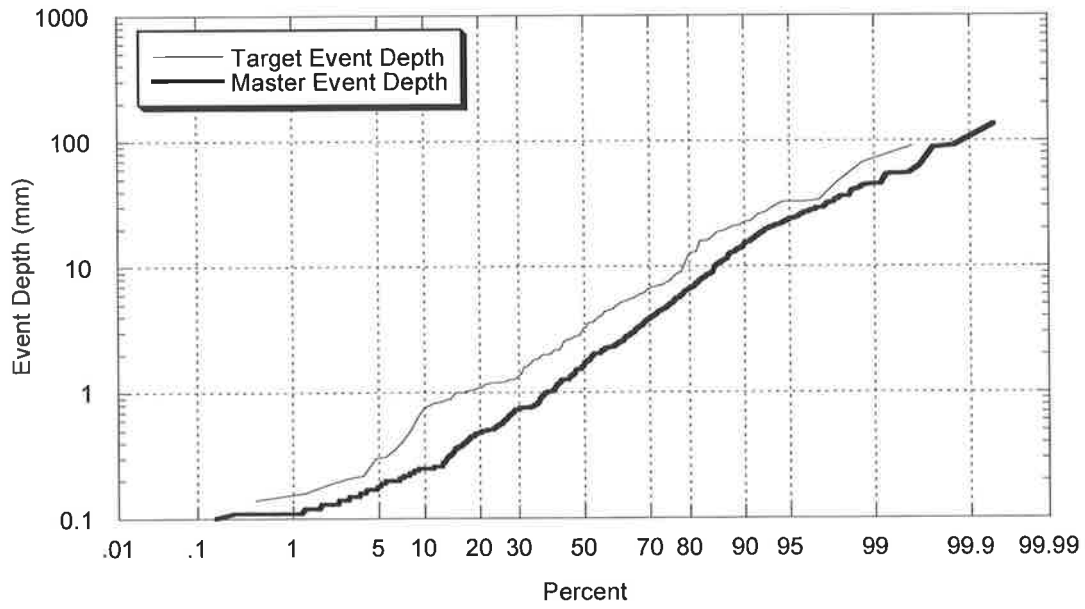

Figure 6.27: Comparison between Storm Event Depth Distributions Data for July, (Master – Melbourne; Target – East Sale)

Figure 6.28 displays a similar result for data from Adelaide (Master) and Williamstown (Target). In this case there is a decreasing difference between sites as the storm depth increases which could not be captured with a simple linear scaling factor.

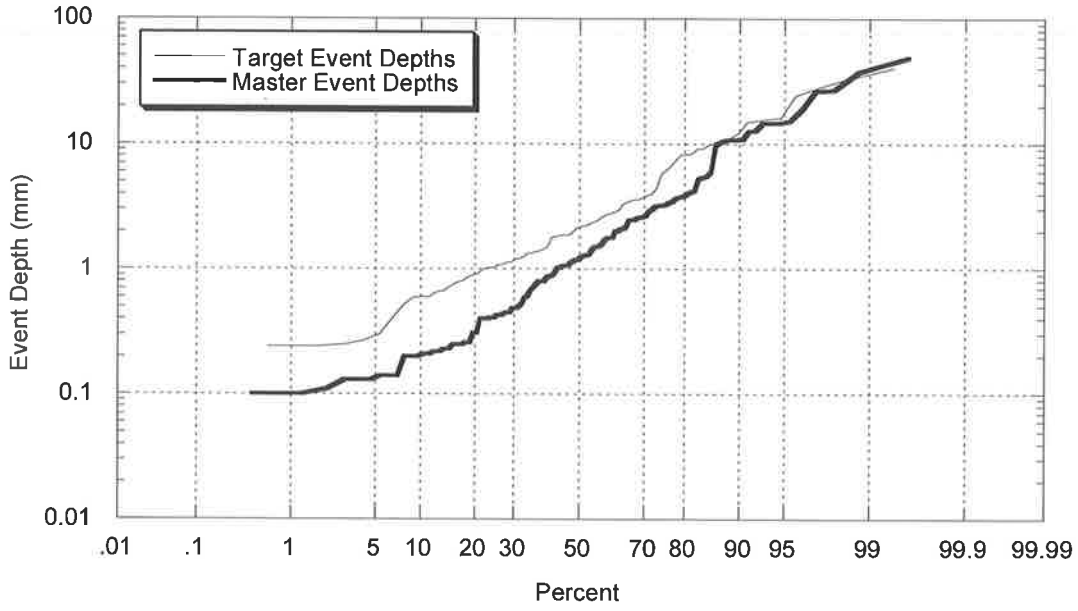

Figure 6.28: Comparison between Storm Event Depth Distributions Data for March, (Master – Adelaide; Target – Williamstown)

The comparisons between master and target storm depth distributions led to the development of a relationship for the scaling factor λ_{Depth} which is dependent on storm event depth. This relationship between λ_{Depth} and storm event depth requires the calibration of 3 parameters. The first is a constant linear scaling factor similar to that used in the calculation of λ for inter-event times and storm duration and is denoted $\lambda_{Depth\text{-}1}$. In addition to this constant factor, two additional parameters determine the location ($\lambda_{Depth\text{-}2}$) and size ($\lambda_{Depth\text{-}3}$) of the triangle peak which provides the regionalisation model for event depths a further degree of freedom. This additional freedom ensures the model is able to obtain a better fit between the master and target sites. A schematic of the model structure is displayed in Figure 6.29 below.

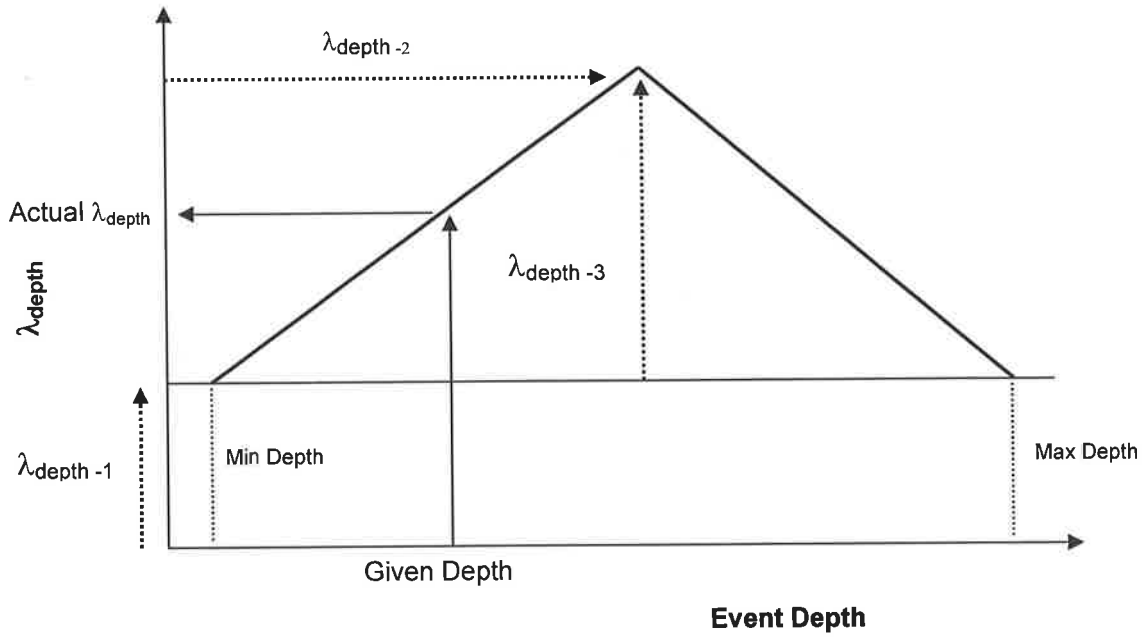

Figure 6.29: Schematic of Scaling Parameter Relationship to Storm Depth

The actual scaling factor during calibration and simulation for a given storm depth is determined by the combination of the three scaling factors as shown in Figure 6.29. In this way the scaling factor has enough freedom to capture the required differences between depth distributions at the master and target sites while still restricting the number of parameters to 3, 15 less than is required for a direct event intensity calibration to pluviograph data.

This approach provides a robust methodology to regionalise the storm event intensity within the model. However, in contrast to the earlier work when regionalising storm durations and inter-event times, there is not an explicit description of the storm event distributions within the model. In order to utilise a maximum likelihood approach to calculating the required scaling parameter between depth distributions, a description or estimate of the storm depth distributions was also required.

6.4.2 Incorporating Non-Parametric Kernel Smoothing Density Estimation

To eliminate the need to incorporate an assumption about the probability distribution of event depths into the model, the probability density was approximated using the method of kernel density estimation or kernel smoothing. This, coupled with a maximum likelihood approach formed the basis of the calibration of event depths at the target site and removed the need to introduce further assumptions into the model structure.

Kernel smoothing is a well-known and accepted method for the non-parametric estimation of probability densities. Rosenblatt (1956) introduced the idea of kernel estimators by proposing to smooth kernel weights on each of the observations. Since then, kernel estimating has been used in numerous applications including hydrologically in the estimation of flood quantiles (Adamowski, 1985,1989,1996,2000; Guo, 1991; Moon *et al.*, 1993 among others). A good introduction to kernel density estimation can be found in Silverman (1986), while Wand and Jones (1995) provide an account of more recent developments. Kernel density estimation is an extension of the histogram, providing a smooth continuous density estimate. The density estimation equation takes the form

$$\hat{f}_X(x) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x - y_i}{h}\right) \tag{6.12}$$

where K() is a kernel function centred at each data point y_i, x is the data value at which the probability density estimate is required and h is known as the bandwidth. The bandwidth sets the degree of smoothing or influence that each individual kernel has on the overall density estimate. The kernel K() must be a probability density function which by definition must have an area under the kernel function equal to one. Often, as was the case in this study, the Gaussian kernel is used

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \tag{6.13}$$

Introducing the required scaling factor λ into the kernel smoothing equation (6.12) yields

$$\hat{f}_X(x) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x - (\lambda y)_i}{h}\right) \qquad (6.14)$$

In this case n denotes the number of data points at the target site while $(\lambda y)_i, i=1,..,n$ is the target depth data. These values of $(\lambda y)_i$ are a product of storm duration t_d, the corresponding duration scaling λ_d, the corresponding target event intensity and the required depth scaling λ.

Figure 6.30 presents a graphical description of the kernel smoothing approach. At the location of each scaled target depth data point a Gaussian kernel is placed. The summation of the n kernels of bandwidth h centered at each target observation $(\lambda y)_i$ in (6.14) forms the kernel probability density estimate of a specific master depth x. At locations with a greater concentration of target data points, there will be more contributions from a number of Gaussian kernels centred at these points giving a greater density estimate to the master data point x and providing the ability to compare the two data distributions.

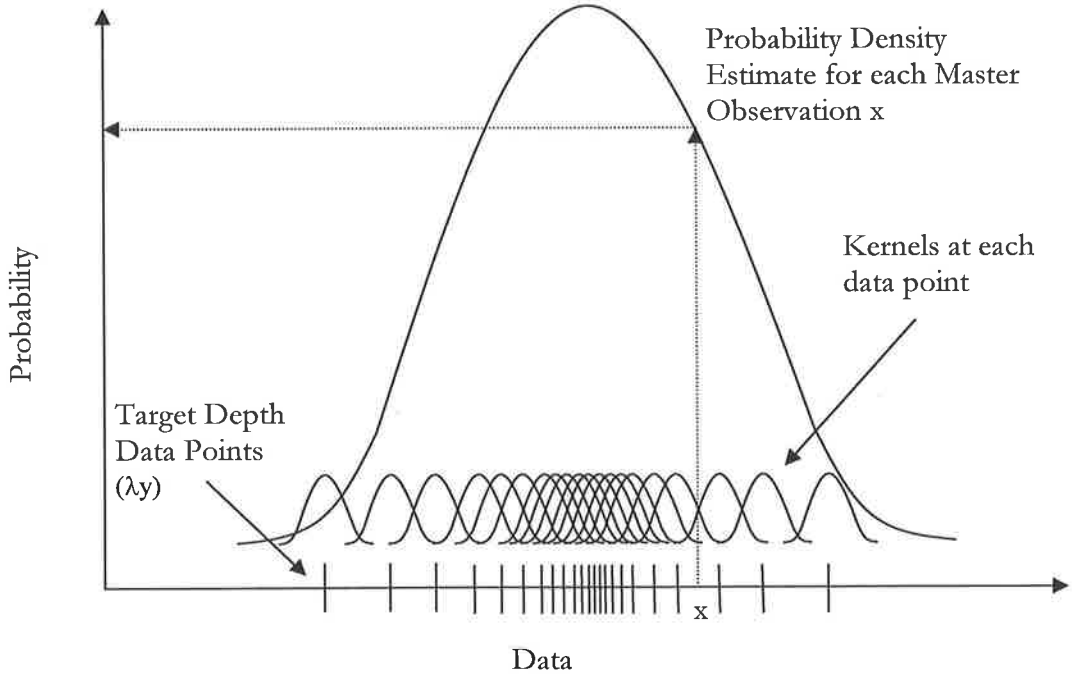

Figure 6.30: Schematic of Kernel Smoothing Calculation

Maximum likelihood techniques are again used to find the optimum value of λ. Given that λ_d is constant (it has been pre-calculated during the storm duration regionalisation work), λ is the only parameter requiring adjustment to produce the maximum likelihood. An iterative process was used whereby the kernel density estimation was recalculated for each iterative change in scaling parameter λ. As is the case in all maximum likelihood approaches, the best fit occurs when the product of $\hat{f}_X(x)$ (or the sum of $\log \hat{f}_X(x)$ as is the case in this research) for all master observations is at a maximum. This ensures that the scaled target depth distribution and the observed master depth distribution are easily compared with the differences described by the scaling parameter λ. Once the optimal scaling parameter is found, the final target simulation is produced using the master parameter set and this resultant scaling factor to provide the required storm event depths at the target site. (The calibration approach has calculated the scaling parameter at the target site to best fit the target depth distribution to the master depth distribution; hence during simulation the simulated master depths must be scaled by the inverse of this parameter to produce the required depth values at the target site.)

6.4.3 Storm Event Depths Results - Introduction

To verify the successful scaling of simulated storm depths two comparisons were used. First, the distributions of event depths from the observed master, target and simulated target can be compared to provide a check that the regionalisation calibration for event depths was successful. Secondly, comparisons to the annual and monthly rainfall distributions provide evidence that the model has captured the bulk rainfall processes successfully over various scales. The observed statistics for this comparison were calculated from the observed daily record at the target site to provide the most accurate monthly and annual statistics. As the distribution of event depths was the basis of the calibration process, the comparison to the non-calibrated monthly and annual rainfall distributions are of great benefit in validating the model structure and performance. An additional comparison between simulated and observed IFD curves provides further validation of the storm depth model output;

however, the IFD curves are also influenced by the internal storm event characteristics and the temporal pattern parameters which are discussed in Section 6.4.6. As a result, IFD comparisons are provided later in Section 6.4.7.

6.4.4 Simulated Storm Event Depth Distribution Results at Selected Target Sites

As the calibration of the scaling parameter for event depths takes into account the event depth distributions the following plots give an indication of how well the model is able to capture the differences between the master and target sites, the expectation being that the observed data will predominantly rest within the simulation bounds. As can be seen below and in Appendix C, this is mostly the case suggesting that the regionalisation model is capable of describing the differences between distributions at pairs of sites. In the selected cases where the observed values did not rest within the bounds, the points are either on the borderline or just outside. Considering the calibration data available and the accuracy of all other reproduced parameters, this is an acceptable result.

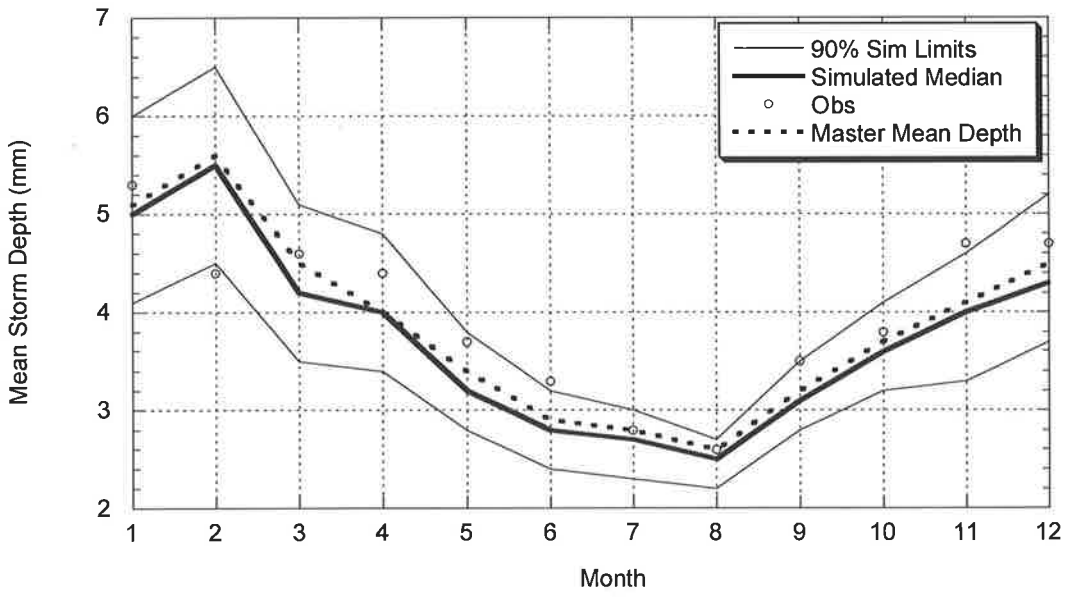

Figure 6.31: Comparison between Observed and Target Simulated Average of Event Depths (Master – Melbourne; Target – East Sale)

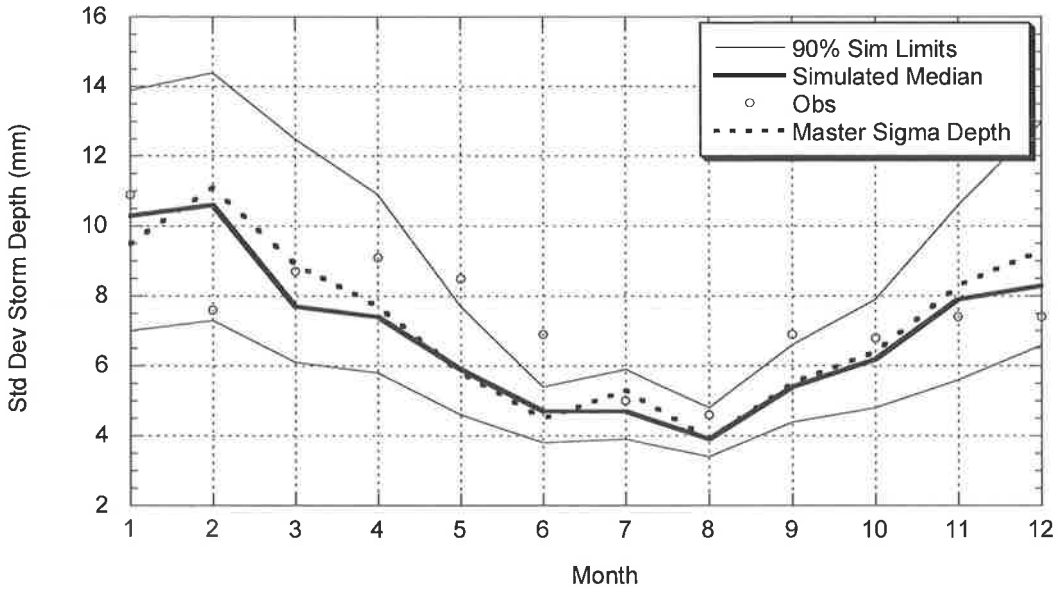

Figure 6.32: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Melbourne; Target – East Sale)

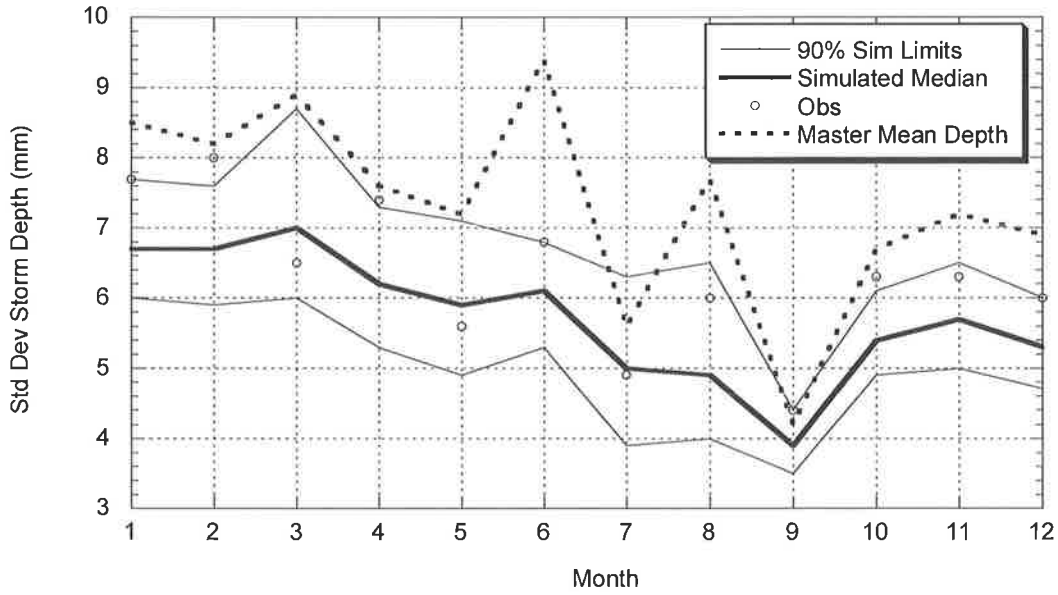

Figure 6.33: Comparison between Observed and Target Simulated Average of Event Depths (Master – Sydney; Target – Richmond)

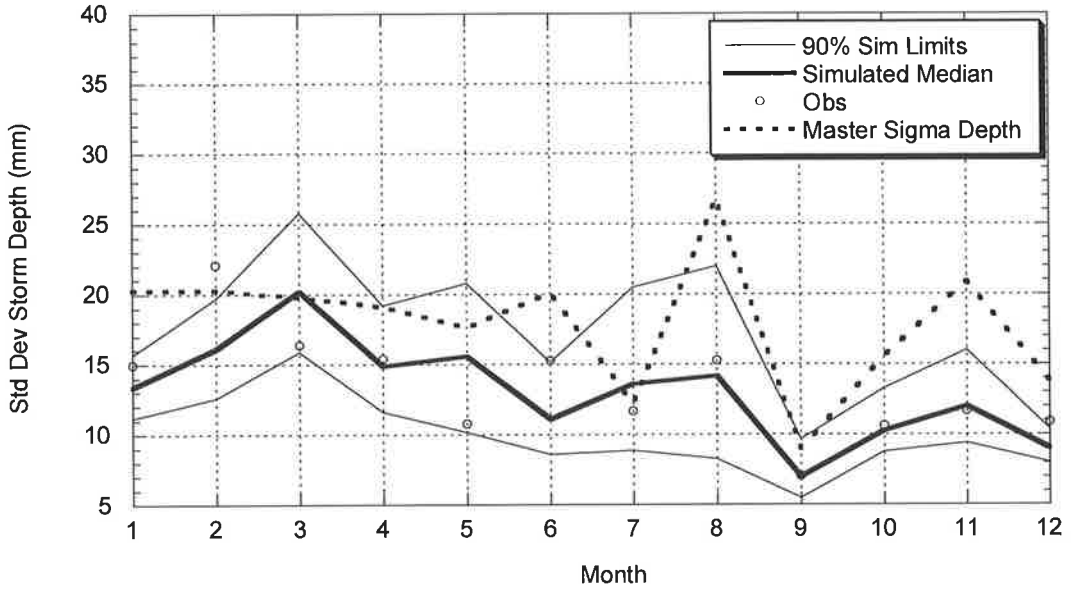

Figure 6.34: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Sydney; Target – Richmond)

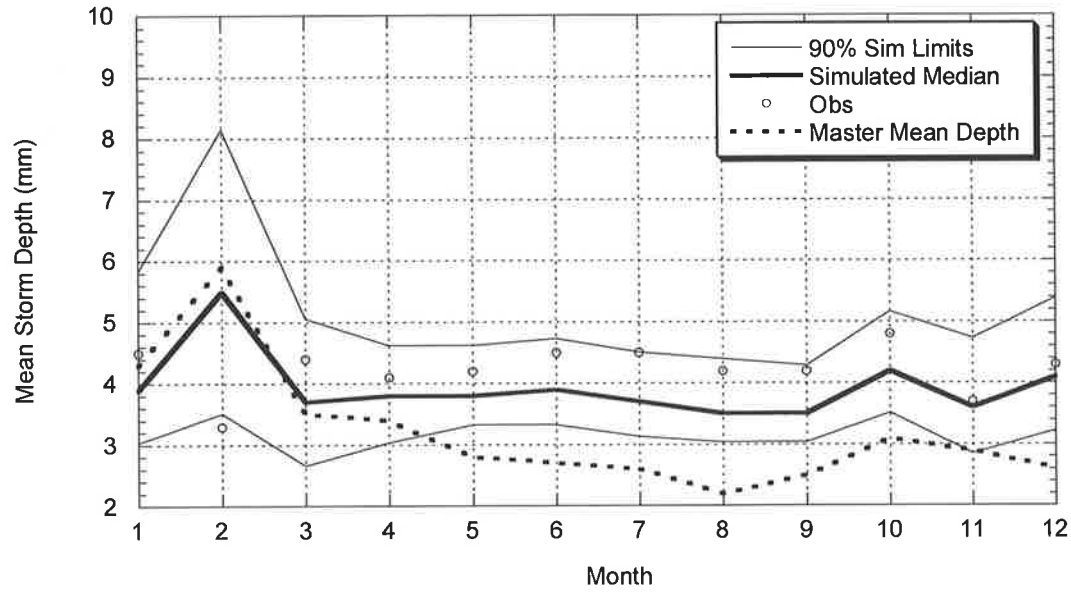

Figure 6.35: Comparison between Observed and Target Simulated Average of Event Depths (Master – Adelaide Airport; Target – Williamstown)

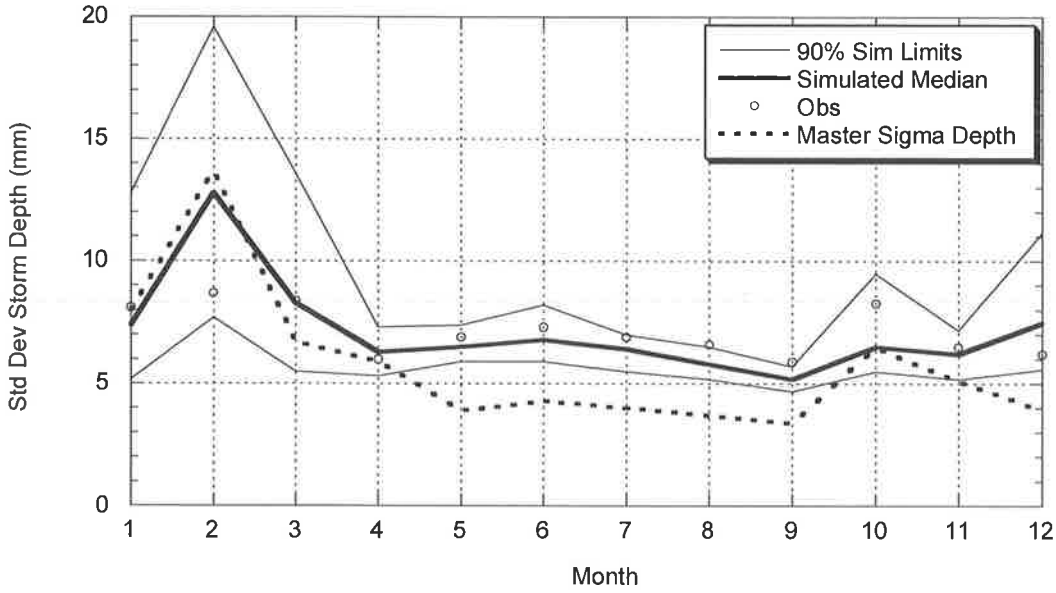

Figure 6.36: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Adelaide Airport; Target – Williamstown)

6.4.5 Monthly and Annual Rainfall Results

To further test the performance of the regionalisation process, the ability of the model to replicate statistics not used during the calibration process was investigated. Reproducing monthly and annual rainfall totals is important for hydrological risk assessment models and are two statistics not used during the calibration process. Their use as a verification tool is further enhanced by the fact that monthly and annual rainfall totals were sourced from daily data records, a completely different data set to the pluviograph records used during the calibration process.

Three pairs of sites were selected for presentation in this section, based on the differences between the master and target rainfall totals. In the first case the master is Adelaide Airport while the target site is Williamstown. Adelaide Airport has an annual rainfall 451mm in comparison to that at Williamstown of 717mm. As can be seen from Figure 6.37 the model has been able to successfully reproduce the mean annual rainfall when scaling from master to the target site. However, consistent with the original Heneker *et al.* (2001) rainfall model the simulation has underestimated the long term persistence structure at the annual time scale. This is not a drawback of

the regionalisation process but rather of event based models generally which lack an inter-annual persistence structure to capture long term climatic effects. Concurrent PhD work by Andrew Frost at the University of Newcastle to identify and introduce a persistence structure into rainfall models of this type should in the future provide a way to improve the reproduction of the standard deviation of annual rainfall in event based models. Nevertheless, the presented result indicates the ability of the regionalisation model to translate between master and target sites.

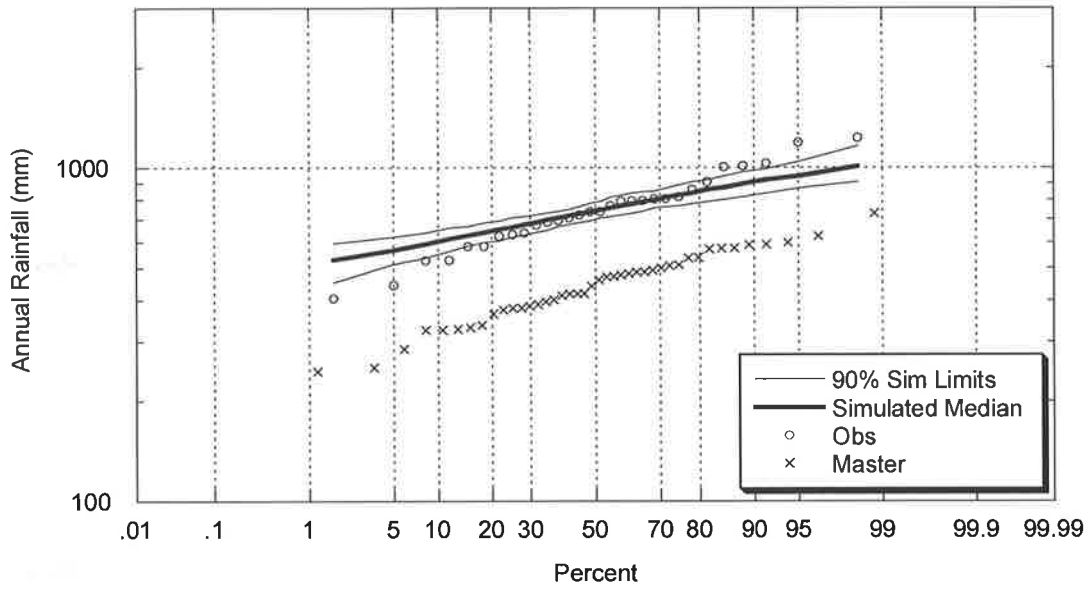

Figure 6.37 Comparison between Observed and Target Simulated Annual Rainfall (Master – Adelaide Airport; Target – Williamstown)

Figure 6.38 displays simulated and observed monthly rainfall for January again demonstrating the models ability to simulate and reproduce non calibrated statistics.

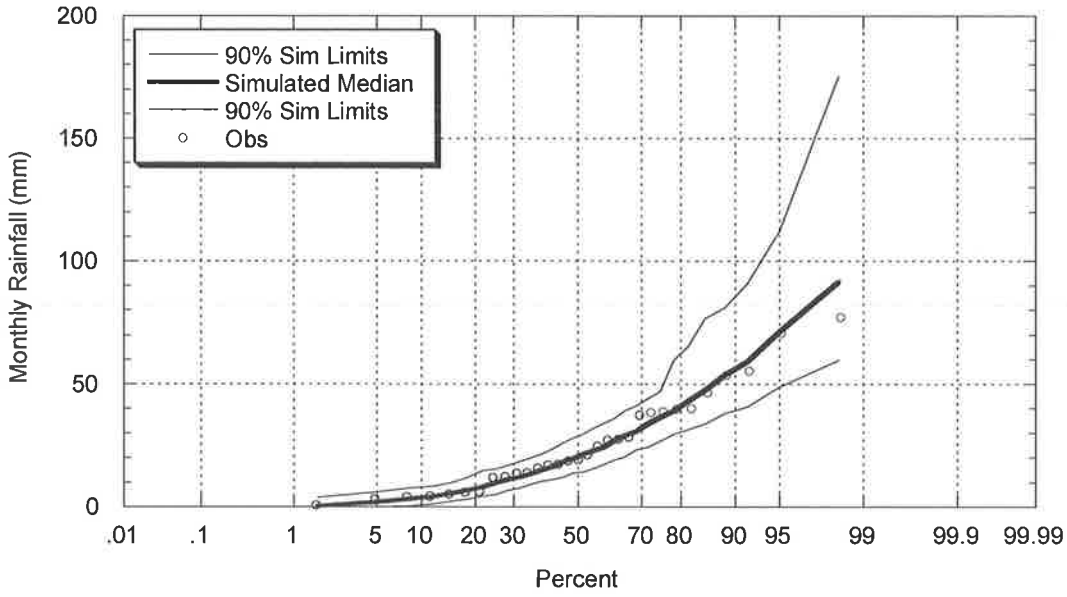

Figure 6.38: Comparison between Observed and Target Simulated January Rainfall (Master – Adelaide Airport; Target – Williamstown)

Figure 6.39 and Figure 6.40 present data from the target site at Kirkleagh in comparison to the master data from Brisbane. These sites have annual mean rainfall which is closer than that between Adelaide Airport and Williamstown but the model can still be seen to successfully reproduce the desired distributions.

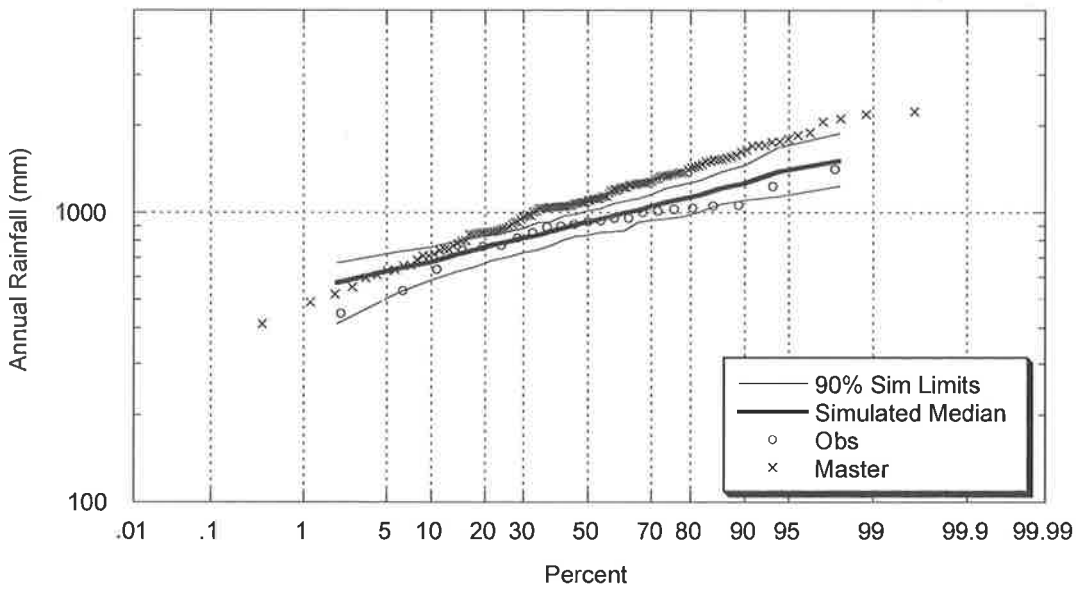

Figure 6.39: Comparison between Observed and Target Simulated Annual Rainfall (Master – Brisbane Regional Office Airport; Target – Kirkleagh)

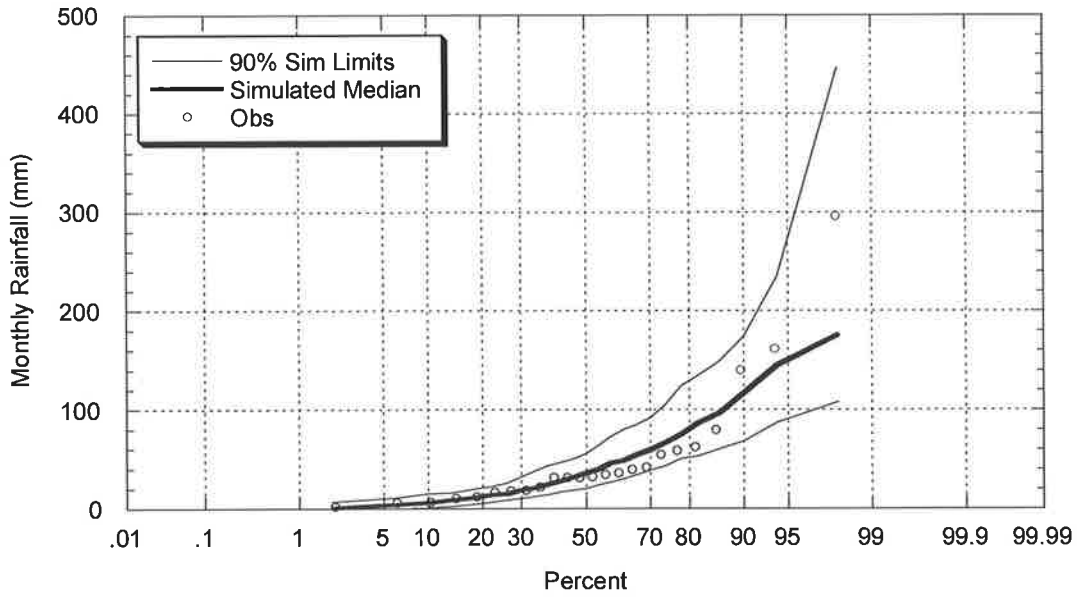

Figure 6.40: Comparison between Observed and Target Simulated July Rainfall (Master – Brisbane Regional Office Airport; Target – Kirkleagh)

Finally Figure 6.41 & Figure 6.42 demonstrate the ability of the model to simulate successfully when there is little change between the master and target sites. Both Sydney and Chichester have similar annual rainfall totals (1217mm and 1295mm respectively), but the model has still captured and reproduced the mean annual rainfall. (Again the standard deviation of annual rainfall has been underestimated)

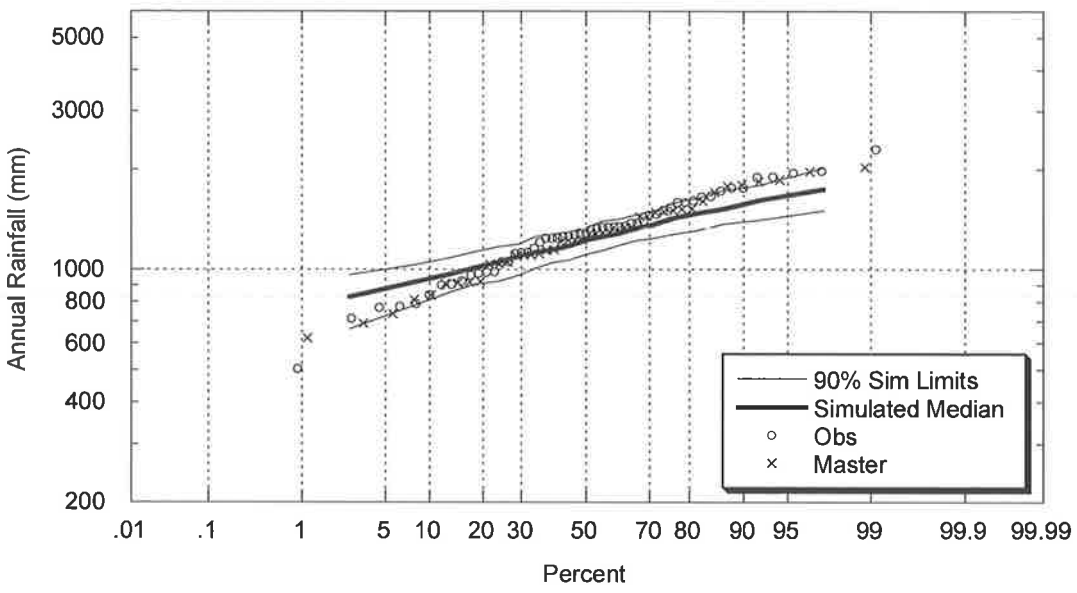

Figure 6.41: Comparison between Observed and Target Simulated Annual Rainfall (Master – Sydney; Target – Chichester)

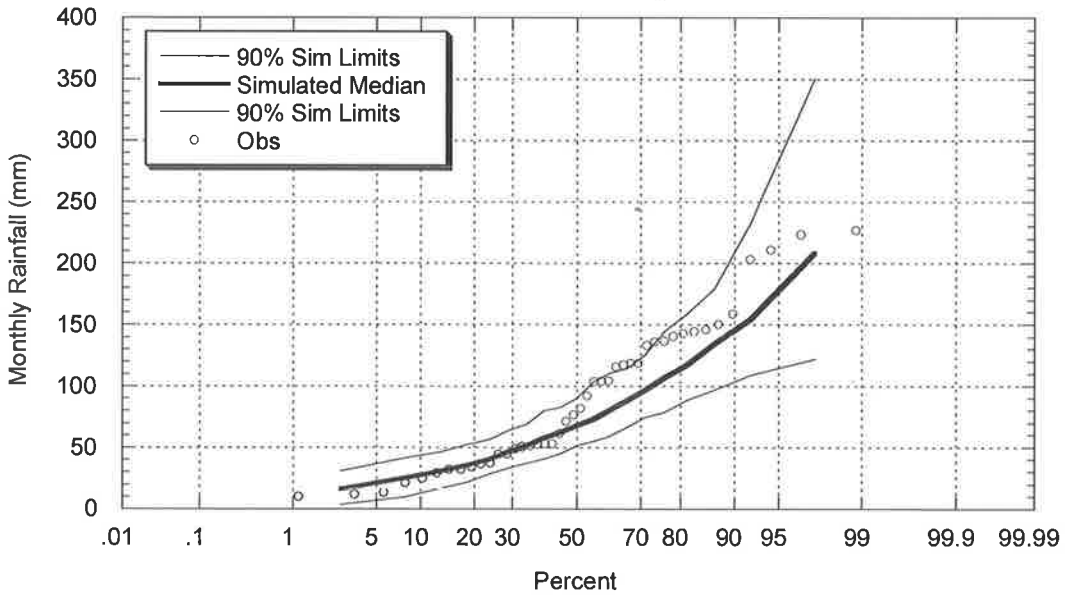

Figure 6.42: Comparison between Observed and Target Simulated November Rainfall (Master – Sydney; Target – Chichester)

6.4.6 Model Development for Storm Temporal Pattern

Once the gross storm characteristics have been successfully simulated at the target site, the temporal pattern model is still required to disaggregate these events to six minute rainfall totals. As discussed previously, the rainfall model uses a disaggregation process expressed as a constrained random walk within non-dimensionalised duration-depth space. Progression of a rainfall trace through this space is characterised by a sequence of jumps described by a truncated log-normal distribution. The parameters that define this log-normal distribution, namely the mean and the standard deviation are conditional on the current position in the non-dimensionalised space and are given by

$$m=(1-\delta_{i-1})(m_1+m_2\delta_{i-1}) \tag{6.15}$$

$$s=(1-\delta_{i-1})(s_1+s_2\delta_{i-1}) \tag{6.16}$$

where $\delta = \frac{d(t)}{d(t_d)}$

The distribution of internal rainfall jumps used during the disaggregation process is therefore predominantly a conditional relationship based on what has occurred previously throughout the storm. When comparing the temporal pattern parameters between two rainfall sites, it is the changes in these distribution parameters that require consideration for regionalisation. In order to understand the potential differences from site to site, a scatter plot of the parameter values against non-dimensionalised duration can be produced for a number of sites. Figure 6.43 displays a typical single site master calibration from Melbourne data.

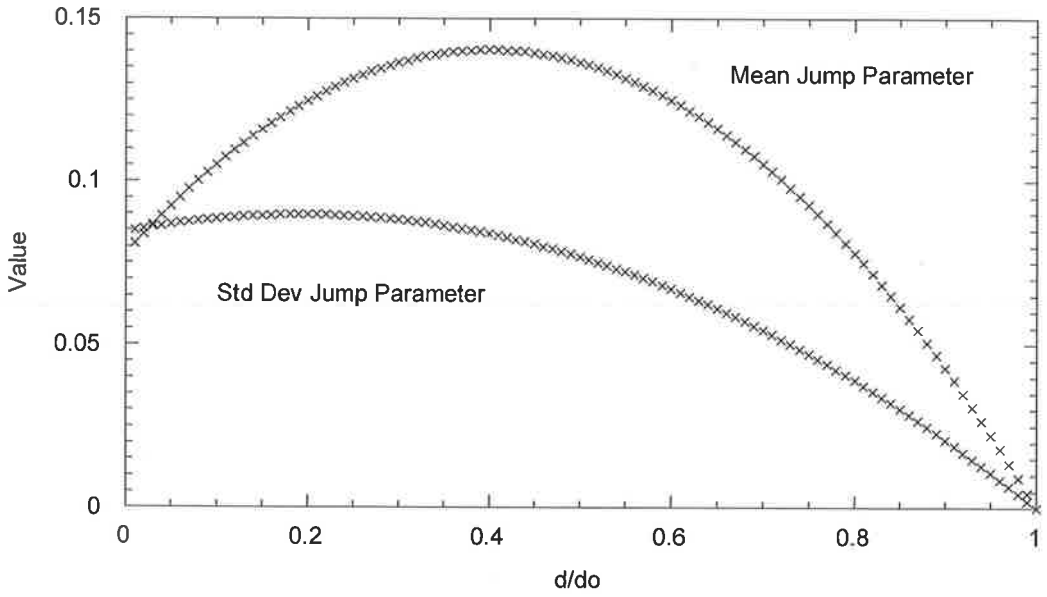

Figure 6.43: Temporal Pattern Parameters (Calibrated to Melbourne Data)

The temporal pattern parameters can be considered as a simple summary of the typical changes in rainfall intensity over time at a given site. If we compare two sites which could be considered similar in terms of the type of rainfall that generally occurs, then it would be logical that the corresponding temporal pattern parameters should remain relatively consistent between these sites. As an example it would not be out of the ordinary to expect that sites within the tropics would exhibit different rainfall patterns to those in temperate climates. To investigate the possibility that rainfall sites within a similar climatic region have similar temporal pattern parameters, a number of sites were calibrated and the temporal pattern parameters at these sites compared.

Once the temporal pattern parameters were calculated for adjacent sites, they were co-plotted to give an insight into potential parameter differences. Figure 6.44 compares the temporal pattern parameters calibrated with data from Sydney, Adelaide, Melbourne and Brisbane. The data used for co-plotting across all sites was adjusted to be the same length and period to ensure that all sites had the chance to experience similar or related weather patterns. While Figure 6.44 displays a similar shape between the parameters at these sites, there are obvious differences in the temporal structure at these sites. For instance Brisbane experiences significantly higher mean intensity bursts and is much more variable than the rainfall experienced

in Melbourne, Sydney or Adelaide. Given Brisbane experiences a tropical climate in comparison to Melbourne, Sydney and Adelaide, this result is feasible.

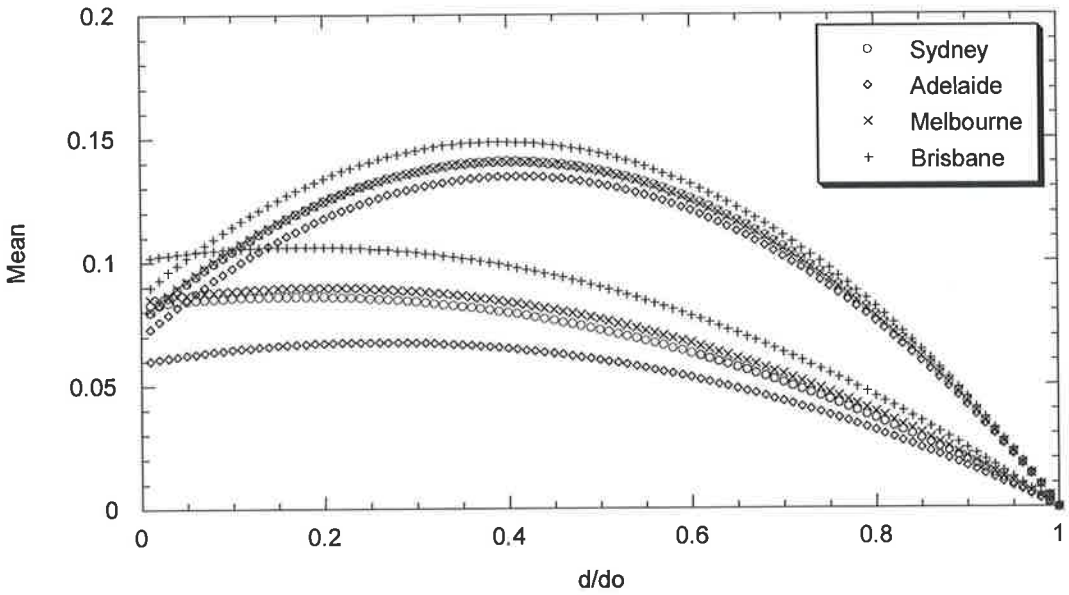

Figure 6.44: Australian City Temporal Pattern Parameters

More significant for the purposes of regionalisation is the comparison between temporal pattern parameters for sites in a specific state or region. Figure 6.45 to Figure 6.48 shows the extent to which these temporal pattern parameters remain consistent when comparing regional/state wide sites.

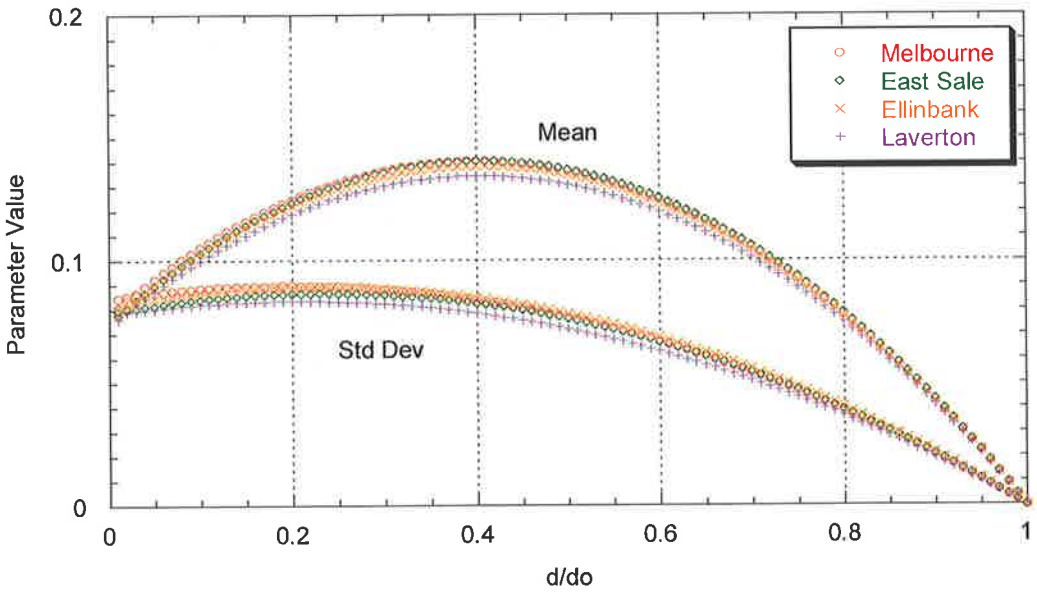

Figure 6.45: Southern Victorian Temporal Pattern Parameters

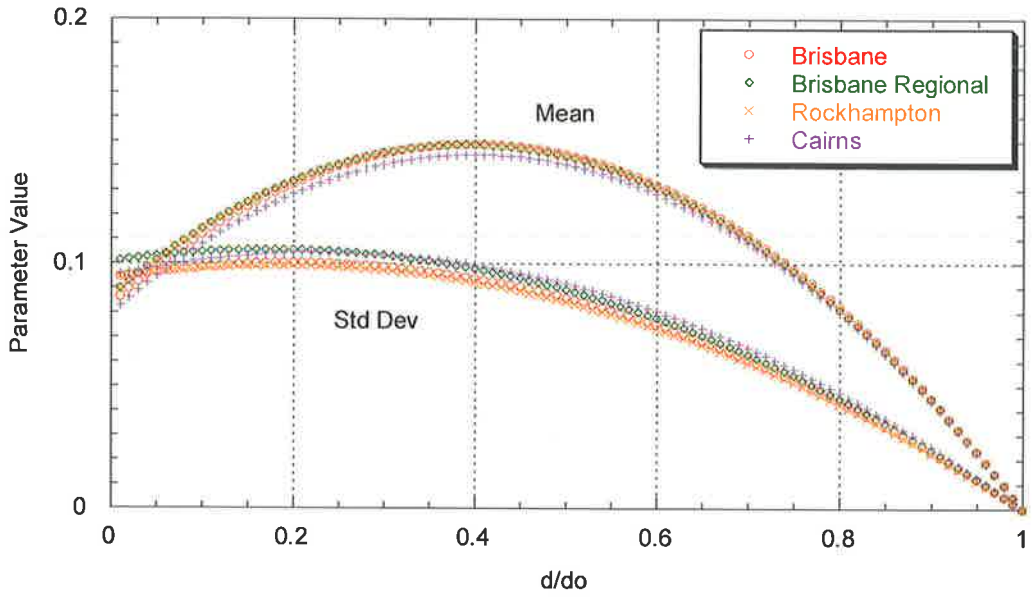

Figure 6.46: Queensland Temporal Pattern Parameters, East of the Divide

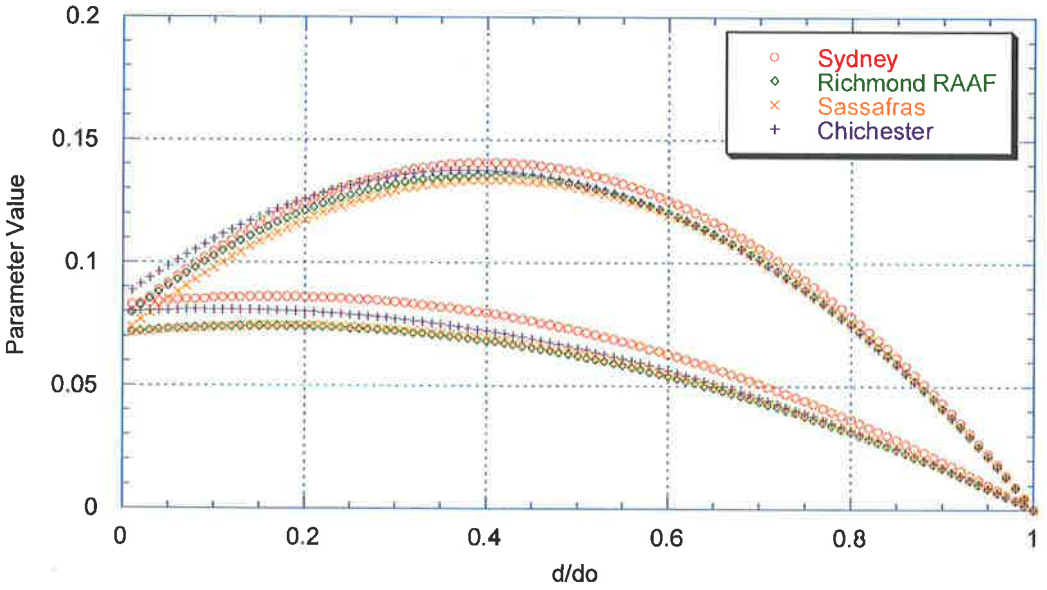

Figure 6.47: New South Wales Temporal Pattern Parameters, East of the Divide

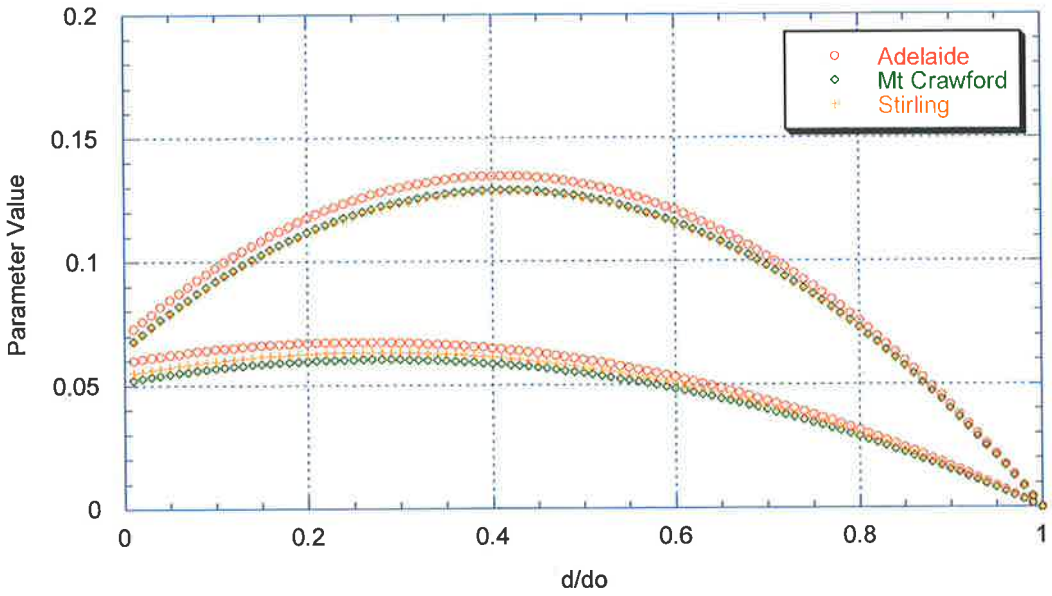

Figure 6.48: Adelaide Temporal Pattern Parameters

Significantly these plots show a high level of consistent results. Not only is there a self-similar process that was initially exploited to describe the progression of rain traces within a storm, these results indicate that the overall process is also similar for groups of sites within similar climatic regions across Australia. Given an understanding of the rainfall variation across these sites and the ability of the model to describe the temporal rainfall patterns, this expected result suggested that disaggregation parameters for the target sites could be adopted without adjustment from the master calibration site. This result was further verified during simulation as presented below and removed a significant hurdle in the regionalisation process. Provided the bulk storm characteristics can be successfully generated at the target site, the internal storm intensity characteristics can be accurately represented.

6.4.7 Intensity Frequency Duration Curve Results

The comparison between observed and simulated IFD curves at the Target site provides a further check on the performance of the regionalisation model. If the comparison is favourable then the decision to keep the temporal pattern parameters constant between master and target sites can be further justified. As presented below

in Figure 6.49 to Figure 6.51, the internal storm characteristics have been successfully reproduced at the target site.

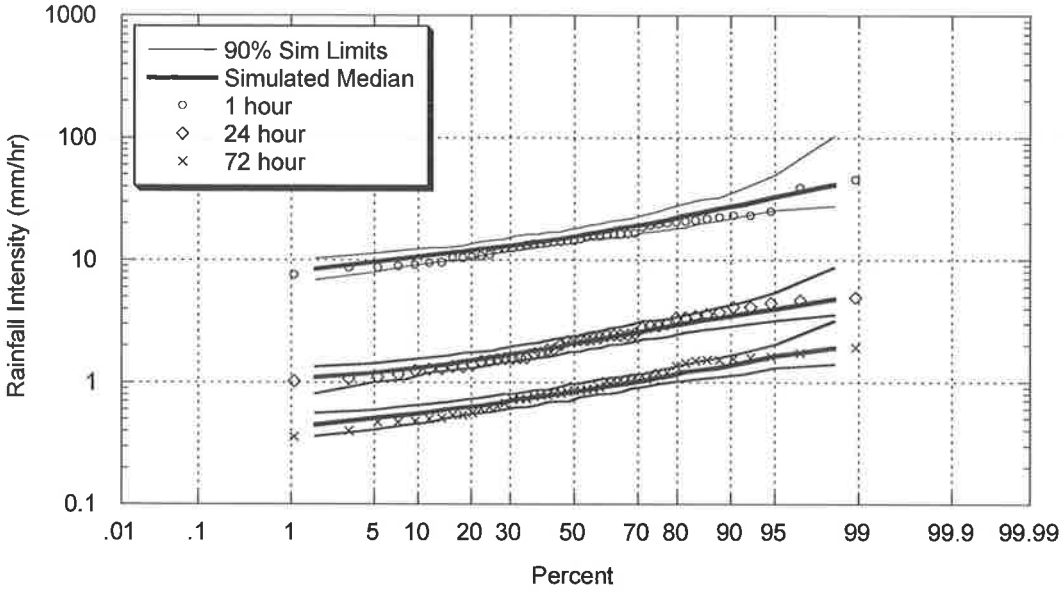

Figure 6.49: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Melbourne; Target – East Sale)

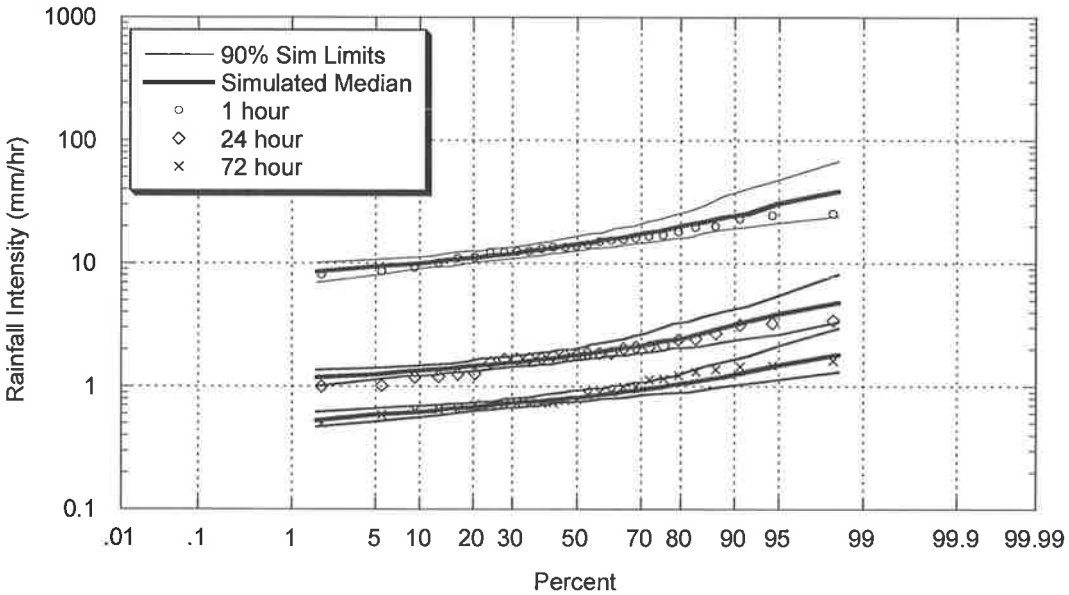

Figure 6.50: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Adelaide; Target – Williamstown)

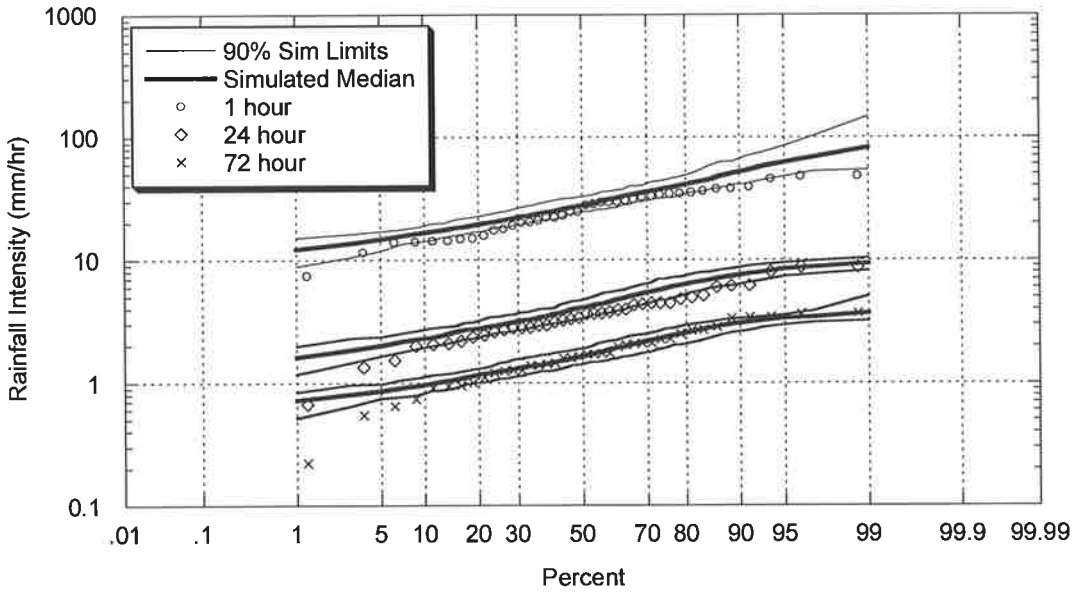

Figure 6.51: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Sydney; Target – Richmond)

6.5 Summary

The introduction of the master - target linear scaling relationship into the Heneker *et al.* (2001) rainfall model has produced a model capable of simulating long-term synthetic pluviograph records at numerous sites around Australia with short-term historical data.

The regionalisation model for inter-event times and storm durations employed a master – target relationship coupled with a linear updating factor between sites. An intermediate calibration was developed to remove any sampling variability issues between the rainfall records at the master and target sites due to their differing lengths and/or periods of record. The use of this intermediate step and the new regionalisation scaling factor enabled a successful translation from simulated inter-event time and storm durations at a master calibration site to a target site with only a short pluviograph record available for updating. Verification against calibrated and non-calibrated statistics has verified the adopted approach.

A similar master-target relationship was used as the basis in the regionalisation model development for storm event intensity. As a result of the requirement to consider the conditional intensity-duration relationship during regionalisation, the event depth distributions were used for comparison between sites rather than the storm intensities directly. The introduction of a non-parametric kernel smoothing technique to estimate the probability distributions for the event depths removed the need to assume a set distribution for this previously un-modeled storm parameter. An investigation into the temporal pattern parameters used in the model indicated they remained consistent within climate regions allowing the master parameters to be used without adjustment at the target site. The model assumptions and performance was verified with comparison between observed and simulated calibration statistics and independent statistics from a daily record at the target sites.

Comparisons between the observed and simulated data at various pairs of sites indicate the model's ability to significantly shift the mean and standard deviation of the bulk storm characteristics in order to reproduce various calibrated and non-calibrated statistics at sites with minimal Pluviograph data sets. These results give validity to the underlying structure of the model. Coupled with the ability to successfully simulate rainfall amounts at various time scales and apply the existing disaggregation process to generate synthetic pluviograph data, this model is now a useful tool capable of being applied to a large number of additional sites across Australia. To enhance the application of the model further, a process was developed to enable model calibration at sites with only daily data available. This work is presented in the following chapter.

CHAPTER 7

REGIONALISATION WITH A DAILY RAINFALL RECORD

7.1 Introduction

High resolution point rainfall models that can be accurately calibrated to daily rainfall data would provide a useful tool for investigating engineering systems. Such models would need to be capable of replicating both daily and sub-daily statistics as well as longer aggregated values to be confident of the structure and robustness of the model. In addition the calibration process should be straightforward and capable of being employed by all model users.

Previously, Chapter 6 demonstrated the ability of the rainfall model to be calibrated at sites with a short pluviograph record. This chapter further develops the model to enable calibration at sites with only daily rainfall data. This work meets the final objective of this study which was "To extend the application of the model to sites with only historical daily data and no pluviograph data available for calibration".

As discussed in Chapter 1, there are more than 1400 sites across Australia that now have active daily rainfall records which contain at least 40 years of historical data (see Figure 7.1).

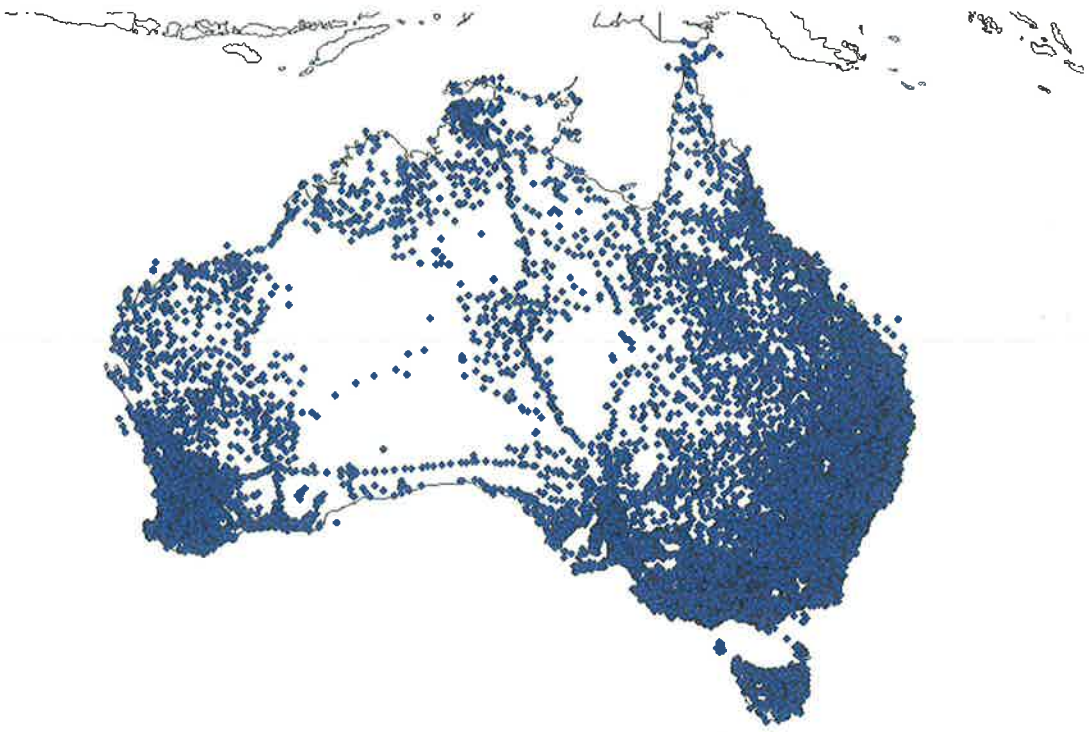

Figure 7.1: Australian Bureau of Meteorology: Daily Recording Stations

This data is readily available and provides a useful resource for researchers and practising engineers alike. To use this information a Master – Target relationship was developed similar in structure to that introduced in Chapter 6. Again an initial calibration of model parameters is undertaken at a long pluviograph master site. In this case, the parameters are then updated through comparison to daily data at the Target site. This new simulated likelihood approach developed in Section 7.2.1, in conjunction with non-parametric density estimation allowed model parameters to be updated directly based on the comparison between simulated daily Master rainfalls and observed daily Target data.

Model validation occurred at numerous pairs of sites which were selected to ensure the target site also contained a significant pluviograph record. Selecting sites with a significant pluviograph record provided adequate information to test the accuracy of reproducing various statistics at the sub daily time scale as well as aggregated statistics and monthly/annual rainfall distributions. The final model is capable of calibration to sites with daily rainfall data (available either through direct data measurement or via the SILO Bureau of Meteorology process for interpolation of data between measurement sites), greatly enhancing the application of the model across Australia

and providing an excellent continuous simulation tool for engineers and designers in areas where no historical short time increment rainfall data is available.

7.2 Development of the Daily Regionalisation Model Structure

In order to develop a storm event model that could be calibrated with daily data, the model parameters and their relationship to the available data required consideration. At locations where only daily data is available at the target site, it is clear that there is a limit to the quantity and quality of detailed information which could be used to calibrate/adjust model parameters. The Master – Target structure presented in Chapter 6 relied on the fact that there was pluviograph calibration data available (albeit limited) which could be used to adjust the master parameters providing a direct update for application at the Target site. While a similar master – target approach was seen to be advantageous in terms of structure and understanding, it required further development for application with daily data.

The statistics that are available within a daily data record include the probability of observing rain on any given day, the distribution of rain given the day is a rain day and the probability a day receives zero rain. These values (Probability of observing dry days, the probability of observing rain days and the distribution of rain totals given a rain day) are all statistics that can be easily extracted from a daily record and while they contain no direct information regarding storm event parameters, their values are influenced by the distributions of inter-event time, storm duration and storm depth. It is this relationship between the storm event parameters and the daily data statistics which led to the adoption of a similar master – target calibration structure as that developed in Chapter 6.

If we consider the distribution of inter-event times, this distribution provides a major influence on the frequency of storm events observed during a given period of time. (i.e. the longer the average inter-event time, the fewer number of storm events in a day/month/year etc). So even though there is no *direct* information regarding inter-

event times in a daily record, daily statistics such as the daily dry probability (or the probability of no rain) or potentially the distribution of consecutive dry days can be used in the calibration process in some form to infer the required parameter changes between master and target sites.

In a similar manner, the model parameters which describe the values of storm depth cannot easily be adjusted directly based on daily data as there is limited detailed information within a daily record regarding the depth of individual storm events. Consider the situation when two storms events occur on a single recording day. As the daily data records the observed depth once a day, there is no additional information available in the daily record to estimate the rain contributions each individual storm event has provided to the overall daily total. Similarly, it is impossible to define and allocate daily rainfall contributions from two storms where one has continued across the arbitrary recording day boundary and another storm starts subsequently on that same day. Only in the rare occurrence when an individual storm event is completely contained within a recording day and it is the only storm for that day can the recorded daily depth be attributed to an individual storm depth. However, the ability of the rainfall model to successfully simulate the mean daily rainfall (when calibrated to pluviograph data) suggests such a statistic could be used to enable comparison between daily information at master and target sites.

To use the available daily data, a method was required to provide an accurate comparison between simulated and observed daily statistics. This would allow the model parameters to be adjusted and updated based on comparisons of these daily data statistics between master and target sites. So, while the regional model developed in Chapter 6 could not be adopted directly, the Master – Target structure was again introduced for use with daily data. This structure had the benefits of having already been demonstrated as a success with pluviograph data and also provided a technique to use an initial accurate model calibration at the master site as the basis for parameter values at the target site.

The introduction of a Master – Target relationship into the model for consideration with daily data is presented schematically in Figure 7.2. Similar in structure to the model in Chapter 6, the master site provides an initial accurate calibration to a long

pluviograph record. In this case however, target daily data is then used to adjust the model parameters producing a simulation model at the Target site. (Daily data is also required at the master site to provide an initial fit which removes issues associated with different record periods. This is discussed in further detail in section 7.2.2)

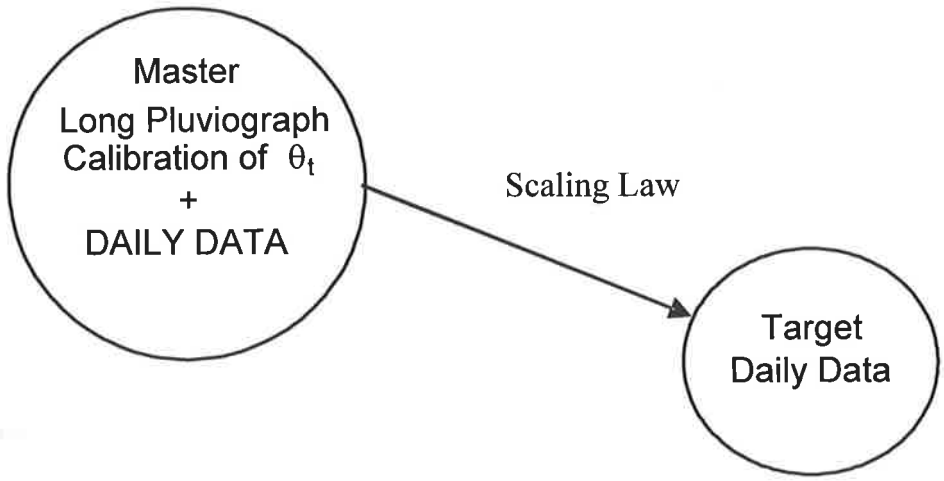

Figure 7.2: Schematic of Regionalisation Structure

One of the requirements of this model structure is the ability to successfully simulate rainfall at timescales other than those used during the calibration process (in particular at the daily scale). With each iterative change in model parameters, a new simulation is required to generate a simulated master daily record which can then be compared to the observed target daily data. This comparison is used to drive changes in the scaling model parameters and improve the fit of the model at the target site. It is important that the model is capable of replicating the daily rainfall data in the first instance so that any identified parameter changes are a result of site requirements and are not reflective of simulation errors. In this case, not only should the model be capable of replicating the daily statistics at the master site after a direct calibration, but it also should be replicating the daily statistics at the target site after a successful regionalisation.

As previously shown in Chapter 6, Figure 7.3 presents a comparison between observed and simulated daily dry probability results when regionalising from a master at Sydney to the target at Richmond using pluviograph data. The fact that the

regionalised model has been able to successfully capture this statistic when calibrated with pluviograph data suggests that this statistic can also be used as part of the calibration when developing the daily regionalisation process.

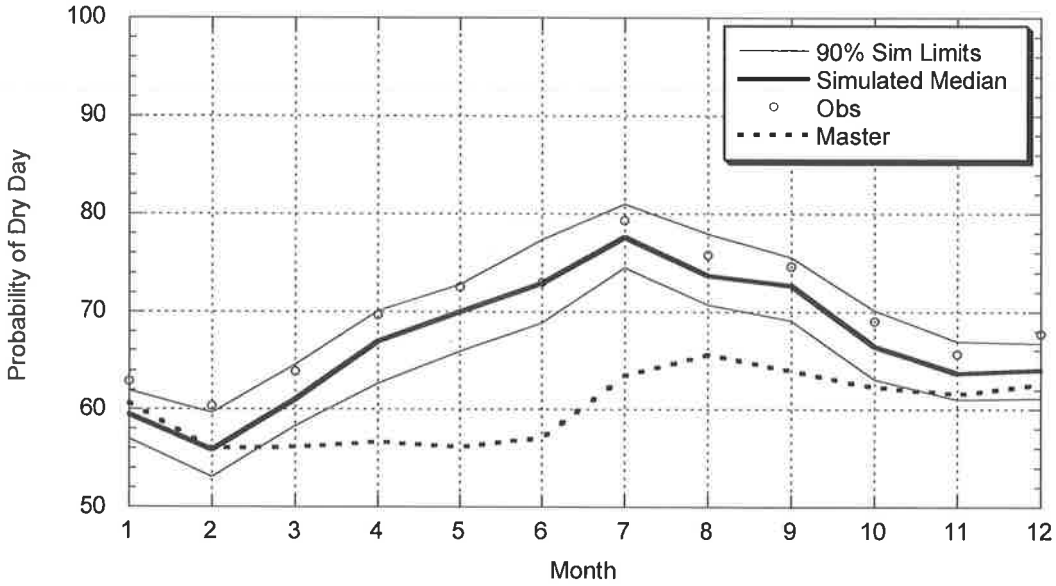

Figure 7.3: Comparison between Observed and Target Simulated Probability of a Dry Day (Master – Sydney; Target – Richmond)

In a similar result, Figure 7.4 displays the ability of the rainfall model to capture the non-calibrated mean daily depths. (It also indicates the ability of the model to reproduce this statistic when shifting from master to target with pluviograph data). Again the model has reproduced this daily statistic and coupled with the previously displayed ability of the model to replicate aggregated monthly and annual rainfall statistics, provides confidence that aggregated statistics can be used to calibrate the storm event parameters in a daily regionalisation model.

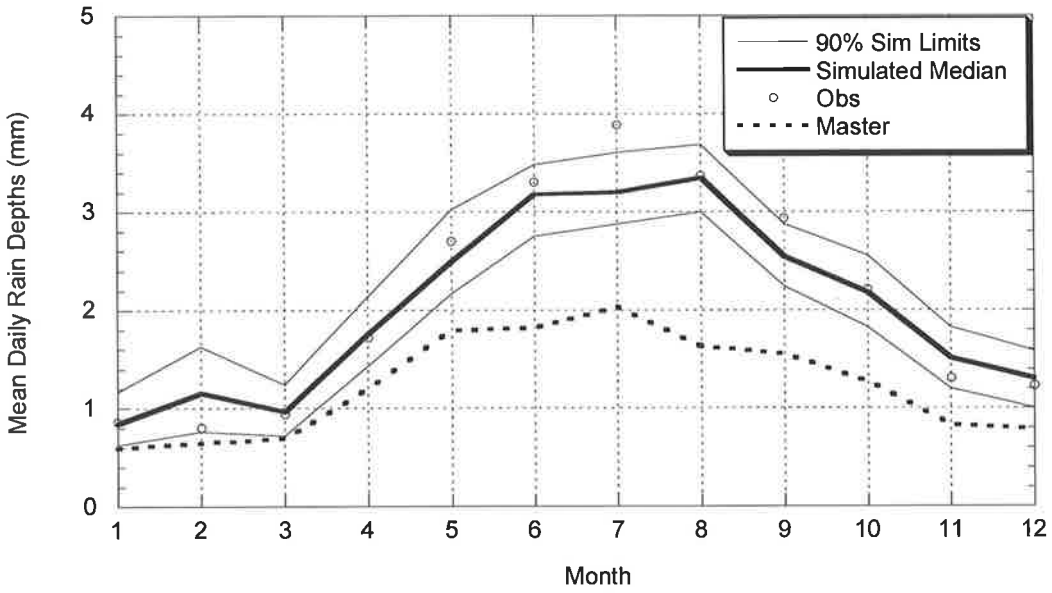

Figure 7.4: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Adelaide Airport; Target – Williamstown)

In order to calibrate the storm event model parameters to daily statistics a new likelihood approach was developed which provides a direct comparison between simulated model output and the observed daily data at the target site. This new likelihood process exploits the ability of the model to successfully simulate these rainfall statistics at the daily timescale and provides a robust comparison between master parameters and observed daily data statistics.

7.2.1 Daily Calibration Model Development and Simulated Likelihood Approach

To develop a master – target likelihood approach to regionalise the rainfall model using daily data, it is appropriate to review the successful regionalisation model that was developed for use with a short pluviograph record. It is this master – target pluviograph regionalisation model that formed the basis for the daily work.

In Chapter 6, a scaling factor λ was introduced to the maximum likelihood equations to provide new likelihood equations for the regionalised storm event calibrations

when dealing with a short pluviograph record. Equations (7.1) and (7.2) display the maximum likelihood equations for inter-event times/storm duration and the non-parametric kernel smoothing approach that was used as the basis for storm depths respectively.

$$F_Y(y \mid \theta_t, \lambda) = F_X(\lambda y \mid \theta_t) = 1 - \exp[-g(\lambda y, \theta_t)] \qquad x > 0 \quad y > 0 \quad \lambda > 0 \qquad (7.1)$$

F_Y denotes the distribution function at the target site while F_X denotes the distribution function at the master site. The scaling factor λ acts as a multiplier on the target pluviograph data y, to provide a best fit between the master calibration and the target data.

$$\hat{f}_X(x) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x - (\lambda y)_i}{h}\right) \qquad (7.2)$$

K() is a kernel function centred at each data point, n denotes the number of data points at the target site while $(\lambda y)_i, i=1,..,n$ is the target depth data. These values of $(\lambda y)_i$ are a product of storm duration t_d, the corresponding duration scaling λ_d, the corresponding target event intensity and the required depth scaling λ.

When using pluviograph data, the scaling value λ acted as a direct multiplier on the observed target pluviograph data (y) ensuring a good fit between the target pluviograph data set (y) and the calibrated parameters at the master site. Maximum likelihood techniques were used to calculate the required scaling parameters.

The development and results of the pluviograph regionalisation work in Chapter 6 provided the impetus for incorporating a similar master – target framework for use with Daily Data. One of the outcomes of the work in Chapter 6 was that of the temporal pattern parameters and the fact that they were shown to be consistent enough between sites within a similar climatic region and location to be kept the same without significant deterioration to the accuracy of the model results. This result also allows these temporal pattern parameters to be kept unchanged when shifting from master to target site with Daily Data. Without such a result, an alternative method of estimating the change in temporal pattern parameters would be

required and would be a major issue as there is no storm information within a daily data record which could adjust the temporal pattern parameters in the model. However, given the consistent temporal pattern parameters and the ability of the model to capture non-calibrated aggregated statistics at the daily level, introducing a master – target approach provides a workable solution for calibrating the model at sites with only daily data.

The daily regionalisation model also introduces a set of scaling parameters (scaling parameters are required for each of inter-event times, storm durations and storm depths and are again denoted λ) as a means of shifting from the master to the target site. However, unlike the pluviograph regionalisation model (and the original model) which used traditional maximum likelihood methods to ensure the best fit between estimated model parameters and observed data, the likelihood description for daily data is not as straightforward. The likelihood model cannot be formulated directly as there is no storm event data available at the target site for comparison. In order to use available information from a daily data record, an alternate likelihood approach was required. This approach has been termed the simulated likelihood approach and provides a flexible model structure capable of dealing with data at significantly different time scales to that usually used for parameter calibration.

Simulated likelihood relies on the ability of the model to produce a simulation for each and every adjustment in regionalisation model parameters during the calibration process. For each iterative step the regionalisation parameters are adjusted within the SCE search routine and then these "trial" parameters are used to generate a new realisation of the model providing the simulated pluviograph data at the target site. This simulated data is then aggregated into a daily time step and compared to the available observed daily data at the target site. If a better fit is required, the parameters are adjusted and the model re-simulated and so forth. In this way the model is capable of being compared and adjusted directly based on its fit to the daily data record.

The likelihood model for daily regionalisation is set up at the daily time step and is significantly more complex than that required for pluviograph data calibration. The first requirement of the model is the capability to compare the distributions of daily

rainfall totals. For each iteration in the calibration process the scaling parameters (λ) are changed and a new simulation is required to provide simulated target data which can be aggregated into a new record of simulated target daily depth data (denoted y). The simulated target data is then compared to the observed target daily data and the maximum likelihood approach drives changes to the scaling parameters until the best fit is found. In order to incorporate the maximum likelihood approach, a description of the probability density for these daily rainfall depth distributions was required. As was the case in Chapter 6 for the distribution of event depths, the adoption of the kernel smoothing approach provides this density estimate and takes the form

$$\hat{f}_X(x) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x - (\lambda y)_i}{h}\right) \tag{7.3}$$

where K() is again a kernel function centred at each scaled simulated target daily data point λy, x is the observed daily target data value at which the probability density estimate is required for use in the maximum likelihood calculations and h is the bandwidth. As before, a Gaussian kernel was used for K().

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \tag{7.4}$$

Using the Kernel Smoothing approach in this way enables a comparison between the master and target distributions of daily rainfall depth totals given the day receives rain. If all that was required was the best fit between the master and target distributions of daily depths, then the maximum likelihood estimators for λ would be obtained when the product of $\hat{f}_X(x)$ (or the sum of $\log \hat{f}_X(x)$) for all master observations was at a maximum with no other considerations. However, the model must also take into account any changes to the probability of observing a dry day. To do this, the dry days are separated from the aggregated daily record and the probability of observing a dry day $P(Y=0 \mid \lambda, \theta_t)$ calculated. Considering this in the likelihood formulation gives the following:

$$\hat{f}_X(x) = P(dry) = P(Y = 0 \mid \lambda, \theta_t) \quad \text{for } x = 0 \tag{7.5}$$

and

$$\hat{f}_x(x) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x-\lambda y_i}{h}\right)(1-P(dry)) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x-\lambda y_i}{h}\right)(P(wet)) \quad \text{for x>0} \qquad (7.6)$$

The scaling parameter set λ is adjusted for each iteration during calibration until $\hat{f}_X(x)$ for all master observations is at a maximum. This ensures the best fit between both the simulated and observed dry probability and the distribution of daily depth totals. The format and structure of the scaling parameters in the daily model are identical to their pluviograph counterparts (i.e. for inter-event times and storm durations a single linear scaling parameter is used while a 3 parameter model conditional on storm depth is used as the basis of storm depth scaling (refer to section 6.4.1)).

The end result was a set of model parameters at the target site which were a combination of the original master parameter set, and any scaling due to the variations between the two sites. However, as was the case for the work with pluviograph regionalisation, variations between the sites were not only a result of climatic differences but also due to differences in data length and non concurrency. This issue required further attention prior to developing the final daily regionalisation model.

7.2.2 Treatment of Sampling Variability between Rainfall Record Time Periods

Sampling variability between rainfall time periods was first identified as an issue during the work on the regionalisation model for use with short pluviograph records. Instances where rainfall records were not concurrent resulted in inaccurate comparisons between the two rainfall sites because one site generally observed a significantly different rainfall period than the other. (As discussed previously, this

may relate to extended wet or dry/drought periods present in one record and not in the other). As a result any statistical comparison was problematic and influenced by not only the differences between the sites (which is what the model was trying to describe), but also the different data periods across each record. To circumvent the issue, the pluviograph regionalisation model used an intermediate "pre-scaling" step to ensure the model was able to compare concurrent data sets between two sites.

The issue of sampling variability and non-concurrent data periods remains a problem when regionalising with daily data. In this case, the pluviograph record at the master site (used for the initial calibration) is usually different in length to that of the target daily record. Introducing an intermediate calibration into the daily regionalisation process again ensures the model compares differences between sites because of local variations only and not the residuals from non-concurrent data periods.

The intermediate calibration step introduced into the daily model is similar in structure to that developed in the previous chapter. The pluviograph model calibrated an intermediate parameter λ_p, which, when used as a pre-multiplication factor on the target pluviograph data, ensured the final regionalisation model was able to determine the 'true' relationship between the master and target sites. Using the intermediate parameter λ_p as the basis for daily data and introducing this into the likelihood equations gives

$$\hat{f}_x(x) = P(dry) = P(Y = 0 \mid \lambda_p, \lambda, \theta_t) \quad \text{for x = 0} \tag{7.7}$$

and

$$\hat{f}_x(x) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x - \lambda_p \lambda y_i}{h}\right)(1 - P(dry)) = \sum_{i=1}^{n} \frac{1}{nh} K\left(\frac{x - \lambda_p \lambda y_i}{h}\right)(P(wet))$$

$$\text{for x>} \tag{7.8}$$

where λ_p is the set (inter-event time, storm duration and storm depth) of intermediate scaling factors.

This new daily intermediate step compares a subset of the observed daily record at the master site (equivalent in length to the target daily record) to the simulated master daily record. By extracting a subset of observed daily data at the master site equivalent in length to the observed target daily record and calculating the intermediate scaling factor, subsequent simulations from the model are pre-adjusted allowing direct comparison between model simulations and the target daily record. The intermediate parameter has described any discrepancies that result from the different time-periods observed between the original pluviograph calibration data and the daily data at the target site.

Once the intermediate factor has been determined, it is applied as a pre-multiplication to the target daily data set. The regionalisation procedure is then applied a second time to compare the master simulated daily record and the 'new' adjusted target data set to determine the 'real' regionalisation scaling factor λ. As the intermediate step takes into account the differences in data periods and therefore any potential changes in the daily rainfall distributions, the influence on the resultant 'real' scaling factor is purely any differences between the master and target sites. Once again the scaling parameters are calculated on a monthly basis therefore requiring the determination of an inter-event, storm duration and storm depth scaling factor for each month.

7.3 Model Calibration Using Daily Data Results

7.3.1 Introduction

Validation of the daily regionalisation model required careful selection of target sites. Not only did it need to be shown that the model could be calibrated to daily data, the model must also be able to reproduce the required sub-daily statistics at the target site. As a result, identical test sites to those used previously to validate the pluviograph regionalisation work in Chapter 6 have been chosen to validate these models (with the exception of Williamstown, South Australia which did not have an

associated long daily record and has been replaced by Rosedale in South Australia). These sites contain both a long daily and a significant pluviograph record providing the required observed statistics and master-target scaling parameters can be compared directly between the pluviograph and daily models. The table of test sites is provided below.

Table 7.1: Master and Target Rainfall Record Details (Daily Model)

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	Melbourne	86071	1900	1995	0
Targets	East Sale	85072	1953	1992	190
	Ellinbank	85240	1961	1992	95
	Laverton	87031	1965	1992	20

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	**Sydney**	**66062**	**1913**	**1991**	**0**
Targets	**Richmond**	**67033**	**1953**	**1993**	**45**
	Chichester	**61151**	**1960**	**1980**	**185**

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	**Adelaide**	**23034**	**1967**	**1997**	**0**
Targets	**Williamstown**	**23763**	**1971**	**1997**	**40**
	Stirling	**23785**	**1964**	**1981**	**15**

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	**Perth**	**9034**	**1946**	**1992**	**0**
Targets	**Esperance**	**9631**	**1963**	**1991**	**580**

	Name	BOM #	Start Year	Finish Year	Distance (km)
Master	**Brisbane RO**	**40214**	**1908**	**1991**	**0**
Targets	**Brisbane AMO**	**40223**	**1949**	**1992**	**10**
	Kirkleagh	**40318**	**1959**	**1990**	**70**

A number of statistical indicators and distributions were chosen to verify the performance of the model. Calibrated statistics such as the daily dry probability and mean daily rainfall are checked to test whether the regionalisation calibration was completed successfully. Comparisons between observed and simulated bulk storm event distributions (inter-event times, storm duration and storm depths) provide an indication that the model was able to successfully simulate storm events and reproduce sub-daily non-calibrated statistics. Further verification that the model has captured the bulk rainfall processes is achieved by comparing annual and monthly rainfall distributions, while IFD curves are compared between the simulated values and the target site, testing the assumed similarity between disaggregation parameters at the two sites of interest. The calculation of observed statistics for annual and monthly rainfall comparisons used the daily rainfall records at the target site as they provide the most accurate monthly and annual statistics. (This is because most pluviograph records have sections of missing data, corrupting the monthly and annual rainfall totals). All other statistics were calculated from the observed target pluviograph. The results from 3 pairs of test sites (Sydney – Richmond, Brisbane – Kirkleigh and Adelaide – Williamstown) have been presented within this Chapter to demonstrate the performance of the model with all other site results presented in Appendix D.

Successful reproduction of statistics and distributions across the different time scales provided evidence that the model was able to be calibrated with daily data and that the assumptions and structure of the regionalisation process was sufficient to capture the required local variations between sites ensuring an accurate synthetic pluviograph record.

7.3.2 Calibrated Daily Statistics

The original rainfall model and the pluviograph-regionalised model were both capable of successfully replicating the probability of observing a dry day and the

mean daily rainfall distribution even when these models did not use this information during the calibration process. Successful calibration of the rainfall model with target daily data should again ensure the reproduction of these daily statistics at all sites. Previously (for the original and pluviograph-regionalised model) this successful reproduction provided confidence in the structure and assumptions within the model, in this case it is a good indication of the success or otherwise of the model calibration.

Figure 7.5 compares simulated and observed daily dry probabilities for data from Richmond (with Sydney as the master and shown for reference). It is evident from this result that the completed calibration was successful and that the model is capable of reproducing this daily statistic even with a significant shift when comparing the master and target statistics. Figure 7.6 reinforces this result with similar results for Kirkleigh (Brisbane master).

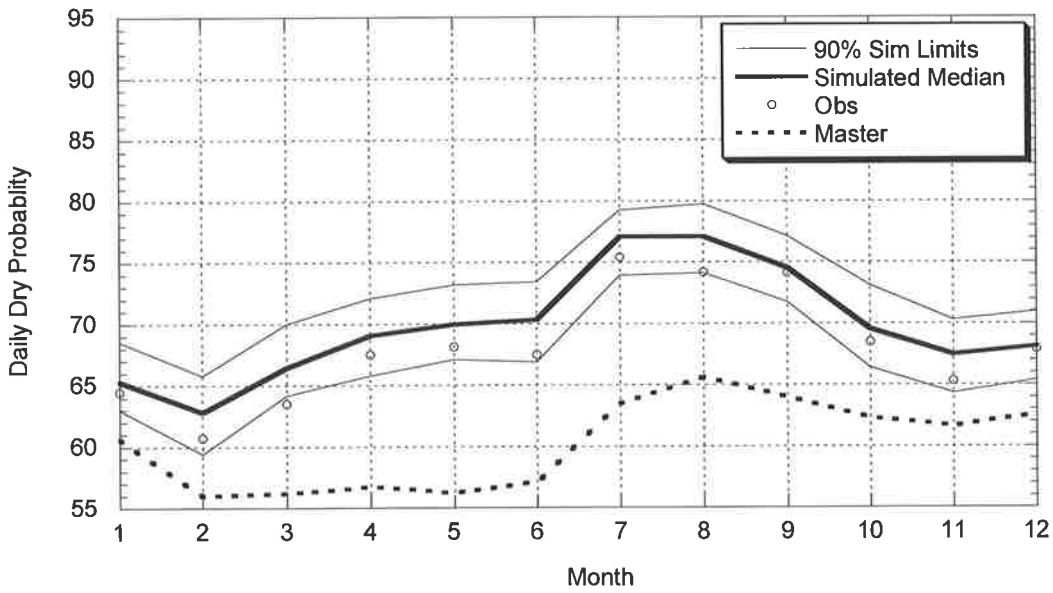

Figure 7.5: Comparison between Observed and Target Simulated Daily Dry Probabilities. (Master – Sydney; Target – Richmond (Daily))

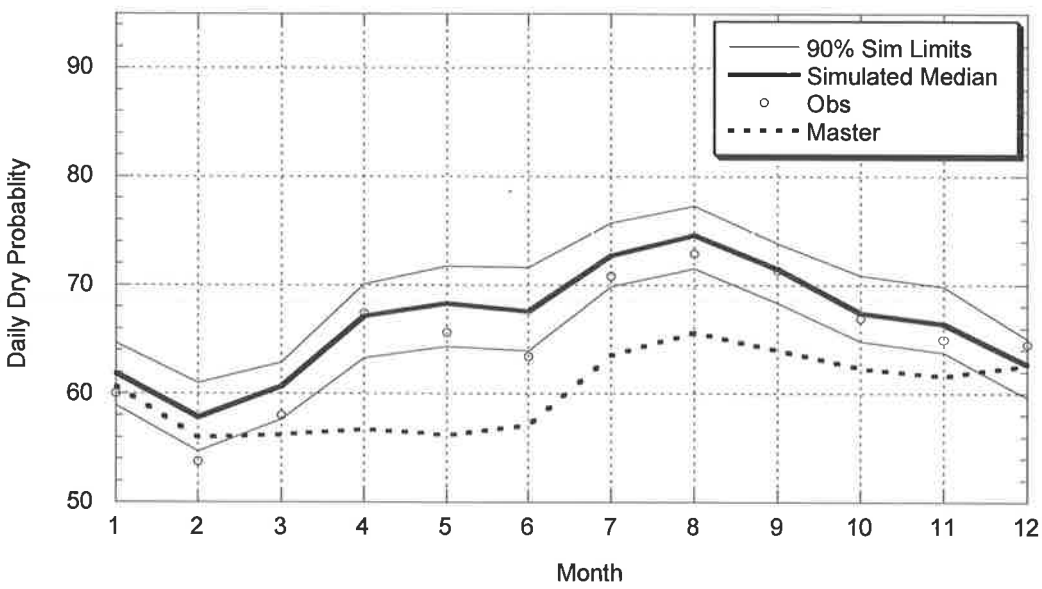

Figure 7.6: Comparison between Observed and Target Simulated Daily Dry Probabilities. (Master – Brisbane; Target – Kirkleagh (Daily))

Figure 7.7 presents data from Williamstown (Adelaide Master).

In contrast to the Richmond and Kirkleigh sites, there was very little difference between the daily dry probabilities at Adelaide (master) and Williamstown (target). As can be seen in Figure 7.7, the observed data still lies within or very close to the simulation limits. These results when coupled with the comparison between daily and pluviograph scaling parameters for inter event times, indicate that the model has been successful in calibrating to daily data and capturing the required variations between master and target sites.

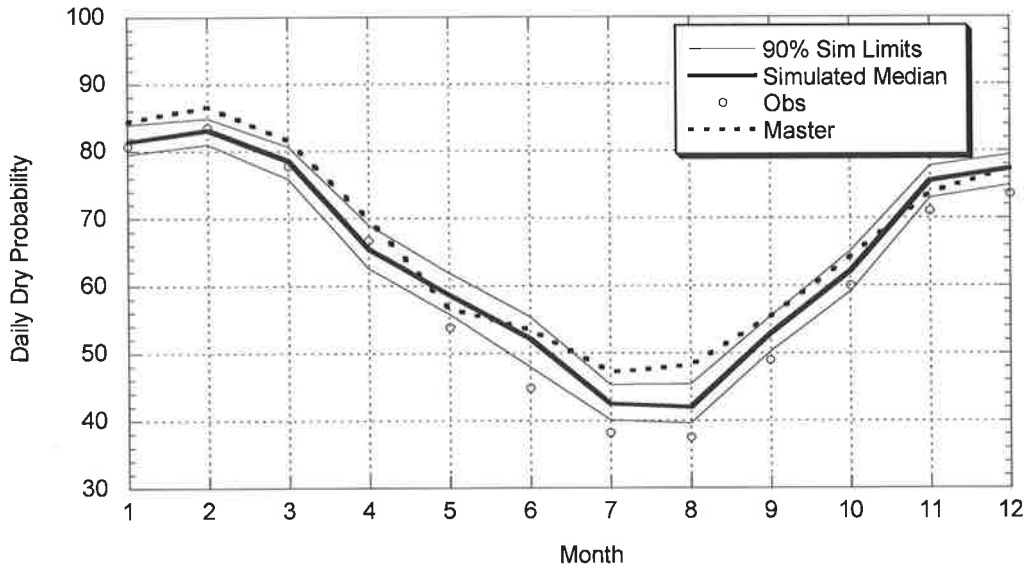

Figure 7.7: Comparison between Observed and Target Simulated Daily Dry Probabilities.
(Master – Adelaide; Target – Williamstown (Daily))

The second statistic used during the calibration process was the mean daily rainfall. Both the storm duration and storm depth parameters have a significant influence on this daily statistic. Figure 7.8 presents this result for Williamstown (with Adelaide master as reference). Again, the model has been able to capture the differences between the two sites and reproduce this calibrated statistic. Most months required significant adjustment and only April sits slightly outside the simulation bounds.

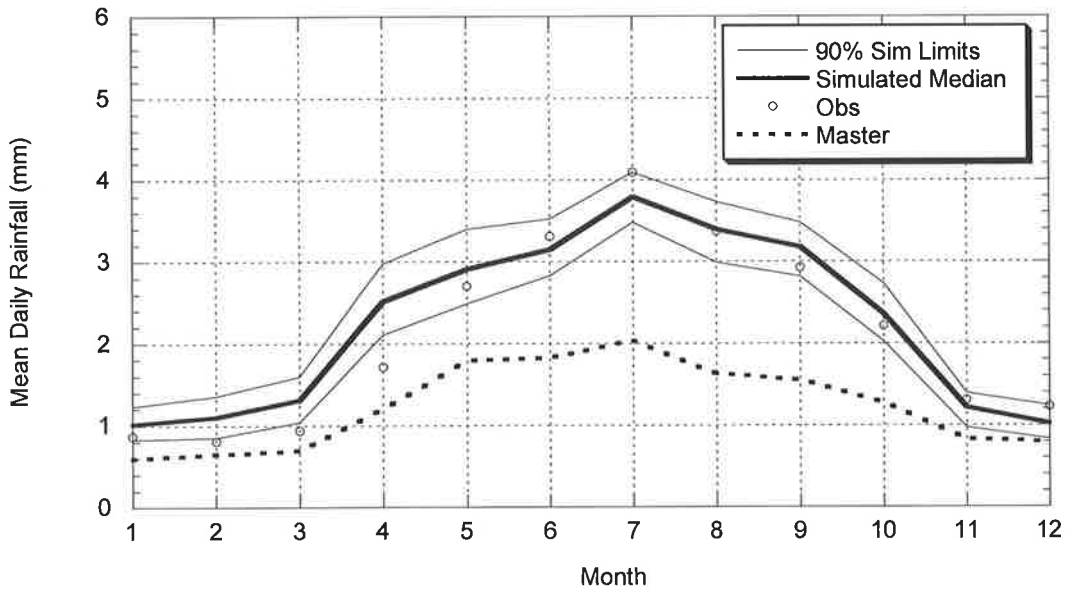

Figure 7.8: Comparison between Observed and Target Simulated Daily Mean Depth.
(Master – Adelaide; Target – Williamstown (Daily))

Figure 7.9 & Figure 7.10 present additional results for Richmond and Kirkleigh with all results falling within the simulation bounds. These results reinforce the model's ability to calibrate to the mean daily depth distribution and successfully reproduce this statistic during simulation.

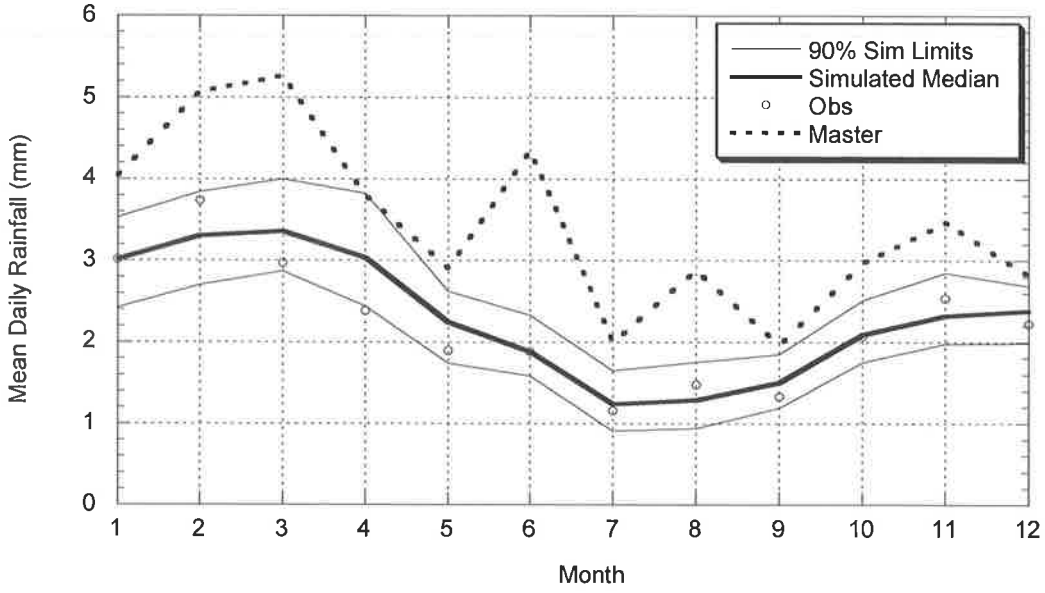

Figure 7.9: Comparison between Observed and Target Simulated Daily Mean Depth.
(Master – Sydney; Target – Richmond (Daily))

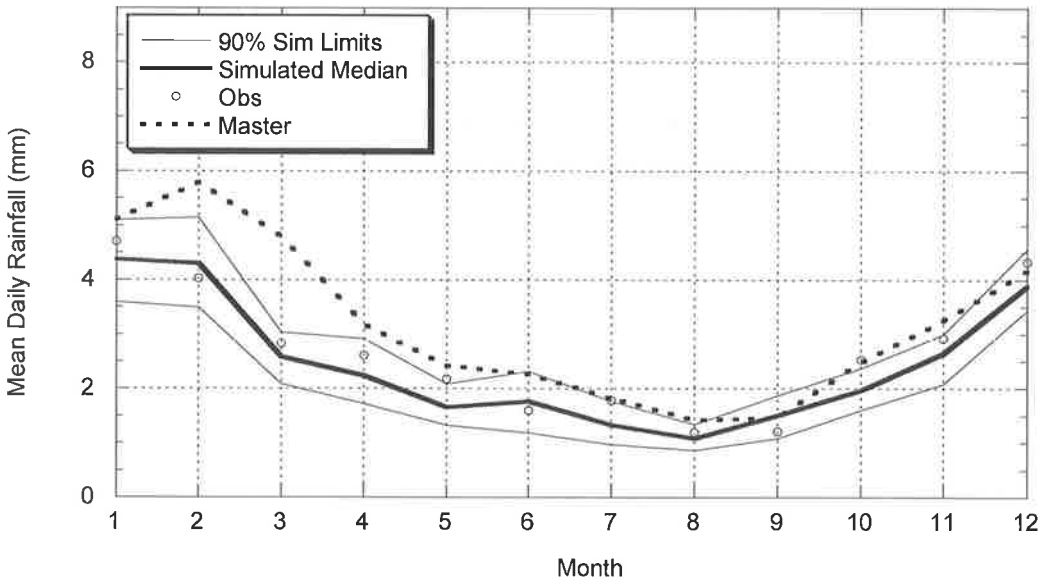

Figure 7.10: Comparison between Observed and Target Simulated Daily Mean Depth.
(Master – Brisbane; Target – Kirkleagh (Daily))

Appendix D provides additional plots across numerous sites with similar results. These results confirm the ability of the model to successfully use daily data for regional calibration (and reproduce these calibrated statistics during simulation). However, the ability of the model to replicate non-calibrated statistics at time scales greater than or less than the daily scale still requires validation. Prior to presenting these results however, it is also important to investigate the accuracy of the model in replicating the storm event distributions of inter-event time, storm duration and storm depth as these are the building blocks of the original model. Successful reproduction of these bulk storm distributions would provide evidence that not only was the calibration process a success, but that the model structure itself is sufficient to use daily data and describe the differences between two sites.

7.3.3 Comparison of Observed and Simulated Annual and Monthly Rainfall Distributions

It is important that any storm event rainfall model is capable of reproducing non-calibrated statistics such as the monthly and annual rainfall, not only as it is a representation of the characteristics of the rainfall site in question, but also specifically in this case as it provides further indication that fitting to the distribution of daily depths has worked successfully.

Comparison to annual rainfall is an effective test of the model structure. Calibration of the model with daily data has introduced a scale shift during calibration in the model from data at a 24 hour time scale down to effectively 6 minute data. If the aggregation back up to annual rainfall (and monthly rainfall) is accurate, then this is further evidence that the model structure and assumptions have merit. In any case, for the model to be accepted as a tool for engineering applications, aggregated statistics must be well reproduced at numerous time scales.

Annual rainfall plots are presented below for Richmond (Master – Sydney) Figure 7.11, Kirkleigh (Brisbane) Figure 7.12 and Williamstown (Adelaide) Figure 7.13. As mentioned previously plots for all other sites can be found in the Appendix D.

These plots demonstrate the ability of the model to successfully reproduce the annual rainfall and provide confidence that the model has been successfully applied to the target site with only daily data for calibration. There is a significant shift from the distributions at the master and target sites for both the Richmond and Williamstown results with Richmond in particular very successful at replicating the observed annual rainfall values. Williamstown annual rainfall has been slightly overestimated by the simulation however the majority of observed data points sit within or just outside the simulation limits. There is a smaller difference between annual rainfall at Brisbane and Kirkleigh, however the model has once again been successful in shifting to the target site as evidenced by the observed points sitting within or on the edge of the simulation bounds.

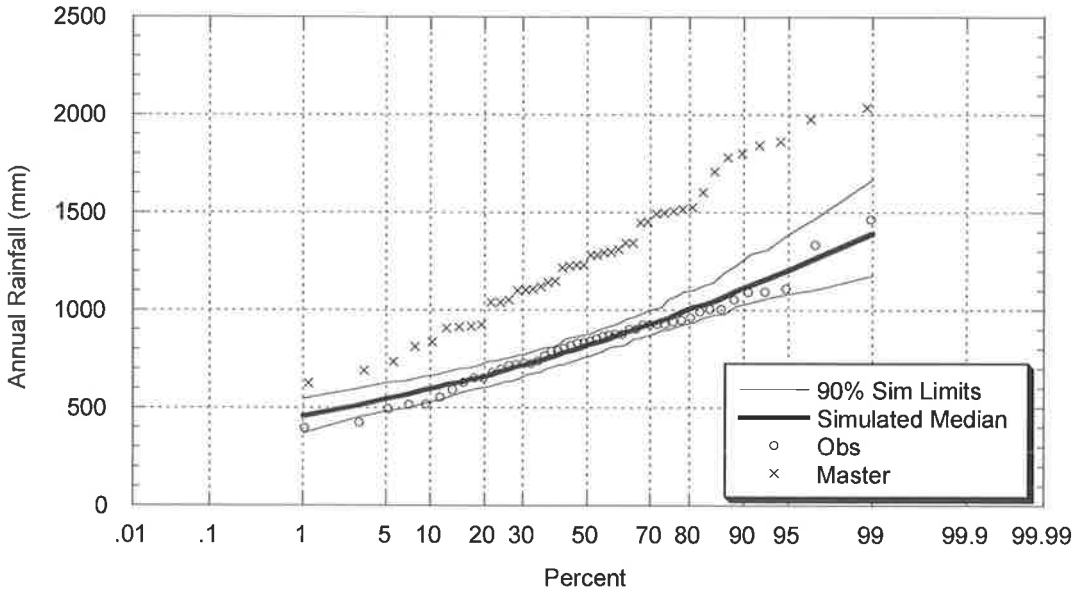

Figure 7.11: Comparison between Observed and Target Simulated Annual Rainfall. (Master – Sydney; Target – Richmond (Daily))

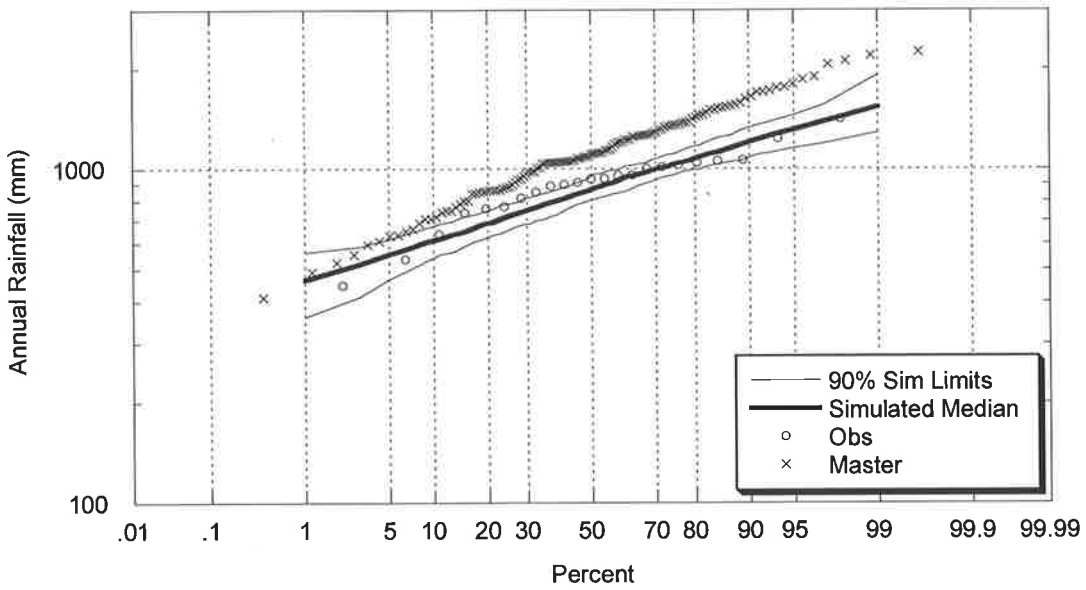

Figure 7.12: Comparison between Observed and Target Simulated Annual Rainfall.
(Master – Brisbane; Target – Kirkleagh (Daily))

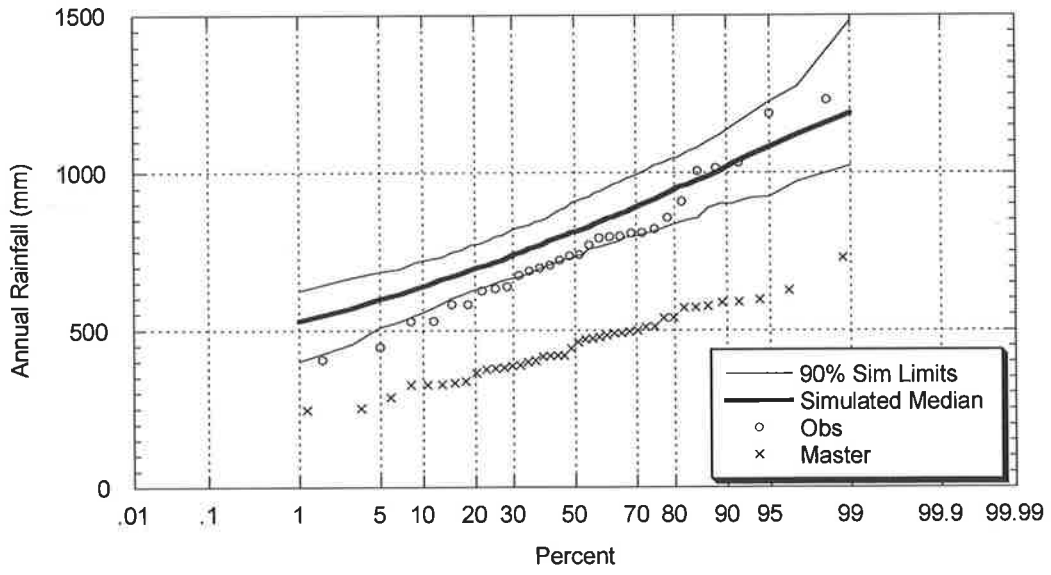

Figure 7.13: Comparison between Observed and Target Simulated Annual Rainfall.
(Master – Adelaide; Target – Williamstown (Daily))

The reproduction of monthly rainfall in April is presented in Figure 7.14 for data from Richmond (Sydney). All observed rainfall totals fall within the simulation limits suggesting the model has reproduced the monthly rainfall pattern.

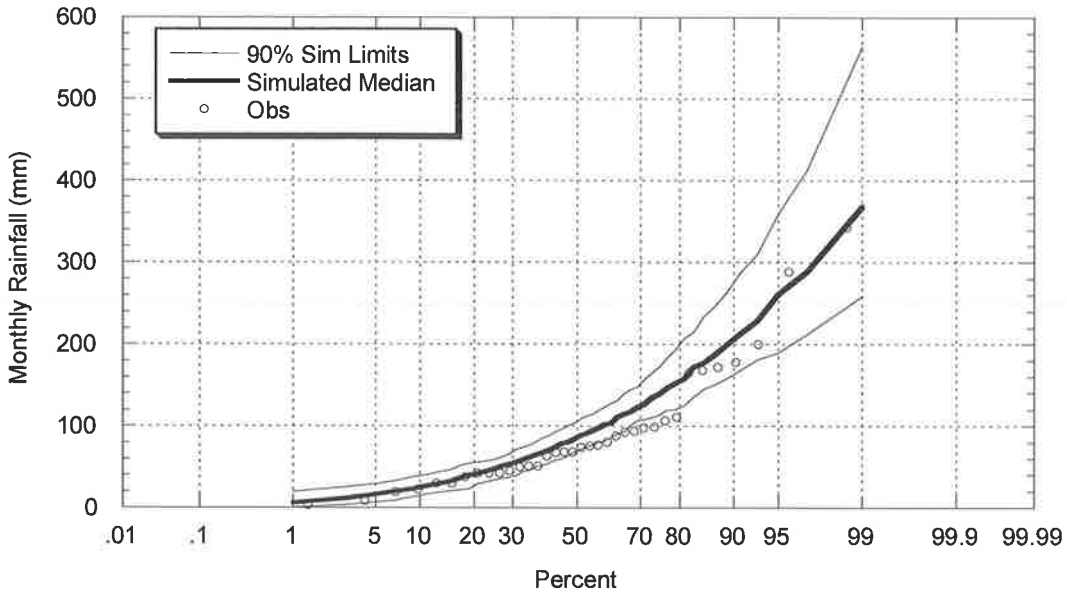

Figure 7.14: Comparison between Observed and Target Simulated April Rainfall. (Master – Sydney; Daily Target – Richmond (Daily))

Additional results for Williamstown (Adelaide) and Kirkleigh (Brisbane) are presented below. With the exception of a select number of points, all observed monthly rainfall totals fall within the simulation limits. Additional monthly rainfall results can be found in Appendix D.

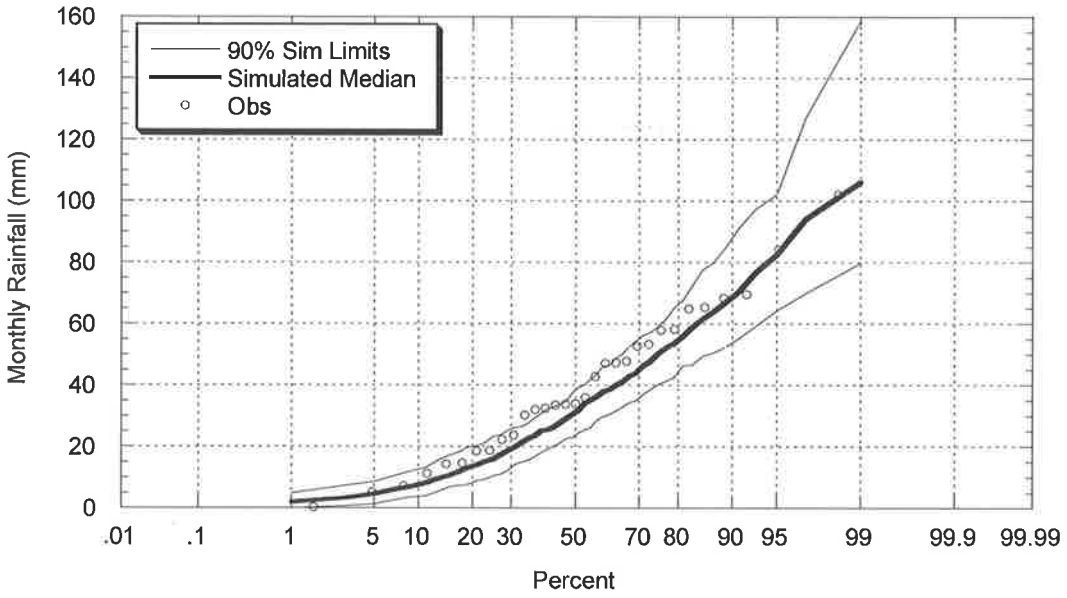

Figure 7.15: Comparison between Observed and Target Simulated Novemner Rainfall. (Master – Adelaide; Daily Target – Williamstown (Daily))

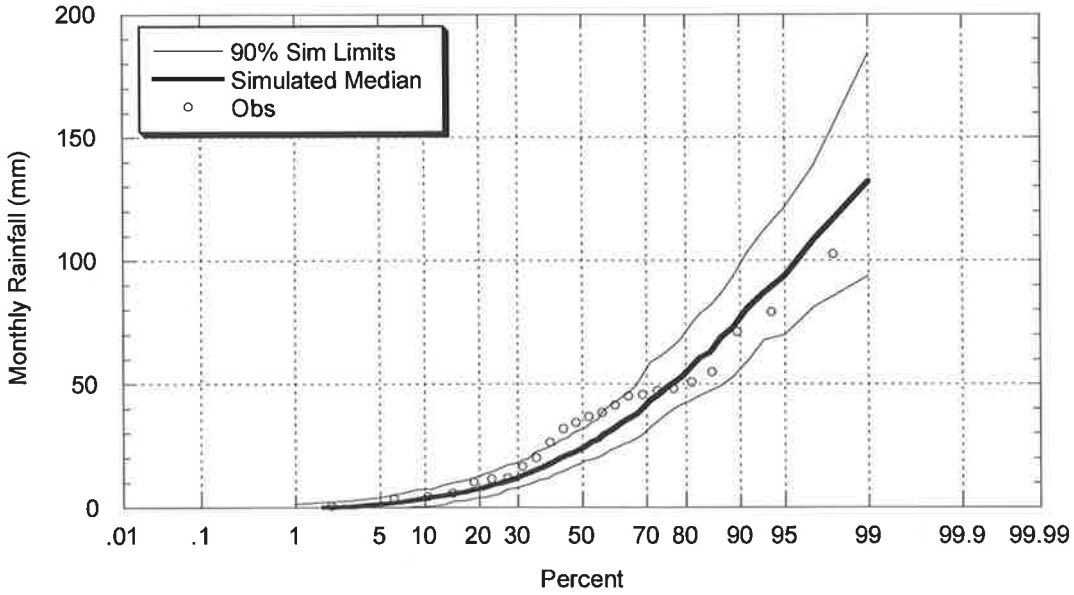

Figure 7.16: Comparison between Observed and Target Simulated August Rainfall.
(Master – Brisbane Regional Office; Daily Target – Kirkleigh (Daily))

The ability of the model to capture the aggregated monthly and annual rainfall distributions along with the aforementioned ability to reproduce the bulk storm characteristics suggest the rainfall amounts and dry periods are well represented in the regionalisation model. The final test is to compare the observed and simulated Intensity Frequency Duration results to ensure the model has been able to capture the temporal pattern within storm events.

7.3.4 Comparison of Observed and Simulated Bulk Storm Event Distributions - Inter-Event Times

The ability of the model to reproduce the bulk storm distributions when calibrated to a site with only daily data is a significant step forward for the application of this model and event based models in general as the theory behind this work can be adapted for use with other event models. Successful reproduction of these distributions would allow model application at sites that have only a daily rainfall record available for calibration. While it is unrealistic to expect the reproduction of bulk storm results to be as accurate as similar results obtained when the model is

calibrated directly with pluviograph data, results presented below indicate that calibration using the regionalisation process with daily rainfall data does provide a workable solution to model calibration in instances when pluviograph calibration data is not available. As with all stochastic models, it is reasonable to expect the user to know and understand the expected difficulties and deficiencies of the model when calibrated to daily data.

The selection of specific sites containing both a historical pluviograph and a historical daily record provides the best comparison between observed and simulated bulk storm distributions. The model was calibrated at the target site using the daily regionalisation model with resultant simulated outputs compared to the observed pluviograph data at the target site.

The comparison between simulated and observed inter-event times is presented in Figure 7.17 for data at Richmond shifted from a Sydney master calibration. As can be seen from this plot, the inter-event times are well reproduced by the daily regionalisation calibration model with the majority of storm events within the simulation limits. There is a slight deviation away from these simulation limits as the inter-event time decreases. This is a result of the lack of inter-event times available at a time scale less than 24 hours when using daily rainfall information.

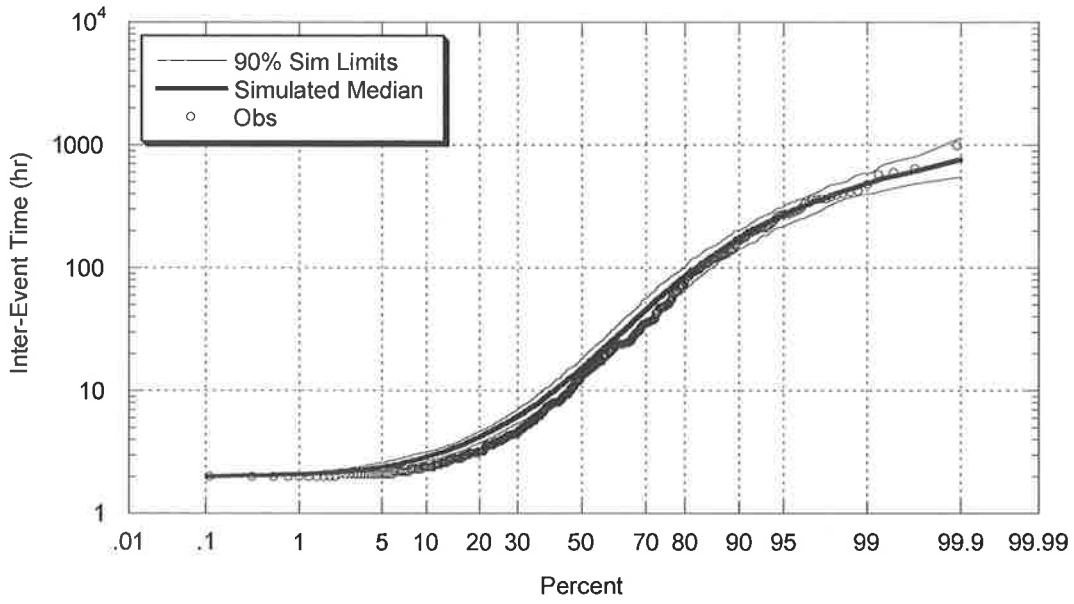

Figure 7.17: Comparison between Observed and Target Simulated March Inter Event Distribution for Richmond. (Master – Sydney; Daily Target – Richmond (Daily))

Figure 7.18 & Figure 7.19 provide further results from Kirkleigh (from Brisbane) and Williamstown (from Adelaide) respectively. Again the large inter-event times are well reproduced with a slight deviation observed for inter-event times less than 24 hours.

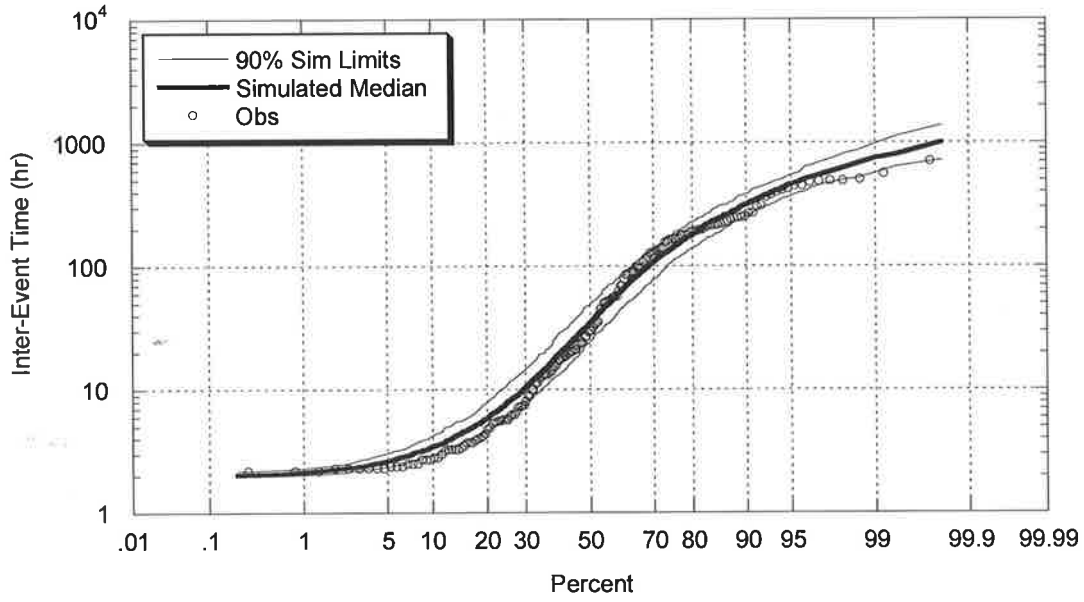

Figure 7.18: Comparison between Observed and Target Simulated September Inter Event Distribution. (Master – Brisbane; Daily Target – Kirkleigh (Daily))

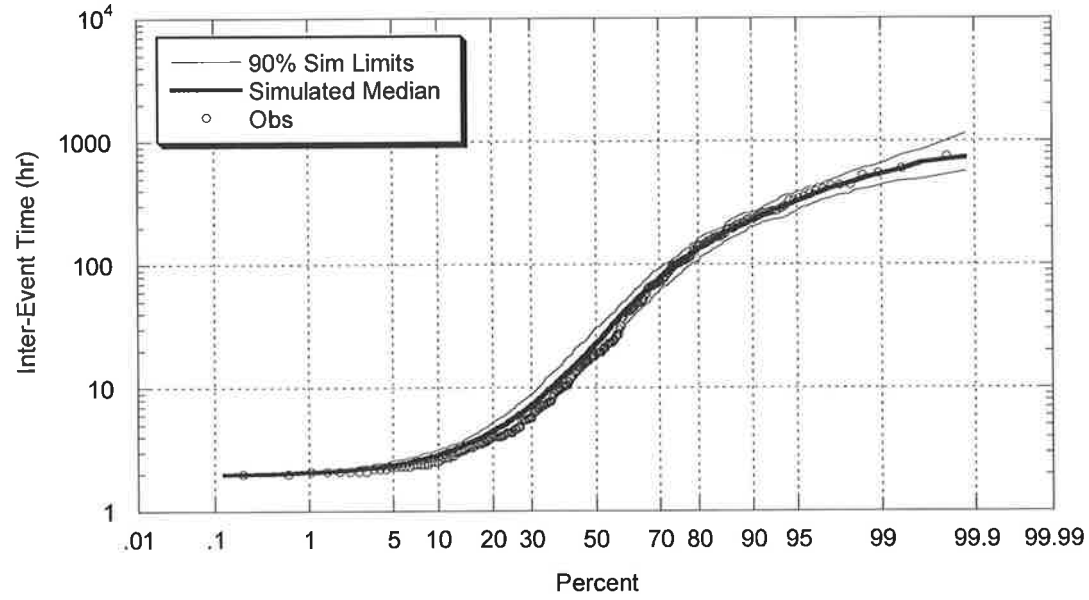

Figure 7.19: Comparison between Observed and Target Simulated December Inter Event Distribution. (Master – Adelaide; Target – Williamstown (Daily))

Additional plots for all sites are displayed in Appendix D. The ability of the model to replicate the storm inter-event is an important result in the context of the overall model performance for two major reasons. First, this result indicates that adjusting the inter-event time regional scaling parameter based on changes to the probability of observing a dry day is justifiable. With the exception of a slight decrease in accuracy for inter-event times less than 24 hours, the distributions are well reproduced. Secondly, (and probably more important for the overall success of the model) this result ensures the model can successfully replicate the storm event time series and as a result the number of storms simulated in a given time period. This is critical for the model if it is aiming to reproduce observed storm depth statistics and aggregated monthly and annual rainfall.

7.3.5 Comparison of Observed and Simulated Bulk Storm Event Distributions – Storm Duration

For the model to capture aggregated rainfall totals such as monthly and annual rainfall, it must be capable of reproducing the monthly storm duration distributions at an event level. With less information in the daily record available for calibration in comparison to inter-event times, it was expected that there would be a less accurate agreement between observed and simulated storm duration distributions. Again, the pluviograph data at the target daily site has provided the information for comparison. Figure 7.20 presents simulated and observed comparisons for data from Richmond (Sydney).

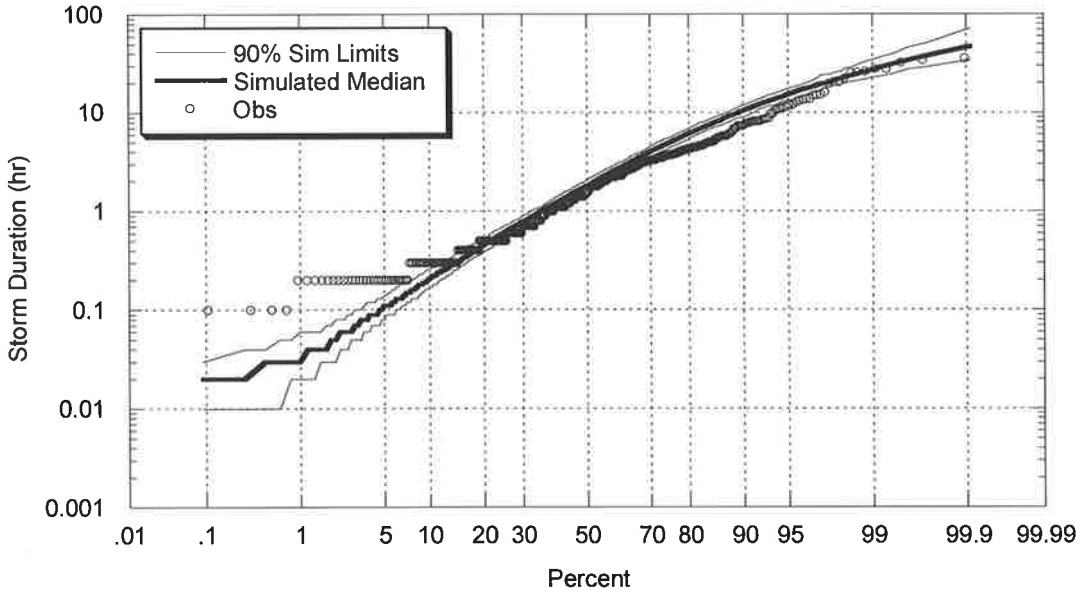

Figure 7.20: Comparison between Observed and Target Simulated March Storm Duration Distribution. (Master – Sydney; Daily Target – Richmond (Daily))

This result suggests an adequate reproduction of the storm event distribution, however there is a slight deviation outside the simulation limits for storms between 4 and 15 hours. Similar results for Williamstown (Adelaide) and Kirkleigh (Brisbane) are seen below in Figure 7.21 & Figure 7.22.

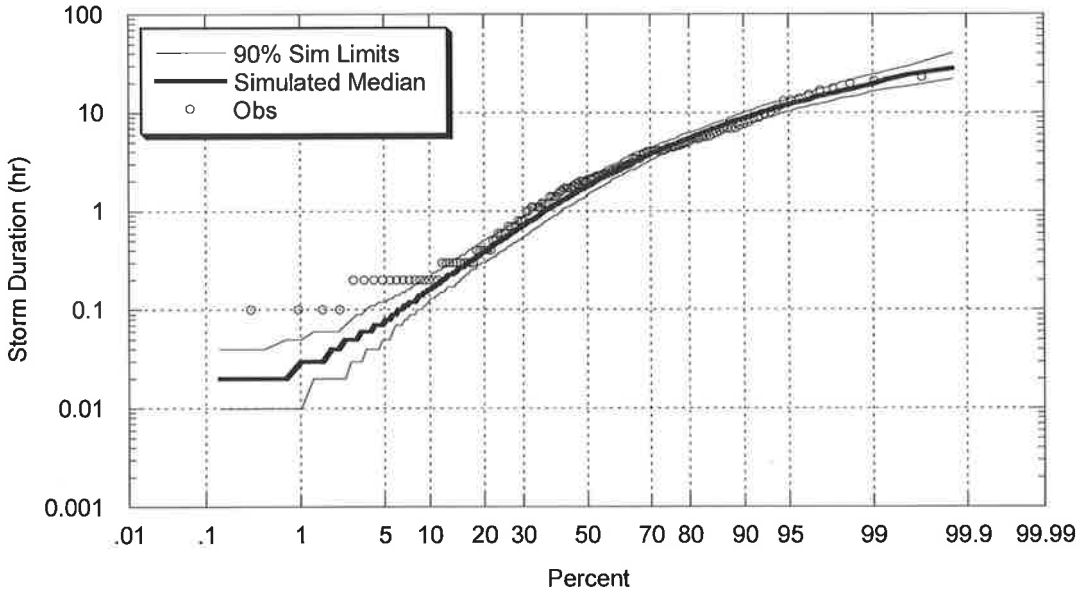

Figure 7.21: Comparison between Observed and Target Simulated June Storm Duration Distribution. (Master – Adelaide; Target – Williamstown (Daily))

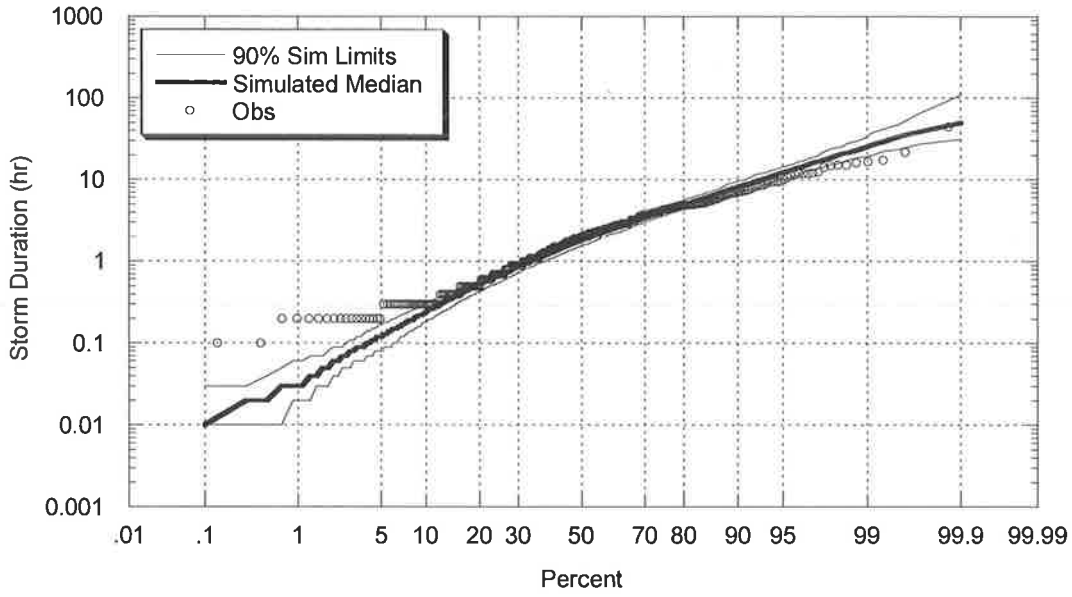

Figure 7.22: Comparison between Observed and Target Simulated December Storm Duration Distribution. (Master – Brisbane; Daily Target – Kirkleagh (Daily))

Results indicate that the reproduction of storm durations is less accurate than the reproduction of inter-event times. This is a result of there being very little information available regarding the length of individual storm events within a daily data record. While the average inter-event time was greater than 24 hours which ensures there is limited information regarding its distribution from a daily record, the average storm duration is less than 24 hours and subsequently only minimal information can be obtained from the daily record. This results in a less accurate re-production of the storm duration characteristics. However, the majority of observed storm events are still contained within or within close proximity to the simulation limits which suggest the regionalisation model is still capable of using daily data for calibration.

One result evident from the figures presented above is the loss in accuracy as the storm duration times approach zero. This is consistent across all sites and is also evident when the model is calibrated at the master site with an extensive pluviograph record as displayed in Figure 7.23 for data from Adelaide.

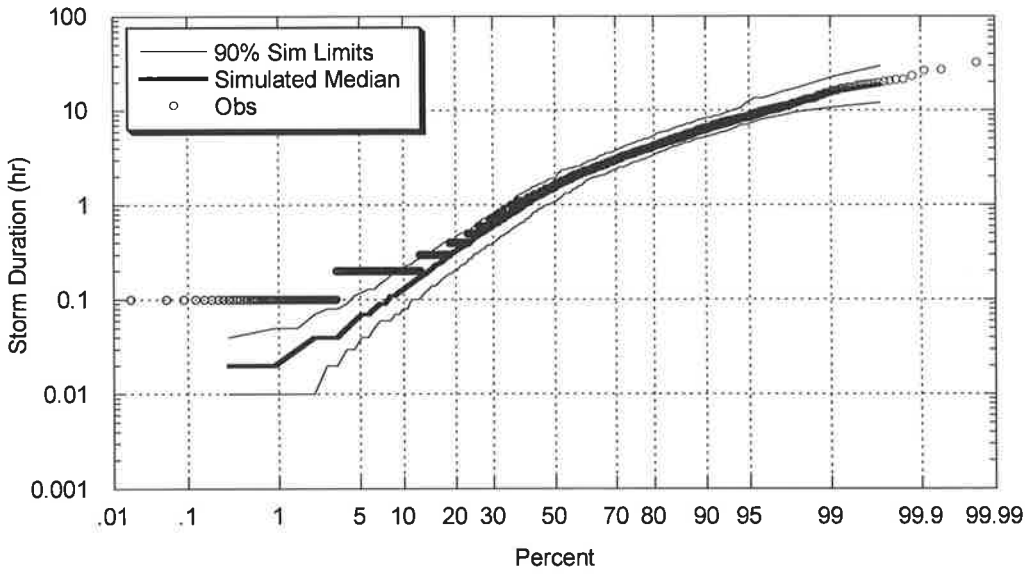

Figure 7.23: Comparison between Full Length Master Observed and Master Simulated March Storm Duration Distribution for Adelaide

As evidenced in this plot the agreement between observed and simulated storm durations decreases as the storm duration approaches zero. This is a result of the rainfall model being free to simulate any storm duration value and consequently produce any storm depth value (in mm). In contrast, rounding occurs in the observed pluviograph data into 0.1mm depth increments. (This is a function of the device). This introduces a slight error between all observed and simulated storm values, which is most evident when comparing small storm depths or very short storm durations. (The majority of small depth rainfall storms are also short duration events). In the case of very small simulated storm depths, the corresponding observed rainfall is either too small to register on the scale (and is therefore excluded from the historical record) or the value is simply recorded with a storm depth of 0.1mm. Consequently, with the majority of small depth rainfall storms also being short duration events, this bias is evident in the comparison of short storm durations with fewer observed short durations in comparison to the model output. This is not a major issue and can be solved by including an appropriate rounding routine on the output data file. However, for the purposes of model validation it is best to compare the raw simulation data.

The other factor influencing the accuracy of these comparisons relates to the scale issues between daily and pluviograph data. As an observed storm event shifts further

and further away from a length equivalent to the daily time step of 24 hours, an individual storm exerts less and less influence on the make up of the daily record. This ensures shorter duration storms have only a minimal influence on the daily data record, making these events very difficult to distinguish and adjust from site to site based on daily data alone. As a result, the reproduction of these smaller storm durations decreases as the storm duration decreases. Nevertheless, and understanding these minor shortcomings, the model has been shown to successfully reproduce the majority of inter-event times and storm durations from only daily data information.

7.3.6 Comparison of Observed and Simulated Bulk Storm Event Distributions – Storm Depth

In a similar manner to the distribution of storm durations, comparisons of the storm depth distributions were expected to be less accurate than those results presented for inter-event times. Figure 7.24 presents results for Kirkleigh (Brisbane) confirming this prediction. Countering the successful reproduction of the majority of the distribution is the small errors evident as the storm depth approaches 0mm. As described earlier this is a result of the available information within a daily record for calibration and the accuracy of the measuring apparatus.

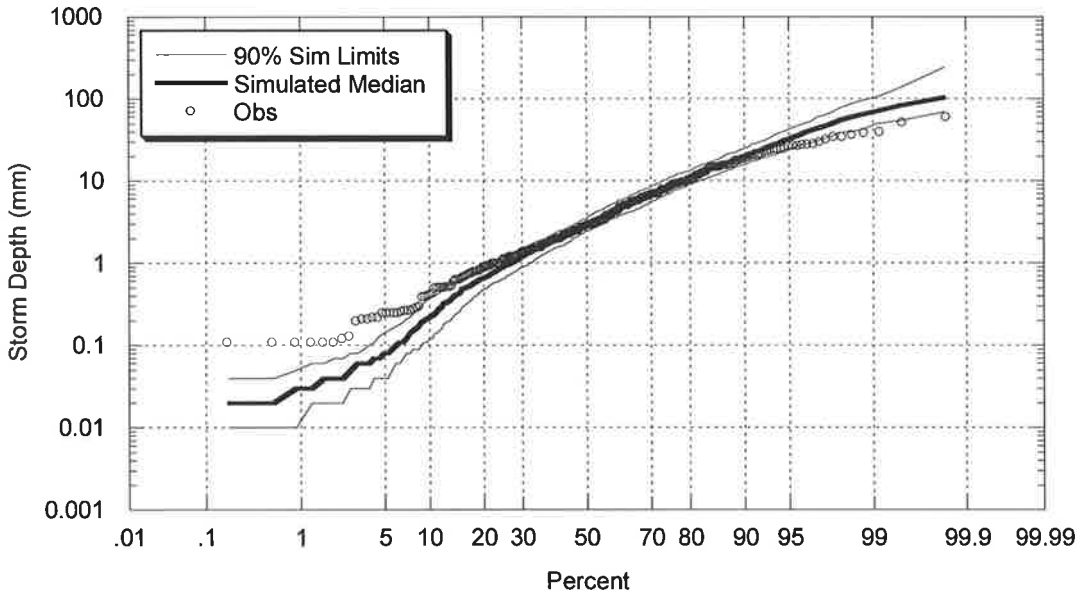

Figure 7.24: Comparison between Observed and Target Simulated September Storm Depth Distribution. (Master – Brisbane; Daily Target – Kirkleigh (Daily))

Similar results are presented below for Richmond (Sydney) and Williamstown (Adelaide) respectively.

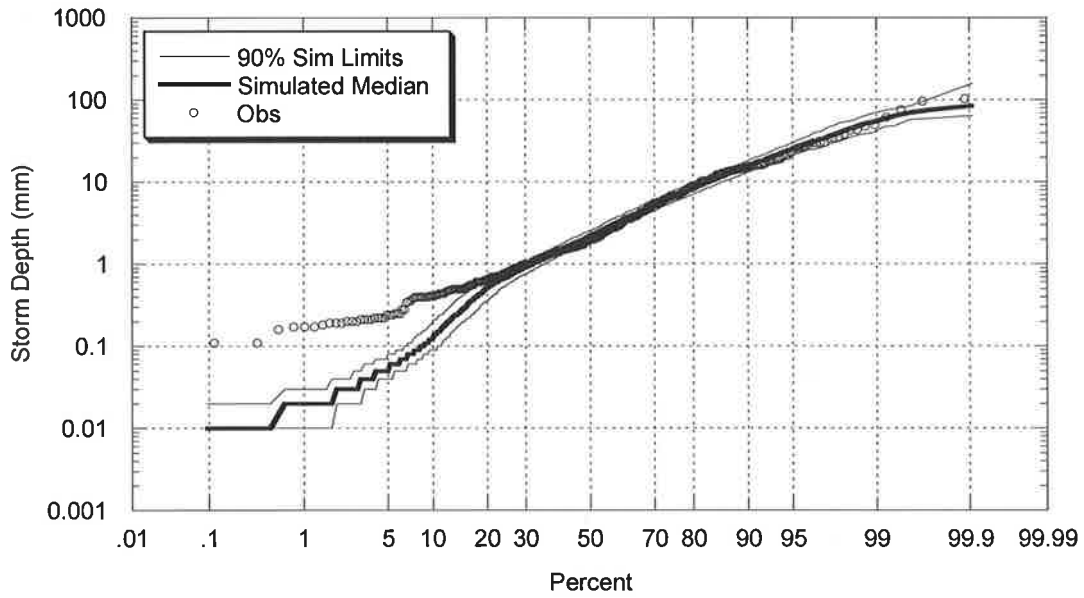

Figure 7.25: Comparison between Observed and Target Simulated December Storm Depth Distribution. (Master – Sydney; Target – Richmond (Daily))

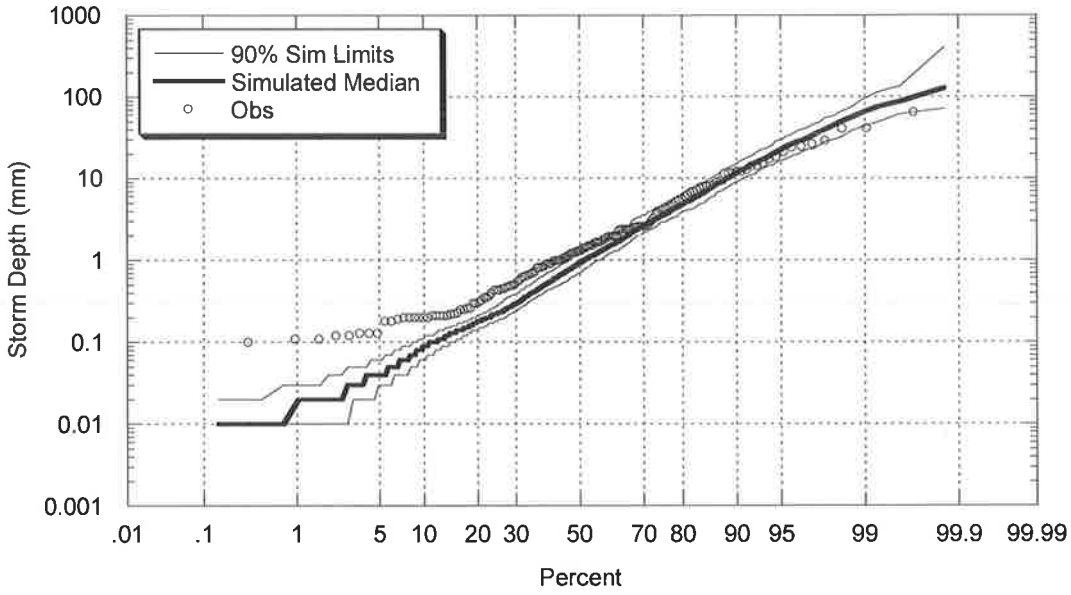

Figure 7.26: Comparison between Observed and Target Simulated March Storm Depth Distribution. (Master – Adelaide; Daily Target – Williamstown (Daily))

The daily regionalisation model is not capable of replicating the bulk storm distributions to the same level of accuracy as the model when calibrated directly to pluviograph data. This was an expected result due to the level of storm information available from a daily rainfall record. Analysis of the depth distribution results indicate that the decrease in accuracy only occurs for depths less than 1mm. For a model calibrated to daily data, this is an exceptional result.

The results presented for the bulk storm characteristics provide evidence that the regional model assumptions were valid and that the model has been well calibrated. A major test of the daily regional model is its ability to reproduce non calibrated monthly and annual rainfall with these results presented in the next section.

7.3.7 Comparison of Observed and Simulated Intensity Frequency Duration Curves

The final comparison between observed and simulated data is the Intensity – Frequency Duration results for each simulated site. Given the results presented for

the pluviograph regionalisation model and the use of the same model assumption (i.e. that the temporal pattern parameters are consistent between master and target sites), it was expected that the IFD results at sites compared in Chapter 5 would again be well reproduced. In almost all cases the IFD results are as good as the pluviograph regionalisation simulations. Figure 7.27 presents the observed and simulated IFD curves for 1 hour, 24 hour and 72 hour for data from Williamstown (Adelaide). Similar results are presented for Kirkleigh (Brisbane) and Richmond (Sydney).

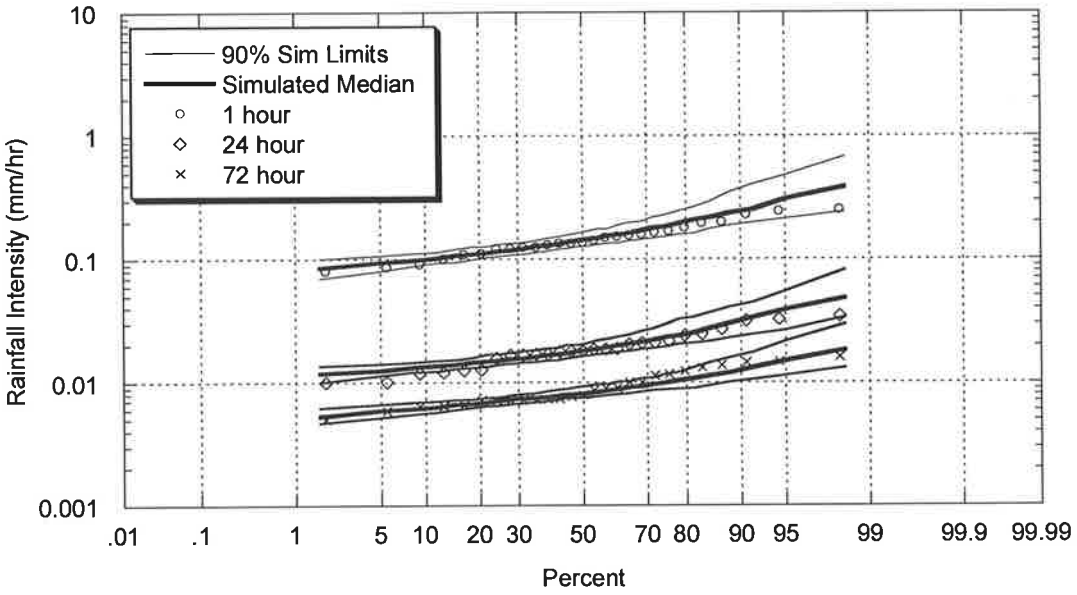

Figure 7.27: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship. (Master – Adelaide; Target – Williamstown (Daily))

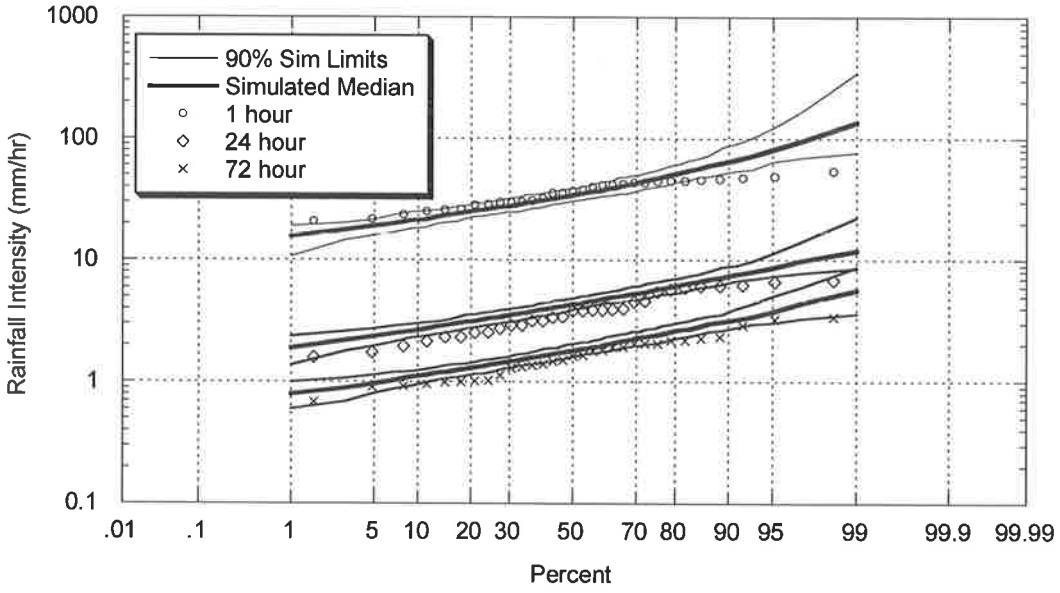

Figure 7.28: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship. (Master – Brisbane; Target – Kirkleigh (Daily))

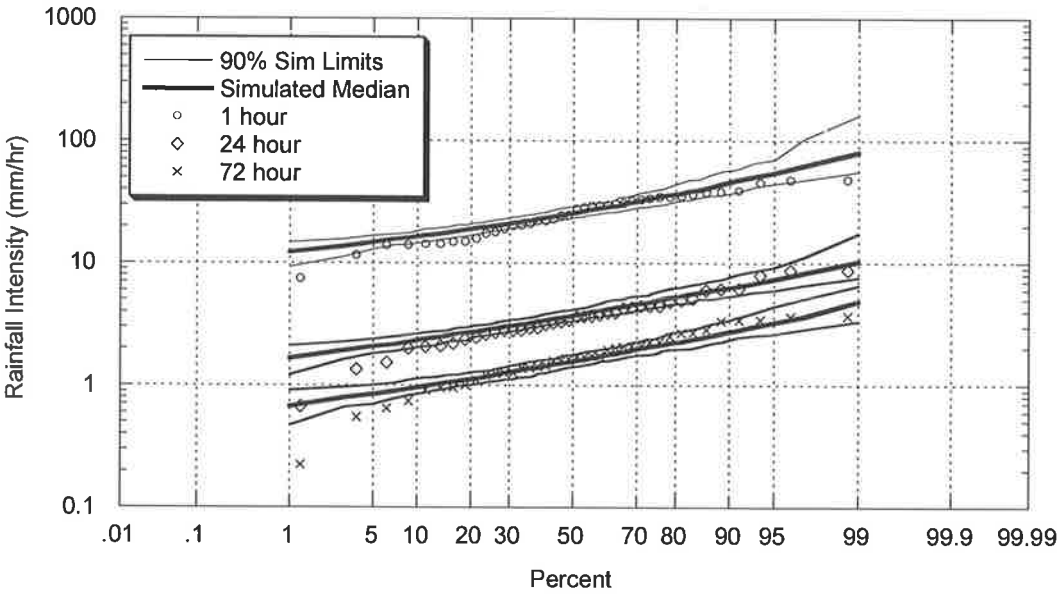

Figure 7.29: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship. (Master – Sydney; Target – Richmond (Daily))

These results suggest that the internal storm characteristics have been well reproduced when simulating using the model calibrated to daily data. Given the parameters at each site were the same as during the pluviograph regionalisation model work and coupled to the success of the daily model in replicating the bulk storm characteristics, this was an expected result.

7.4 Summary

Adapting the master - target linear scaling relationship introduced into the Heneker *et al.* (2001) rainfall model in Chapter 6 for use with Daily data has produced a model capable of simulating long-term synthetic pluviograph records at numerous daily sites around Australia.

The daily regionalisation model uses a master – target relationship coupled with a linear updating factor between sites. The development of a simulated likelihood approach using a non-parametric kernel smoothing technique to estimate the probability distributions provided the basis for calibrating the updating factor when shifting between the master and target sites. As was the case in Chapter 6 for regionalisation with a short pluviograph record, the introduction of an intermediate calibration removes the issues associated with sampling variability between the rainfall records at the master and target sites due to their differing lengths and/or periods of record. The use of this intermediate step and the new regionalisation-scaling factor enabled a successful translation from simulated daily rainfall records at a master calibration site to a target site with only daily data available.

The regionalisation model used two main statistics easily calculated from a daily record to drive the calibration process. The selection of the probability of observing a dry day and the mean daily depth was based on the proven ability of the model to reproduce these non-calibrated statistics during simulation and the ease at which they can be calculated from observed data. Successful reproduction of these calibrated statistics verifies the selection of these values for use in calibration.

Comparisons between the observed and simulated data at various pairs of sites indicate the model's ability to successfully capture the shift in mean and standard deviations of the bulk storm characteristics in order to reproduce various calibrated and non-calibrated statistics within acceptable degrees of accuracy. These results give validity to the underlying structure of the model and the calibration process.

Coupled with the ability to successfully simulate rainfall amounts at various time scales and applying the existing disaggregation process to generate synthetic pluviograph data, this model is now a useful tool capable of being applied to a large number of sites. Coupled with the application to sites with short historical pluviograph records, the new regionalised rainfall model has the potential to be a powerful tool for application in hydrological risk analysis across Australia.

CHAPTER 8

CONCLUSION AND RECOMMENDATIONS

8.1 Overview

The estimation of flood risk relies on joint probability theory where the combination of stochastic inputs such as rainfall and a description of the hydrological/hydraulic runoff process determine the probability distribution of flooding events. To date continuous simulation through a Monte Carlo approach has provided a workable method for deriving flood probability distributions. The Monte Carlo approach uses the idea that a long model simulation will eventually sample almost all possible joint probability interactions (i.e. all combinations of rainfall input and runoff model conditions etc). If this is the case, the derived flood distribution can be viewed as an accurate inference of the true flood distribution.

Despite the theoretical superiority of continuous simulation, in practice designers use a far simpler approach within Australia. The Australian procedure for risk-based hydraulic design is typically described in Australian Rainfall and Runoff (referred to as ARR) (Institution of Engineers Australia, 1987) and is known as the design storm approach. The method for evaluating flood risk probabilities is based on this design storm for which "the intention is to derive a flood of selected probability of exceedance from a design rainfall of the same probability" [ARR, p6]. This approach relies on the assumption that median values of all other variables other than rainfall (such as losses, base flow, temporal patterns and hydrograph model parameters) can be used and still provides an accurate runoff representation. When using this method there is no indication that the design storm approach produces floods with the same exceedance probability as the rainfall. Indeed ARR admits that "there is a need for research to test this approach". Meanwhile designers across Australia

continue to adopt the design rainfall method as the design technique of choice. Without an appropriate working continuous simulation alternative, designers will continue with this approach.

Continuous simulation also relies on the availability and length of historical rainfall records. This is particularly important if we consider the tails of the flood probability distribution where it is unlikely that a 15-year historical record can provide accurate estimates of a 100 year flood event. While the advent of numerous rainfall models provides a method of extending historical rainfall records, without significant historical rainfall data available for calibration, their accuracy is often questionable.

This study was initiated by the desire to provide a rainfall simulation model which could successfully simulate accurate synthetic pluviograph records at sites across Australia with minimal or no historical pluviograph data. This would provide a workable rainfall model solution for application within a continuous simulation flood estimation framework or for use in situations where water volumes are important i.e. Water Sensitive Urban Design approaches to stormwater treatment and disposal, stormwater detention etc. To achieve this objective five aims were developed:

(1) To develop or select a rainfall model capable of simulating synthetic pluviograph data;

(2) To refine and improve the rainfall model by including uncertainty and a Monte Carlo simulation structure ensuring the calibration process is robust and comparison to observed data is accurate.

(3) To verify the accuracy of the rainfall model by analysing its performance and structure at sites with significant pluviograph records for calibration and comparison;

(4) To extend the application of the model to sites with minimal historical pluviograph data available for calibration and finally;

(5) To extend the application of the model to sites with only historical daily data and no pluviograph data available for calibration.

8.2 Stochastic Rainfall Simulation Model

8.2.1 Summary

The review of potential models for continuous simulation of short time increment rainfall focussed on previously reported results, data requirements, structure, parameter estimation and calibration techniques. The objective was to determine which model (if any) was capable of accurately simulating short time increment rainfall and whether such a model could be further developed for application at sites with little or no historical records available for calibration.

The model presented by Heneker *et al.* (2001) had been previously shown to be capable of generating synthetic rainfall data down to time resolutions in the order of minutes. The reproduction of short duration IFD values at most sites validated the effectiveness of the disaggregation procedure, while the overall model structure was validated through comparisons between observed and simulated inter-event times, storm durations and mean annual rainfall. As a result this model was selected for further investigation/development in this study.

A number of shortcomings in the original model presented by Heneker *et al.* (2001) were investigated and improved as part of this study. The resultant model is based on this original work with the following enhancements:

- the incorporation of the Metropolis Algorithm which enables the identification of calibrated parameter distributions and potential parameter correlations. This lead to the discovery of significant correlations within the distribution descriptions for inter-event times and storm durations.
- the inter-event time and storm duration distributions still modelled using a generalised exponential distribution (with the kernel defined using a combination of the generalised Pareto and power law distributions), but with

the exponential parameter value now set to a constant. This has removed the issues associated with the aforementioned parameter correlations and also reduced the number of monthly parameter values requiring calibration for each distribution from four (4) to three (3).

- the average inter-event intensity conditional model now being described by a piece wise linear model with a constant set of break points across all sites. This removes the need for the manual selection of breakpoints and the automatic calibration process has been completed by using a hybrid continuous function to describe the conditional relationship and provide initial parameter values at set breakpoint positions.
- the incorporation of Monte Carlo simulation and parameter sampling uncertainty. A Monte Carlo framework has been included into the simulation model which enables multiple realisations of the model to be generated with minimal effort. Coupled with the addition of parameter sampling uncertainty which describes the accuracy of calibration given the available data set, the simulation model is now able to provide simulation limits providing an improved ability to compare observed and simulated results.

8.2.2 Conclusions and Recommendations

The synthetic rainfall generated by the original model had previously been shown to provide a good representation of observed rainfall over a range of climatic regions. Improvements to the model have further refined its structure and calibration processes to provide a robust and efficient rainfall simulation model. Model validation presented in Chapter 5 showed that with the exception of select aggregation statistics at long time frames (annual variability for example) the synthetic rainfall data produced for all sites was considered satisfactory. In particular:

- the calibrated distributions of inter-event times and storm duration were reproduced satisfactorily with the three (3) parameter distribution model. Monthly parameters were used to account for seasonal variability.
- the conditional average event intensity and storm duration was successfully modelled using the piece wise linear model with set breakpoints. The

simulated rainfall adequately replicated the observed marginal distribution and aggregated statistics across a number of sites.

- the model continued to successfully reproduce non-calibrated statistics such as the Intensity-Frequency-Duration (IFD) curves and aggregated rainfall statistics at a range of time scales. In particular the ability to reproduce the mean daily rainfall and daily dry probability was instrumental in the later development of a calibration routine with daily data.
- it was again observed that similarity in the model parameters within the temporal pattern generator offered an ability to use this model for regionalisation to sites with little or no historical data available for calibration.

8.3 Regionalisation with a Short Pluviograph Record

8.3.1 Summary

Chapter 6 presented a significant leap forward in the application of rainfall models calibrated to pluviograph data. A new regional calibration process was introduced which enabled the model storm event parameters to be calibrated at sites with only a short historical pluviograph data set. The resultant model uses:

- a master - target relationship coupled with a linear updating factor between sites for the regionalisation of inter-event times and storm durations. The master calibration requires a regular calibration with a long (estimated at >30 year) pluviograph record.
- a similar master - target relationship coupled with a linear 'triangle' model for updating the factor between sites for the regionalisation of storm event depths. Event depths were chosen instead of using event intensities directly due to the need to consider the conditional intensity-duration relationship when regionalising. By developing the storm duration relationship between sites first and then considering the event depth relationship, the conditional relationship requirements are taken into account.
- an intermediate calibration step which enabled the developed model to capture any data variations that exist between two pluviograph records as a

result of their different record lengths. This in turn enabled any sampling variability issues between the rainfall records at the master and target sites due to their differing lengths and/or periods of record to be removed during the final calibration of the linear updating factors.

- consistent temporal pattern parameters between sites within a climatic region. This expected outcome enabled the regionalisation process to continue without adjustment to the temporal pattern parameters between master and target sites.

8.3.2 Conclusions and Recommendations

The regionalisation model for inter-event times and storm durations employed a master – target relationship coupled with a linear updating factor between sites. The use of this intermediate step and the new regionalisation scaling factor enabled a successful translation from simulated inter-event time and storm durations at a master calibration site to a target site with only a short pluviograph record available for updating. Test sites were chosen with long historical records to enable thorough examination of simulated statistics while a short sub-set (10 years) was used for calibration purposes. Comparison of calibrated and non-calibrated statistics (observed and simulated) has verified the adopted approach. In particular the model:

- captured IFD statistics at the target site using only the short sub-set of the overall target record during testing. This verifies the adoption of consistent temporal pattern parameters between sites.
- reproduced observed inter-event time and storm duration distributions at the target site. While these were used during the calibration process, the model was only calibrated to 10 years of record and then compared to the statistics obtained from the entire target pluviograph.
- successfully captured the mean annual rainfall totals, while slightly underestimating the annual rainfall variance. This is typical of event based models which use independence criteria to define storm events. Further work in this area is suggested and may be based on recently finished work by Frost (2002).

- satisfactorily reproduced both the observed daily rainfall distribution and the probability of observing zero rain on a given day. These statistics were very important to the further development of the model for calibration at sites with daily data only as presented in Chapter 7.

8.4 Regionalisation with a Daily Record

8.4.1 Summary

Adapting the master - target linear scaling relationship introduced into the Heneker *et al.* (2001) rainfall model in Chapter 7 for use with Daily data has produced a model capable of simulating long-term synthetic pluviograph records at numerous daily sites around Australia. In particular the daily model uses:

- a master – target relationship coupled with a linear updating factor between sites to adjust all storm event variables (inter-event times, storm duration and storm depth). This is based on the same model structure as that for calibrating with short pluviograph records.
- a simulated likelihood approach which uses a non-parametric kernel smoothing technique to estimate the probability distributions of important simulated daily statistics. For calibration purposes the mean daily depth and probability of observing zero rain were chosen due to the models ability to accurately reproduce these values when successfully calibrated with a long term pluviograph. Observed daily values are compared against the estimated density for each iterative simulation providing a method to calibrate the bulk event linear updating factors.
- an intermediate calibration step which enabled the developed model to capture any data variations that exist between the master pluviograph and target daily record as a result of their different record lengths. This in turn enabled any sampling variability issues between the rainfall records at the master and target sites due to their differing lengths and/or periods of record to be removed during the final calibration of the linear updating factors.
- the structure of the pluviograph regionalisation model keeping consistent temporal pattern parameters between sites within a climatic region.

8.4.2 Conclusions and Recommendations

The regionalisation model compared two main statistics easily calculated from a daily record to drive the calibration process. The selection of the probability of observing a dry day and the mean daily depth were based on the proven ability of the model to reproduce these non-calibrated statistics during simulation and the ease at which they can be calculated from observed data. Successful reproduction of these calibrated statistics verifies the selection of these values for use in calibration.

Comparisons between the observed and simulated data at various pairs of sites indicate the model's ability to significantly shift the mean and standard deviation of the bulk storm characteristics in order to reproduce various calibrated and non-calibrated statistics within acceptable degrees of accuracy. These results give validity to the underlying structure of the model and the calibration process. Coupled with the ability to successfully simulate rainfall amounts at various time scales and apply the existing disaggregation process to generate synthetic pluviograph data, this model is now a useful tool capable of being applied to a large number of sites. In particular the model

- satisfactorily reproduced both the observed daily rainfall distribution and the probability of observing zero rain on a given day. These statistics were used during the calibration and their subsequent reproduction during simulation verifies the structure and success of the calibration process.
- adequately captured IFD statistics at the target site. This again verifies the adoption of consistent temporal pattern parameters between sites.
- successfully reproduced observed inter-event times and adequately reproduced storm duration distributions at the target site. These statistics were not used during the calibration process and validated the use of the simulated likelihood approach and adoption of the daily mean rainfall and dry probability as indicators of a good calibration.
- successfully captured the mean annual rainfall totals, while again slightly underestimating the annual rainfall variance.

Coupled with the application to sites with short historical pluviograph records, the new regionalised rainfall model has the potential to be a powerful tool for application in hydrological risk analysis across Australia. It is recommended that future work with this model focus on:

1) Applying the model to the majority of pluviograph and daily rainfall sites across Australia with a view to possibly developing a parameter contour map or similar which would enable application of the model via interpolation between data recording stations.
2) Incorporating an inter-annual persistence model to capture the effects of El Nino and other long term climatic influences within the Australian climate.

CHAPTER 9

REFERENCES

Acreman, M. C. & Wiltshire, S. E. (1989) "The Regions are Dead: Long Live the Regions. Methods of identifying and Dispensing with Regions for Flood Frequency Analysis", IAHS Publ.187: 175-188

Acreman, M. C. (1990) "A simple stochastic model of hourly rainfall for Farnborough, England", Hydrological Sciences Journal 35(2) 119-148.

Adamowski, K (1985) "Nonparametric Kernel Estimation of Flood Frequencies" Water Resources Research 21(11): 1585-1590

Adamowski, K (1996) "Nonparametric Estimation of Low-Flow Frequencies", Journal of Hydrologic Engineering 122:46-49

Adamowski, K (1989) "A Monte Carlo Comparison of Parametric and Nonparametric Estimation of Flood Frequencies", Journal of Hydrology 108: 295-308

Adamowski, K (2000) "Regional Analysis of Annual Maximum and Partial Duration Flood Data by Nonparametric and L-Moment Methods", Journal of Hydrology 229: 219-231

Arnbjerg-Nielsen, K., Harremoes, P. & Spliid, H. (1996) "Interpretaion of Regional Variation of Extreme Values of Point Precipitation in Denmark", Atmospheric Research 42: 99-111

Benson, M. A. (1962) "Factors Influencing the Occurrence of Floods in a Humid Region of Diverse Terrain", U.S. Geological Survey Water Supply Paper 1580B: 61 pp.

Beven, K.J. & Binley, A.M. (1992) "The Future of Distributed Models: Model Calibration and Predictive Uncertainty", Hydrological Processes 6: 279-298

Birikundavyi, S. & Rouselle, J. (1997) "Use of Partial Duration Series for Single-Station and Regional Analysis of Floods", Journal of Hydrologic Engineering April: 68-75

Boughton, W. C. (1984) "A Simple Model for Estimating the Water Yield of Unguaged Catchments", Civil Engineering Transactions Institution of Engineers Australia

Box, G. E. P., & Cox, D. R. (1964) "An Analysis of Transformations (with Discussion)", Journal of the Royal Statistical Society Series B 26: 211-252

Boyd, M. J. (1978) "Regional Flood Frequency Data for N.S.W. Streams", Civil Engineering Transactions Institution of Engineers Australia

Burlando, P. & Rosso, R. (1991) "Comment on "*Parameter Estimaion and Sensitivity Analysis for the Modified Bartlett-Lewis Recangular Pulses Model of Rainfall" by S. Islam et al*", Journal of Geophysical Research 96(D5): 9391-9395

Burlando, P. & Rosso, R. (1993) "Stochastic Models of Temporal Rainfall: Reproducibility, Estimation and Prediction of Extreme Events" in Stochastic Hydrology and its Use in Water Resources Systems Simulation and Optimization, eds J.B. Marco, R. Harboe and J.D. Salas. Kluwer Academic Publishers pp 137-173

Burn, D. H. & Boorman, D. B. (1993) "Estimation of Hydrological Parameters at Unguaged Catchments", Journal of Hydrology 143: 429-454

Burn, D. H. (1990) "Evaluation of Regional Flood Frequency Analysis with a Region of Influence Approach", Water Resources Research 26(10): 2257-2265

Burn, D. H., Zrinji, Z. & Kowalchuk, M. (1997) "Regionalization of Catchments for Regional Flood Frequency Analysis", Journal of Hydrologic Engineering, April: 76-82

Burroughs, W.J. (1999) "The Climate Revealed", Cambridge University Press, New York

Calenda, G. & Napolitano, F. (1999) "Parameter Estiation of Neyman-Scott Processes for Temporal Point Rainfall Simulation", Journal of Hydrology 225: 45-66

Cameron, D., Beven, K. & Tawn, J. (2000) "An Evaluation of Three Stochastic Rainfall Models", Journal of Hydrology 228: 130-149

Cameron, D., Beven, K. & Tawn, J. (2001) "Modelling Extreme Rainfalls Using a Modified Random Pulse Bartlett-Lewis Stochastic Rainfall Model (with Uncertainty)", Advances in Water Resources 24: 203-211

Cameron, D., Beven, K., Tawn, J., Blazkova, S. & Naden, P. (1999) "Flood Frequency Estimation for a Guaged Upland Catchment (with Uncertainty)", Journal of Hydrology 219: 169-187

Cao, C. (1974) "A Contribution to Statistical Depth-Duration-Frequency Analysis", Journal of Hydrology 22: 109-129

Caskey, J.E. (1963) "A Markov Chain Model for the Probability of Precipitation Occurrence in Intervals of Various Lengths", Monthly Weather Review 91(6): 298-301

Chiew, F.H.S., Piechota, T.C., Dracup, J.A. & McMahon, T.A. (1998) "El Nino/Southern Oscillation and Australian Rainfall, Streamflow and Drought: Links and Potential for Forecasting", Journal of Hydrology 204: 138-149

Cho, H. & Chan, D.S.T. (1987) "Mesoscale Atmospheric Dynamics and Modeling of Rainfall Fields", Journal of Geophysical Research 92(D8): 9687-9692

Cho, H. (1985) "Stochastic Dynamics of Precipitation: An Example", Water Resources Research 21(8): 1225-1232

Cong, S., Li, Y., Vogel, J. L. & Schaake, J. C. (1993) "Identification of the Underlying Distribution Form of Precipitation by Using Regional Data", Water Resources Research 29(4): 1103-1111

Cooley, H.W. & Lohnes, P. R., (1971). "Multivariate Data Analysis" Wiley, New York, N.Y.

Cowpertwait, P. S. P., O'Connell, P. E., Metcalfe, A. V. & Mawdsley, J. A. (1996a) "Stochastic Point Process Modelling of Rainfall. I Single-site Fitting and Validation", Journal of Hydrology 175: 17-46

Cowpertwait, P. S. P., O'Connell, P. E., Metcalfe, A. V. & Mawdsley, J. A. (1996b) "Stochastic Point Process Modelling of Rainfall. II Regionalisation and Disaggregation", Journal of Hydrology 175: 47-65

Cowpertwait, P.S.P (1991a) "The Stochastic Generation of Rainfall Time Series", PhD Thesis, University of Newcastle upon Tyne

Cowpertwait, P.S.P (1991b) "Further Developments of the Neyman-Scott Clustered Point Process for Modeling Rainfall", Water Resources Research 27(7): 1431-1438

Cowpertwait, P.S.P., O'Connell, P.E., Metcalfe, A.V. & Mawdsley, J.A. (1996) "Stochastic Point Process Modelling of Rainfall I. Single-Site Fitting and Validation", Journal of Hydrology 175: 17-46

Cox, D. R. & Isham, V. (1980) "Point Processes" Chapman and Hall, London

Dalrymple, T. (1960) "Flood Frequency Analyses", U.S. Geological Survey Water Supply Paper 1543A: 11-51

DeCoursey, D.G. & Deal, R.B., (1974) "General Aspects of Multi-Variate Analysis with Applications to Some Problems in Hydrology", Proceedings Symposium on Statistical Hydrology. Misc. Pub. No. 1275, USDA-ARS, pp. 47-68

DeCoursey, D.G., (1973) "Objective Regionalisation of Peak Flow Rates" in: Floods and Droughts. Proceedings of the Second International Symposium in Hydrology, September 11-13, 1972, Fort Collins, Colorado. Water Resources Publications, Fort Collins, Colorado, pp. 395-405.

DeGaetano, A. T. (1998) "A Smirnov Test-Based Clustering Algorithm with Application to Extreme Precipitation Data", Water Resources Research 34(2): 169-176

Duan, Q., Sorooshian, S. & Gupta, V. (1992) "Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models", Water Resources Research 28(4): 1015-1031

Eagleson, P. (1972) "Dynamics of Flood Frequency", Water Resources Research 8: 878-898

Eagleson, P.S.(1978) "Climate, Soil and Vegetation : 2. The Distribution of Annual Precipitation Derived From Observed Storm Sequences", Water Resources Research 14(5) : 713-721

El-Jabi, N., Ashkar, F. & Hebabi, S. (1998) "Regionalisation of floods in New Brunswick (Canada)", Stochastic Hydrology and Hydraulics 12: 65-82

Entekhabi, D., Rodriguez-Iturbe, I. & Eagleson, P. (1989) "Probabilistic Representation of Temporal Rainfall Processes by a Modified Neyman-Scott Rectangular Pulses Model: Parameter Estimation and Validation", Water Resources Research 25(2): 295-302

Feyerherm, A.M. & Bark, L.D. (1967) "Goodness of Fit of a Markov Chain Model for Sequences of Wet and Dry Days", Journal of Applied Meteorology 6: 770-773

Foufoula-Georgiou, E. & Guttorp, P. (1986) "Compatibility of Continuous Rainfall Occurrence Models with Discrete Rainfall Observations", Water Resources Research 22(8):1316-1322

Foufoula-Georgiou, E. & Lettenmaier, D.P. (1987) "A Markov Renewal Model for Rainfall Occurrences", Water Resources Research 23(5): 875-884

Frost, A. (2004) "Spatio-Temporal Hidden Markov Models for Incorporating Interannual Variability in Rainfall" PhD University of Newcastle, Australia

Frost, A., Jennings, S., Thyer, M., Lambert, M. & Kuczera, G. (2000) "Floods, Droughts and Everything Else In Between", Proceedings from the 3rd International Hydrology and Water Resources Symposium of the Institute of Engineers, Australia, Perth

Gabriel, K.R. & Neumann, J. (1957) "On a Distribution of Weather Cycles by Lengths", Quarterly Journal of the Royal Meteorological Society 83: 375-380

Gabriel, K.R. & Neumann, J. (1962) "A Markov Chain Model for Daily Rainfall Occurrences at Tel Aviv", Quarterly Journal of the Royal Meteorological Society 88: 90-95

Gelman, A., Carlin, J. B., Stren, H. S. & Rubin, D. B. (1997) "Bayesian Data Analysis", Chapman and Hall, London, Ch 11

Gilman C.S. 1964. "Handbook of Applied Hydrology", Section 9 Rainfall, V.T. Chow, ed. New York, NY: McGraw-Hill Book Co.

Grace, R.A. & Eagleson, P.S. (1966) "The Synthesis of Short-Time-Increment Rainfall Sequences" Hydrodynamics Lab. Report 91, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Grace, R.A. & Eagleson, P.S. (1967) "A Model for Generating Synthetic Sequences of Short-Time-Interval Rainfall Depths" Proceedings International Hydrology Symposium, Fort Collins, Colorado: pp. 268-276

Grayman, W.M. & Eagleson, P.S. (1969) "Streamflow record length for modelling catchment dynamics" Hydrodynamics Lab. Report 114, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Green, J.R. (1964) "A Model for Rainfall Occurrence" Journal Royal Statistics Society B 26: pp.354-353

Green, J. R. (1965) "Two Probability Models for Sequences of Wet or Dry Days" Monthly Weather Review 93: 155-156

Guame, E., Villeneuve, J-P. & Desbordes, M. "Uncertainty Assessment and Analysis of the Calibrated Parameter Values of an Urban Storm Water Quality Model", Journal of Hydrology 210: 38-50

Guetzkow, L. C. (1977) "Techniques for Estimating Magnitude and Frequency of Floods in Minnesota" U.S. Geological Survey Water Resources Invsestigations 77-31, 33 pp

Guo, S.L. (1991) "Nonparametric Variable Kernel Estimation with Historical Floods and Paleoflood Information", Water Resources Research 27(1): 91-98

Gyasi-Agyei, Y. & Willgoose, G. R. (1997) "A Hybrid Model for Point Rainfall Modeling" Water Resources Research 33(7): 1699-1706

Gyasi-Agyei, Y. & Willgoose, G. R. (1999) "Generalisation of a Hybrid Model for Point Rainfall" Journal of Hydrology 219: 218-224

Gyasi-Agyei, Y. (1999) "Identification of regional parameters of a stochastic model for rainfall disaggregation" Journal of Hydrology 223: 148-163

Haan, C. T., Allen, D. M. & Street, J. O. (1976) "A Markov Chain Model of Daily Rainfall" Water Resources Research 12(3): 443-449

Hanson, K. M. (1999) "A Framework for Assessing Uncertainties in Simulation Predictions", Physica D 133: 179-188

Hastings, W. K. (1970) "Monte Carlo Sampling Methods Using Markov Chains and Their Applications", Biometrika 57: 97-109

Hay, L., McCabe, G.J., Wolock, D.M. & Ayers, M.A. (1991) "Simulation of precipitation by weather type analysis", Water Resources Research 27: 493-501

Heneker, T (2002) "An Improved Engineering Design Flood Estimation Technique: Removing the Need to Estimate Initial Loss" PhD Adelaide University, Australia

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

Hopkins, J.W. & Robillard, P. (1964) "Some Statistics from the Canadian Prairie Provinces", Journal of Applied Meteorology 3: 600-602

Hosking, J. R. M. (1986) "The Theory of Probability Weighted Moments", Research Report RC12210, IBM Research Division, Yorktown Heights, N.Y.

Hosking, J. R. M. (1990) L-Moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics", Journal of the Royal Statistical Society, Series B, 52: 105-124

Hosking, J. R. M. & Wallis, J. R. (1988) "The Effect of Intersite Depedence on Regional Flood Frequency Analysis", Water Resources Research 24(4): 588-600

Hosking, J. R. M. & Wallis, J. R. (1993) " Some Statistics Useful in Regional Frequency Analysis", Water Resources Research 29(2): 271-281

Hosking, J. R., Wallis, J. R. & Wood, E. F. (1985) "An Appraisal of the Regional Flood Frequency Procedure in the UK Flood Studies Report", Hydrological Science Journal 30(1): 85-109

Hughes, J. P. & Guttorp, P. (1994) "A Class of Stochastic Models for Relating Synoptic Atmospheric Patterns to Regional Hyrologic Phenomena", Water Resources Research 30(5): 1535-1546

Hughes, J.P. & Guttorp, P. (1994) "A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena", Water Resources Research 30(5): 1535-1546

Hutchinson, M. F. (1990) "A Point Rainfall Model Based on a Three-State Continuous Markov Occurrence Process", Journal of Hydrology 114: 125-148

Institute of Engineers Australia (1987) "Australian Rainfall and Runoff Volume 1 & 2", Institute of Engineers Australia.

Islam, S., Entekhabi, D., Bras, R.L. & Rodriguez-Iturbe, I. (1990) "Parameter Estimation and Sensitivity Analysis for the Modified Bartlett-Lewis Rectangular Pulses Model of Rainfall", Journal of Geophysical Research 95(D3): 2093-2100

Jin, M. & Stedinger, J. R. (1989) "Flood Frequency Analysis with Regional and Historical Information", Water Resources Research 25: 925-936

Jimoh, O. D. & Webster, P. (1999) "Stochastic Modelling of Daily Rainfall in Nigeria: Intra-Annual Variation of Model Parameters", Journal of Hydrology 222: 1-17

Johnson, S. C. (1967) " Hierarchical Clustering Schemes", Psychometrika 32:241-254

Jones, P.G. & Thornton, P.K. (1999) "Fitting a third-order Markov rainfall model to interpolated climate surfaces", Agricultural and Forest Meteorology 97:213-231

Kavvas, M.L. & Delleur, J.W. (1981) "A Stochastic Cluster Model of Daily Rainfall Sequences", Water Resources Research 17(4) : 1151-1160

Koutsoyiannis, D. and Foufoula-Georgiou, E. (1993), "A scaling model of a storm hyetograph", Water Resources Research, 29(7), 2345-2361.

Koutsoyiannis, D. and Onof, C. (2001), "Rainfall Disaggregation using Adjusted Procedures on a Poisson Cluster Model", Journal of Hydrology 246: 109-122

Koutsoyiannis, D., C. Onof, & H. S. Wheater (2003), "Multivariate Rainfall Disaggregation at a Fine Time Scale", Water Resources Research 39(7): 1-18

Koutsoyiannis, D. & Pachakis, D. (1996) "Deterministic chaos versus stochasticity in analysis and modeling of point rainfall series", Journal of Geophysical Research 101(D21) 26441-26451.

Koutsoyiannis, D. and Xanthopoulos, T. (1990) "A dynamic model for short-scale rainfall disaggregation", Hydrological Sciences Journal 35(3) 303-322.

Kroll, C. N. & Stedinger, J. M. (1998) "Regional Hydrologuc Analysis: Ordinary and Generalized Least Squares Revisited", Water Resources Research 34: 121-128

Kroll, C. N. & Stedinger, J. M. (1999) "Development of Ragional Regression Relationships with Censored Data", Water Resources Research 35(3): 775-784

Kroll, C.N. & Stedinger, J.R. (1998) "Regional hydrologic analysis: Ordinary and generalized least squares revisited", Water Resources Research 34(1) : 121-128

Kroll, C.N. & Stedinger, J.R. (1999) "Development of regional regression relationships with censored data", Water Resources Research 33(3) : 775-784

Kuczera, G. (1983) "Improved Parameter Inference in Catchment Models 1. Evaluating Parameter Uncertainty", Water Resources Research 19(5): 1151-1162

Kuczera, G. & Parent, E. (1998) "Monte Carlo Assessment of Parameter Uncertainty in Conceptual Catchment Models: The Metropolis Algorithm", Journal of Hydrology 211: 69-85

Lall, U., Rajagopalan, B. & Tarboton, D. G. (1996) "A Nonparametric Wet/Dry Model for Resampling Daily Precipitation", Water Resources Research 9: 2803-2823

Lambert, M & Kuczera, G. (1996) "A statistical model of rainfall and temporal patterns", Stochastic Hydraulics '96, Proceedings of the 7th IAHR International Symposium, pp 317-324, A.A. Balkema, Roterdam, Netherlands.

Lambert, M. & Kuczera, G. (1998) "Seasonal generalized exponential probability models with application to interstorm and storm durations", Water Resources Research 34(1) : 143-148

Lance, G. N. & Williams, W. T. (1967) " A General Theory of Classificatory Sorting Strategies, I. Hierarchical Systems", Computer Journal 9: 373-380

Lawrance, A. J. (1972) "Some Models for Stationary Series of univariate Events", in Stochastic Point Process: Statistical Analysis Theory and Applications, 1st Edition, edited by P.A.W. Lewis, Wiley-Interscience, New York

Lee, P. M. (1989) "Bayesian Statistics – An Introduction" J. Wiley & Sons, Chichester

Lettenmaier, D. P., Wallis, J. R. & Wood, E. F. (1987) "Effect of Regional Heterogeneity on Flood Frequency Estimation", Water Resources Research 23(2): 313-323

Madsen, H., Pearson, C. P. & Rosjberg, D. (1997) "Comparison of Annual Maximum Series and Partial Duration Series Methods for Modeling Extreme Hydrologic Events 2. Regional Modeling", Water Resources Research 33(4): 759-769

Matalas, N. C. & Gilroy, E. J. (1968) "Some Comments on Regionalisation in Hydrologic Studies", Water Resources Research 4(6): 1361-1369

Metcalfe, A.V. (1997) "Statistics in Civil Engineering", Arnold, London

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E. (1953) "Equation state calculations by fast computing machine", Journal of Chemical Physics 21 : 1087-1091

Moon, Y., Lall, U. & Bosworth K. (1993) "A Comparison of Tail Probability Estimators for Flood Frequency Analysis" Journal of Hydrology 151: 343-363

Mosley, M. P., (1981) "Delimitation of New Zealand Hydrologic Regions", Journal of Hydrology 49: 173-192

Nathan, R. J. & McMahon, T. A. (1990) "Identification of Homogeneous Regions for the Purposes of Regionalisation", Journal of Hydrology 121: 217-238

Nelder, J.A. & Mead, R. (1965) "A simplex method for function minimization", Journal of Computing 7(4) : 308-313

Neyman, J.E. & Scott, E.L. (1958) "A Statistical Approach to Problems of Cosmology", Journal Royal Statistics Society B(20) : 1-43

Onof, C. & Wheater, H. S. (1993) "Modelling of British Rainfall Using a Random Parameter Bartlett-Lewis Rectangular Pulse Model", Journal of Hydrology 149: 67-95

Onof, C. & Wheater, H. S. (1994) "Improvements to the Modelling of British Rainfall using a Modified Random Pulse Parameter Bartlett-Lewis Rectangular Pulse Model", Journal of Hydrology 157: 177-195

Park, E. S., Oh, M. & Guttorp, P. (2002) "Multivariate Receptor Models and Model Uncertainty", Chemometrics and Intelligent Laboratory Systems 60: 49-67

Pattison, A. (1965) "Synthesis of Daily Hourly Rainfall Data", Water Resources Research 1(4): 489-498

Pilgrim, D.H. (1987) "Australian Rainfall and Runoff : A guide to flood estimation", The Institution of Engineers, Australia

Rajagopalan, B., Lall, U. & Tarboton, D.G. (1996) "Nonhomogeneous Markov Model for Daily Precipitation", Journal of Hydrologic Engineering, January: 33-40

Raudkivi, A.J. & Lawgun, N. (1974) "Simulation of Rainfall Sequences", Journal of Hydrology 22: 271-294

Restrepo-Posada, P.J. & Eagleson, P.S. (1982) "Identification of Independent Rainstorms", Journal of Hydrology 55: 303-319

Ribeiro-Correa, J., Cavadias, G. S., Clement, B. & Rouselle, J. (1995) "Identification of Hydrological Neighborhoods using Canonical Correlation Analysis", Journal of Hydrology 173: 71-89

Riggs, H. C. (1973) "Regional Analyses of Streamflow Characteristics", Techniques of Water Resources Investigations, Book 4, Ch. B3, US Geological Survey, Washington, D.C.

Rodriguez-Iturbe, I., Cox, D.R. & Isham, V. (1987a) "Some models for rainfall based on stochastic point processes", Proceedings of the Royal Society, London A410 : pp. 269-288

Rodriguez-Iturbe, I., Cox, D.R. & Isham, V. (1988) "A Point Process Model for Rainfall: Further Developments", Proceedings of the Royal Society, London A447 : pp. 283-298

Rodriguez-Iturbe, I., Febres de Power, B. & Valdes, J.B. (1987b) "Rectangular Pulses Point Process Models for Rainfall: Analysis of Empirical Data", Journal of Geophysical Research 92(D8): 9645-9656

Rodriguez-Iturbe, I., Gupta, V.K. & Waymire, E. (1984) "Scale considerations in the modelling of temporal rainfall", Water Resources Research 20(11) : 1611-1619

Roldan, J. & Woolhiser, D.A. (1982) "Stochastic Daily Precipitation Models : 1. A Comparison of Occurrence Processes", Water Resources Research 18(5) : 1451-1459

Rosjberg, D., Madsen, H. & Rasmussen, P. F. (1992) "Prediction in Partial Duration Series with Generalized Pareto-Distirbuted Exceedences", Water Resources Research 28(11): 3001-3010

Rosenblatt, M. (1956) "Remarks on some nonparametric estimates of a density function" Annals of Mathematical Statistics 27(3):832-837

Sariahmed, A. & Kisiel, C.C. (1968) " Synthesis of sequences of summer thunderstorm volumes for the Atterbury watershed in the Tucson area.", Proceedings International Association Hydrologic Sciences Symposium on Use of Analog and Digital Computers in Hydrology, 2: 439-447

Sefton, C. E. M. & Howarth S. M. (1998) "Relationships Between Dynamic Response Characteristics and Physical Descriptors of Catchments in England and Wales", Journal of Hydrology 211: 1-16

Seibert, J (1999) "Regionalisation of parameters for a conceptual rainfall-runoff model", Agricultural and Forest Meteorology 98-99 : 279-293

Shamir, U. (1965) "Probability Distributions for Some Precipitation Variables", M.I.T. Course 1.71 Term Paper, January, 1965

Sharma, A., Lall, U. & Tarboton, D.G. (1998) "Kernel bandwidth selection for a first order nonparametric streamflow simulation model", Statistic Hydrology and Hydraulics 12 : 33-52

Silverman, B.W. (1986) "Density Estimation", Chapman and Hall, London

Slade, J.J., Jr. (1936) "The Reliability of Statistical Methods in the Determination of Flood Frequencies", U.S. Geological Survey, Water Supply Paper 771,: pp 421-432

Small, M.J. & Morgan, D.J. (1986) "The Relationship Between a Continuous-Time Renewal Model and a Discrete Markov Chain Model of Precipitation Occurrence", Water Resources Research 22(10) : 1422-1430

Smith, E.R. & Schreiber, H.A. (1973) "Point Processes of Seasonal Thunderstorm Rainfall, Part 1, Distribution of Rainfall Events", Water Resources Research 9(4): 871-884

Smithers, J. C. & Schulze, R. E. (2001) "A Methodology for the Estimation of Short Duration Design Storms in South Africa using a Regional Approach based on L-Moments", Journal of Hydrology 241: 42-52

Sokal, R. R. & Sneath, P. H. A. (1963) "Principles of Numerical Taxonomy", W. H. Freeman and Co., San Francisco, California

Srikanthan, R. & McMahon, T.A. (1985) "Stochastic Generation of Rainfall and Evaporation Data", Australian Water Resources Council Technical Paper No.84, Department of Resources and Energy, Australian Water Resources Council.

Tasker, G. D., (1982) "Comparing Methods of Hydrologic Regionalisation", Water Resources Bulletin 18(6): 965-970

Tasker, G. D., Hodge, S. A. & Barks, C. S. (1996) "Region of Influence Regression for Estiating the 50-Year Flood at Ungaged Sites", Water Resources Bulletin 32(1): 163-170

Thom, H.G. (1975) "A Note on the Gamma Distribution", Monthly Weather Review 86(4): 117-122

Thomas, D. M. & Benson, M. A. (1970) "Generalisation of Streamfow Characteristics from Drainage-Basin Characteristics", U.S. Geological Survey Water Supply Paper 1975

Thomas, D. M. & Cervione, M. A. (1970) "A Proposed Streamflow Data Program for Connecticut", Conn. Water Resources Bulletin 23

Thyer, M. & Kuczera, G. (1999) "Modelling Long Term Persistence in Rainfall Time Series : Sydney Rainfall Case Study", Proceedings 25th Hydrology and Water Resources Symposium, July 1999 Queensland Australia

Todorovic, P. & Yevjevich, V. (1967) "A Particular Stochastic Process as Applied to Hydrology", The International Hydrology Symposium, Colorado State University, Fort Collins, Colorado

Todorovic, P. & Yevjevich, V. (1969) "Stochastic processes of precipitation", Colorado State University Hydrology Paper 35: 1-61

Todorovic, P. (1968) "A Mathematical Study of Precipitation Phenomena", Report CER 67-68 PT65, Engineering Research Centre, Colorado State University, Fort Collins, Colorado

Valdés, J.B., Rodriguez-Iturbe, I. & Gupta, V.K. (1985) "Approximations of Temporal Rainfall from a Multidimensional Model", Water Resources Research 21(8): 1259-1270

Valdés, J.B., Vicéns, G.J. & Rodriguez-Iturbe, I. (1979) "Choosing Among Alternative Hydrologic Regression Models", Water Resources Research 15(2): 347-358

Van Straten, G. & Keesman, K.J. (1991) "Uncertainty Propogation and Speculation in Projective Forecasts of Environmental Change: a Lake-Eutrophication Example", Journal of Forecasting 10: 163-190

Velghe, T., Troch, P.A., De Troch, F.P. & Van de Velde, J. (1994) "Evaluation of Cluster-Based Rectangular Pulses Point Process Models For Rainfall", Water Resources Research 30(10): 2847-2857

Vere-Jones, D. (1970) "Stochastic Models for Earthquake Occurrence", Journal of the Royal Statistical Society, Series A 32(1): 1-62

Vogel, R. M. & Kroll, C. N. (1992) "Regional Geohydrologic-Geomorhpic Relationships for the Estimation of Low-Flow Statistics", Water Resoures Research 28: 2451-2458

Wand, M.P. & Jones, M.C. (1995) "Kernel Smoothing", Chapman & Hall, London

Wandle, S. W. (1977) "Estimating the Magnitude and Frequency of Floods on Natural-Flow Streams in Massachusetts", U.S. Geological Survey Water Resources Investigations 77-39, 27 pp

Waymire, E. & Gupta, V. K., (1981a) "The Mathematical Structure of Rainfall Representations 1. A Review of the Stochastic Rainfall Models", Water Resources Research 17(5): 1261-1272

Waymire, E. & Gupta, V. K., (1981b) "The Mathematical Structure of Rainfall Representations 2. A Review of the Theory of Point Processes", Water Resources Research 17(5): 1273-1285

Waymire, E. & Gupta, V. K., (1981c) "The Mathematical Structure of Rainfall Representations 3. Some Applications of the Point Process Theory to Rainfall Processes", Water Resources Research 17(5): 1287-1294

Weiss, L.L. (1964) "Sequences of Wet and Dry Days Described by a Markov Chain Model", Monthly Weather Review 93: 511-516

Wigley, T.M.L., Lough, J.M. & Jones, P.D. (1984) "Spatial Patterns of Precipitation in England and Wales and a Revised Homogeneous England and Wales Precipitation Series", Climatology 4: 1-25

Wong, T. H. F. (1996) "Synthetic Generation of Storm Sequence using a Multi-Module Data Generation Technique", 23rd Hydrology and Water Resources Symposium, Hobart, Tasmania 21-24 May.

Woolhiser, D. A. & Osborn, H. B. (1985), "A Stochastic Model of Dimensionless Thunderstorm Rainfall". Water Resources Research, 21(4), 511-522.

Wotling, G., Bouvier, Ch., Danloux, J. & Fritsch, J. M. (2000) "Regionalization of Extreme Precipitation Distribution unsing the Principal Components of the Topographical Environment", Journal of Hydrology 223: 86-101

Yevjevich, V. (1974). "Determinism and Stochasticity in Hydrology", Journal of Hydrology 22: 225-238

A High Resolution Point Rainfall Model Calibrated to Short Pluviograph or Daily Rainfall Data

Appendices

Shane Anthony Jennings

December 2006

Ph.D. Thesis

Department of Civil and Environmental Engineering
Adelaide University
Australia

TABLE OF APPENDICES

APPENDIX A

Rainfall Data Site Details and Recording Station Information

A.1 South Australia

Station name: Adelaide Aero (23034)

State: South Australia

Elevation: 6 metres

Latitude: 34.96° South

Longitude: 138.53° East

Annual Rainfall: 453.4mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.60	0.65	0.69	1.21	1.79	1.82	2.03	1.63	1.55	1.27	0.83	0.79

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.84	0.87	0.82	0.70	0.57	0.54	0.47	0.48	0.55	0.64	0.74	0.78

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
4.73	3.71	5.69	9.02	13.47	13.96	16.36	16.07	13.38	11.02	7.71	6.82

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
18.11	17.88	21.53	36.14	55.61	54.70	63.03	50.48	46.53	39.28	24.46	23.83

Station name: Rosedale (23343)

State: South Australia

Elevation: 115 metres

Latitude: 34.56° South

Longitude: 138.83° East

Annual Rainfall: 468.0mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.59	0.56	0.69	1.19	1.71	1.86	2.05	1.90	1.90	1.51	0.99	0.74

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.87	0.89	0.84	0.75	0.61	0.55	0.48	0.48	0.57	0.66	0.77	0.81

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
3.26	2.35	3.67	5.31	8.46	9.11	10.80	10.91	8.73	7.06	4.73	4.11

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
14.16	11.69	15.99	25.22	36.64	37.33	42.55	39.51	38.08	31.36	19.84	16.16

Station name: Stirling (23785)

State: South Australia

Elevation: 496 metres

Latitude: 35.00° South

Longitude: 138.72° East

Annual Rainfall: 1118.2mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.19	1.33	1.75	3.24	4.26	3.96	5.66	4.78	4.00	2.98	2.06	1.52

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.76	0.81	0.69	0.58	0.47	0.47	0.36	0.37	0.45	0.55	0.65	0.71

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
7.14	5.23	9.27	11.91	15.27	14.41	18.91	18.64	15.59	13.18	9.96	8.64

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
35.16	35.82	51.81	92.32	121.66	107.29	167.10	140.47	114.23	87.71	58.86	44.87

Station name: Williamstown (23763)

State: South Australia

Elevation: 395 metres

Latitude: 34.71° South

Longitude: 138.94° East

Annual Rainfall: 755.7mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.87	0.81	0.94	1.72	2.70	3.31	4.09	3.37	2.93	2.21	1.31	1.23

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.81	0.84	0.78	0.67	0.54	0.45	0.38	0.38	0.49	0.60	0.71	0.74

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
5.97	4.67	6.88	9.79	13.67	15.67	17.79	18.06	14.39	11.52	8.33	7.94

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
26.96	22.83	29.06	50.52	80.16	94.00	117.78	97.32	82.59	63.53	37.67	36.95

A.2 Queensland

Station name: Brisbane RO (40214)

State: Queensland

Elevation: 38.0 metres

Latitude: 27.48° South

Longitude: 153.03° East

Annual Rainfall: 1146.4mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
5.11	5.79	4.79	3.18	2.42	2.26	1.82	1.42	1.47	2.49	3.26	4.17

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.58	0.52	0.52	0.62	0.68	0.74	0.76	0.78	0.75	0.70	0.66	0.62

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
12.60	13.13	14.24	10.97	9.43	7.60	6.99	6.52	7.16	9.12	9.78	11.17

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
152.57	157.41	141.94	91.73	72.18	65.24	53.92	42.45	42.40	74.21	94.27	123.56

Station name: Brisbane AMO (40223)

State: Queensland

Elevation: 4.0 metres

Latitude: 27.42° South

Longitude: 153.11° East

Annual Rainfall: 1185.4mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
5.16	6.18	4.47	3.01	3.19	2.37	2.02	1.38	1.16	3.05	3.24	4.04

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.58	0.49	0.54	0.63	0.66	0.74	0.76	0.78	0.77	0.68	0.67	0.63

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
12.77	14.12	13.88	10.88	10.49	7.67	7.35	6.77	6.98	10.06	9.90	11.18

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
156.80	171.19	135.75	88.62	96.82	71.19	62.56	42.68	34.92	94.46	95.70	122.71

Station name: Kirkleigh (40318)

State: Queensland

Elevation: 103.6 metres

Latitude: 27.03° South

Longitude: 152.56° East

Annual Rainfall: 912.6mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
4.70	4.02	2.83	2.61	2.18	1.60	1.79	1.20	1.22	2.54	2.93	4.33

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.67	0.56	0.66	0.73	0.71	0.79	0.79	0.81	0.83	0.74	0.71	0.69

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
7.46	8.94	7.74	6.20	6.31	4.89	5.00	4.40	3.54	5.91	6.40	6.77

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
106.18	81.78	64.34	59.75	47.64	36.93	41.88	27.56	25.87	57.65	64.20	93.52

A.3 Victoria

Station name: Melbourne (86071)

State: Victoria

Elevation: 31.2 metres

Latitude: 37.80° South

Longitude: 144.97° East

Annual Rainfall: 653.2mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.57	1.69	1.65	1.93	1.85	1.67	1.56	1.63	1.97	2.17	2.00	1.91

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.73	0.74	0.70	0.60	0.52	0.48	0.47	0.48	0.50	0.54	0.61	0.66

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
8.26	7.38	9.36	11.90	14.95	15.72	16.39	16.26	15.04	14.26	11.77	10.46

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
48.39	47.26	50.86	57.89	57.25	50.08	48.22	50.47	59.04	67.37	59.74	58.92

Station name: East Sale (85072)

State: Victoria

Elevation: 4.6 metres

Latitude: 38.11° South

Longitude: 147.13° East

Annual Rainfall: 611.1mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.56	1.45	1.71	1.59	1.80	1.61	1.35	1.54	1.83	1.97	2.12	1.82

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.72	0.71	0.67	0.57	0.51	0.44	0.49	0.47	0.48	0.52	0.58	0.66

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
8.65	7.93	10.02	12.63	15.25	16.75	15.95	16.32	15.35	14.54	12.42	10.51

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
47.54	40.17	52.01	47.11	55.72	48.41	41.76	46.96	53.85	60.07	62.35	55.40

Station name: Ellinbank (85240)

State: Victoria

Elevation: 167.0 metres

Latitude: 38.25° South

Longitude: 145.93° East

Annual Rainfall: 1092.9mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2.06	1.64	2.21	2.88	3.48	3.51	3.71	3.69	3.91	3.47	3.09	2.75

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.69	0.74	0.63	0.56	0.44	0.40	0.36	0.38	0.40	0.48	0.53	0.61

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
9.58	7.23	11.33	13.10	17.03	17.33	19.30	18.68	17.38	15.63	13.65	11.75

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
63.04	45.53	68.23	85.96	106.37	101.21	112.29	110.82	113.76	104.46	90.02	82.44

Station name: Laverton (87031)

State: Victoria

Elevation: 16.0 metres

Latitude: 37.86° South

Longitude: 144.76° East

Annual Rainfall: 557.3mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.36	1.67	1.14	1.57	1.62	1.31	1.30	1.49	1.75	1.91	1.73	1.51

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.76	0.75	0.72	0.62	0.53	0.48	0.47	0.49	0.50	0.54	0.60	0.68

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
7.25	6.92	8.70	11.09	14.34	15.48	16.12	15.71	14.68	13.93	11.86	9.81

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
41.33	46.42	34.68	46.36	49.30	39.17	39.83	45.38	51.48	58.27	50.94	46.12

A.4 Western Australia

Station name: Perth (9034)

State: Western Australia

Elevation: 19.0 metres

Latitude: 31.95° South

Longitude: 115.87° East

Annual Rainfall: 869.4mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.27	0.48	0.62	1.52	3.91	6.13	5.60	4.36	2.68	1.76	0.72	0.45

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.91	0.90	0.86	0.74	0.55	0.42	0.41	0.44	0.53	0.64	0.78	0.86

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2.85	2.68	4.34	7.49	13.51	16.81	17.74	16.73	13.61	10.76	6.26	4.08

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
8.04	13.21	18.73	44.51	116.96	177.45	167.65	130.35	77.63	52.78	20.87	13.43

Station name: Esperance (9631)

State: Western Australia

Elevation: 158.0 metres

Latitude: 33.61° South

Longitude: 121.78° East

Annual Rainfall: 497.5mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.73	0.89	0.78	1.24	1.80	2.04	2.13	1.98	1.76	1.36	1.03	0.60

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.85	0.80	0.80	0.68	0.61	0.55	0.52	0.53	0.57	0.67	0.75	0.83

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
4.45	5.51	6.20	9.33	11.80	13.18	14.14	13.90	12.49	10.14	7.22	5.29

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
21.98	24.46	23.70	36.08	54.57	59.46	62.60	58.90	50.61	41.12	30.15	18.29

A.5 New South Wales

Station name: Sydney (66062)

State: New South Wales

Elevation: 39.0 metres

Latitude: 33.86° South

Longitude: 151.20° East

Annual Rainfall: 1217.0mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
3.36	4.14	4.24	4.27	3.92	4.37	3.20	2.67	2.34	2.50	2.76	2.54

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.61	0.56	0.56	0.57	0.56	0.57	0.64	0.66	0.64	0.62	0.62	0.63

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
12.12	12.34	13.48	12.89	13.49	12.77	11.31	10.66	10.81	11.69	11.46	11.51

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
103.34	116.02	130.46	127.08	120.63	130.27	99.24	82.68	70.09	77.51	82.30	78.21

Station name: Richmond (67033)

State: New South Wales

Elevation: 19.0 metres

Latitude: 33.6° South

Longitude: 150.78° East

Annual Rainfall: 810.3mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
3.01	3.74	2.97	2.38	1.90	1.88	1.16	1.48	1.33	2.07	2.54	2.22

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.65	0.61	0.64	0.68	0.68	0.68	0.75	0.74	0.74	0.69	0.65	0.68

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
9.69	9.78	9.79	8.60	8.85	8.43	6.72	7.05	6.76	8.60	9.16	8.61

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
82.17	92.96	79.76	62.98	52.62	48.79	31.61	40.30	34.77	56.45	67.06	59.66

Station name: Chichester (61151)

State: New South Wales

Elevation: 194.0 metres

Latitude: 32.24° South

Longitude: 151.68° East

Annual Rainfall: 1313.5mm

Monthly Rainfall Statistics

AVERAGE_DAILY_RAINFALL(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
5.84	6.38	5.40	3.26	3.09	3.54	1.66	2.01	2.13	3.01	3.19	4.06

DAILY_DRY_PROBABILITIES(%)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.60	0.54	0.58	0.67	0.66	0.63	0.71	0.73	0.71	0.67	0.65	0.65

MEAN_WET_DAYS

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
12.17	12.81	12.81	9.62	10.48	10.59	8.90	8.26	8.47	10.09	10.26	10.81

MEAN_MONTHLY_RAIN(mm)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
177.98	177.02	164.62	96.18	94.07	102.35	50.49	61.18	62.75	91.56	93.66	123.52

APPENDIX B

Improved Rainfall Model Validation (Master Sites)

B.1 Adelaide, South Australia (BOM# 23034)

B.1.1 Simulated and Observed Storm Event Characteristics

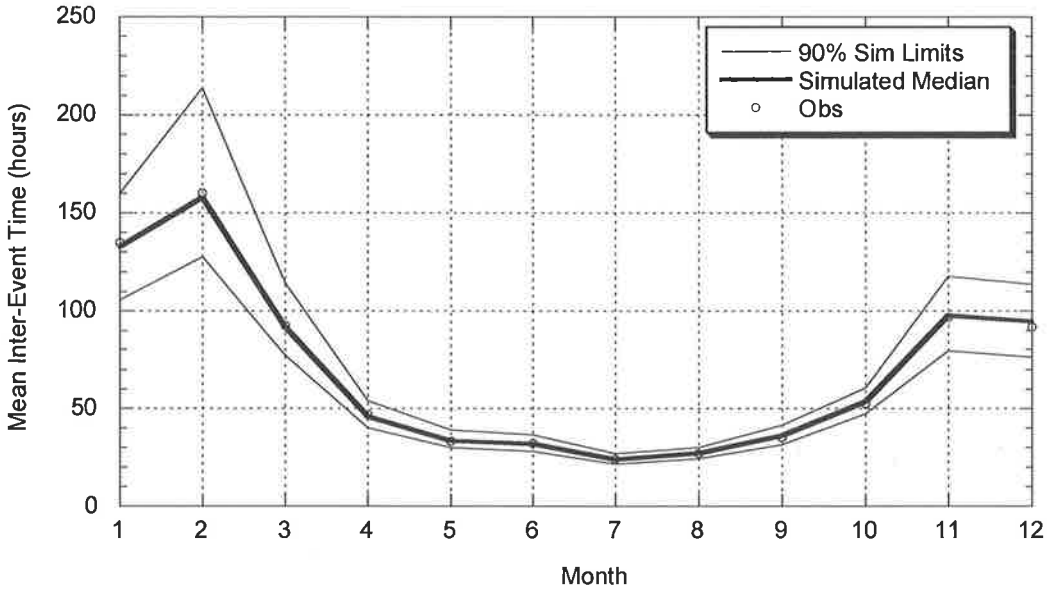

Figure B.1.1: Comparison between Observed and Simulated Mean Inter-Event Times (Adelaide)

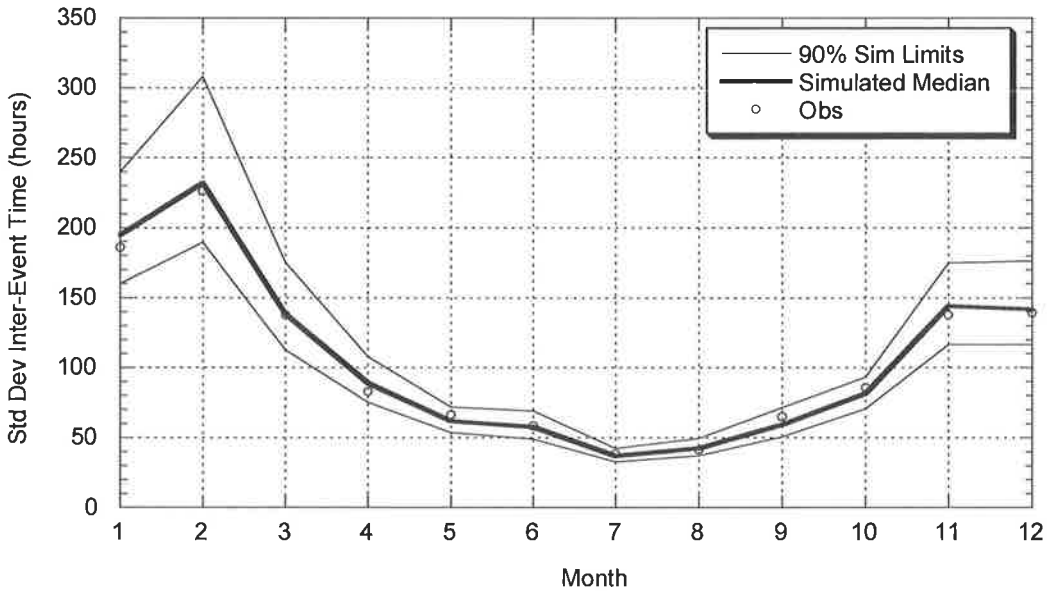

Figure B.1.2: Comparison between Observed and Simulated Standard Deviation of Inter-Event Times (Adelaide)

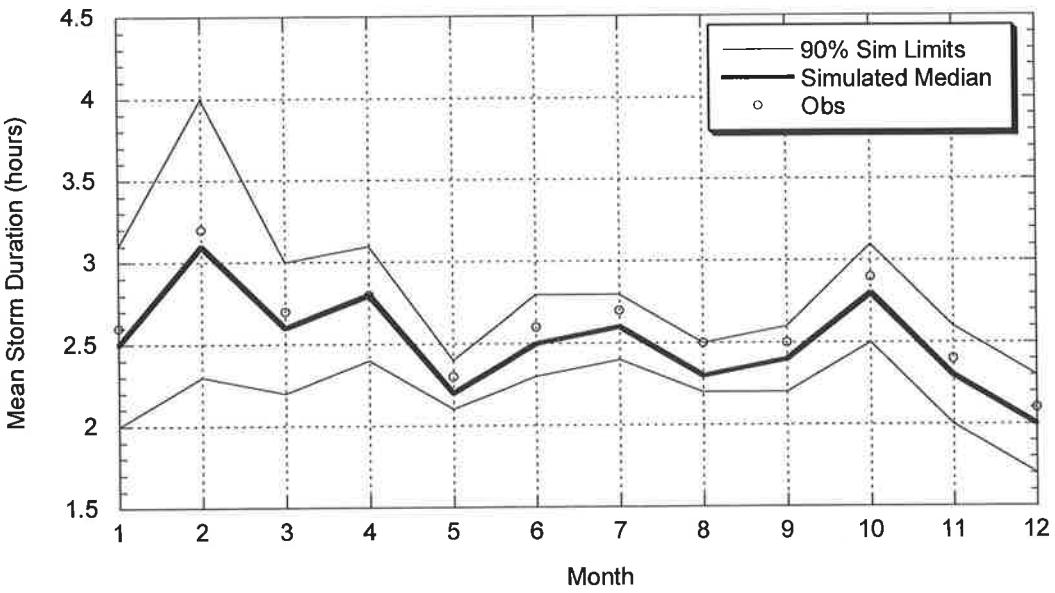

Figure B.1.3: Comparison between Observed and Simulated Mean of Event Storm Durations (Adelaide)

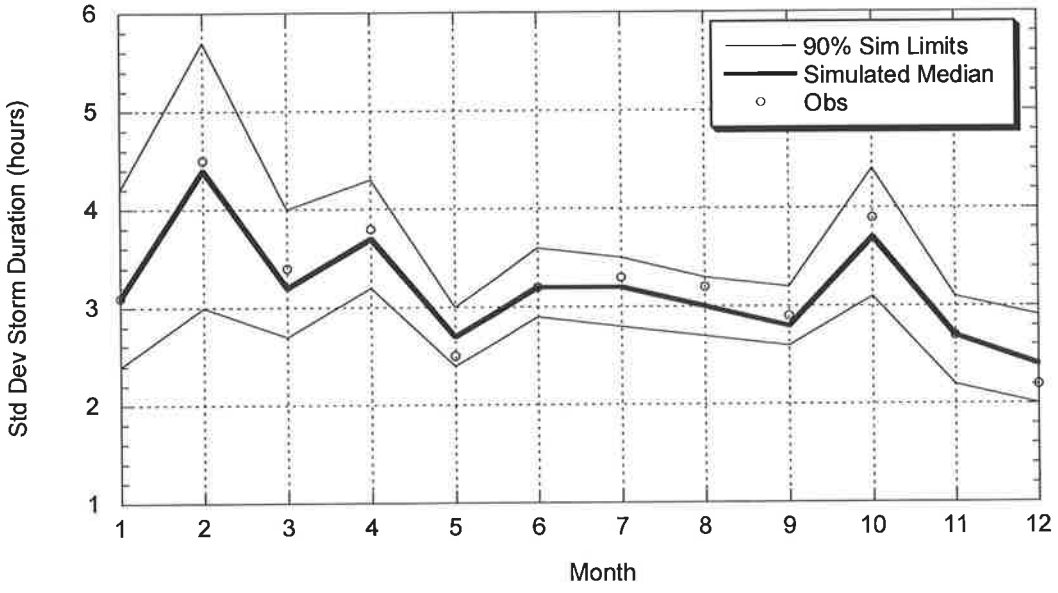

Figure B.1.4: Comparison between Observed and Simulated Standard Deviation of Event Storm Durations (Adelaide)

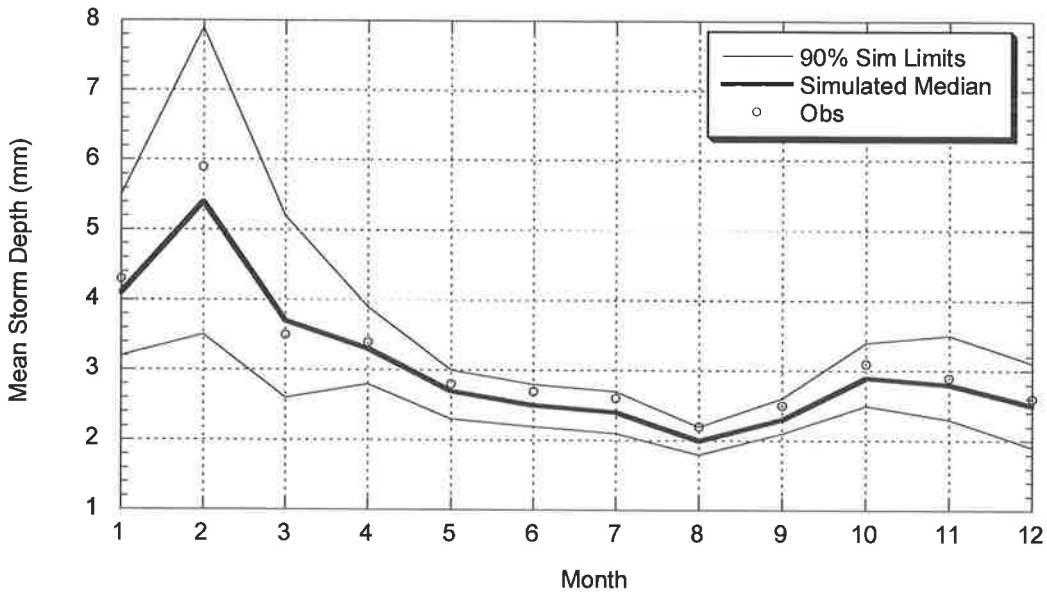

Figure B.1.5: Comparison between Observed and Simulated Average of Event Depths (Adelaide)

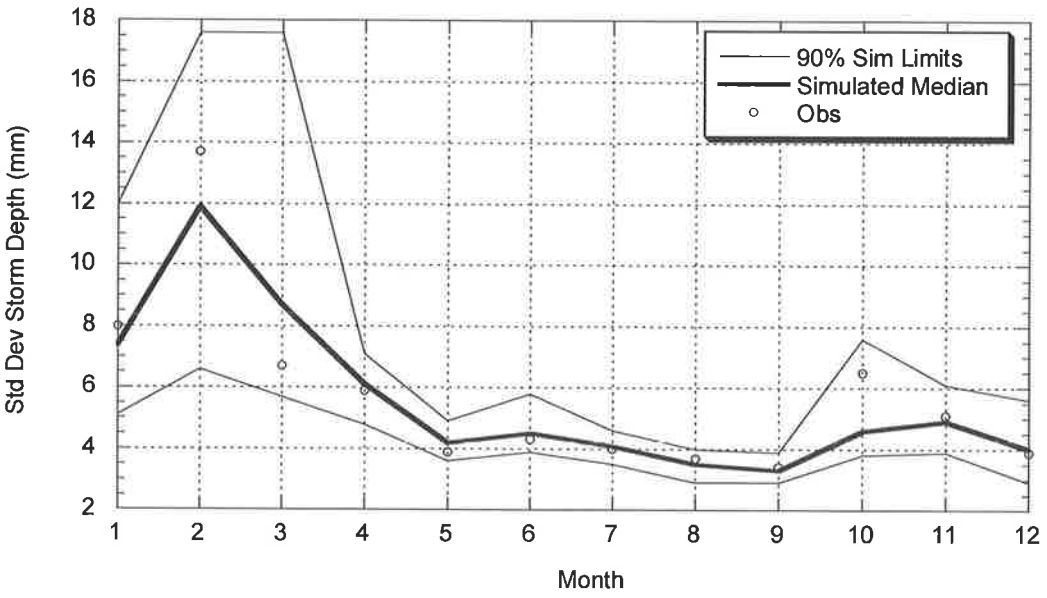

Figure B.1.6: Comparison between Observed and Simulated Standard Deviation of Event Depths (Adelaide)

B.1.2 Simulated and Observed Daily Statistics

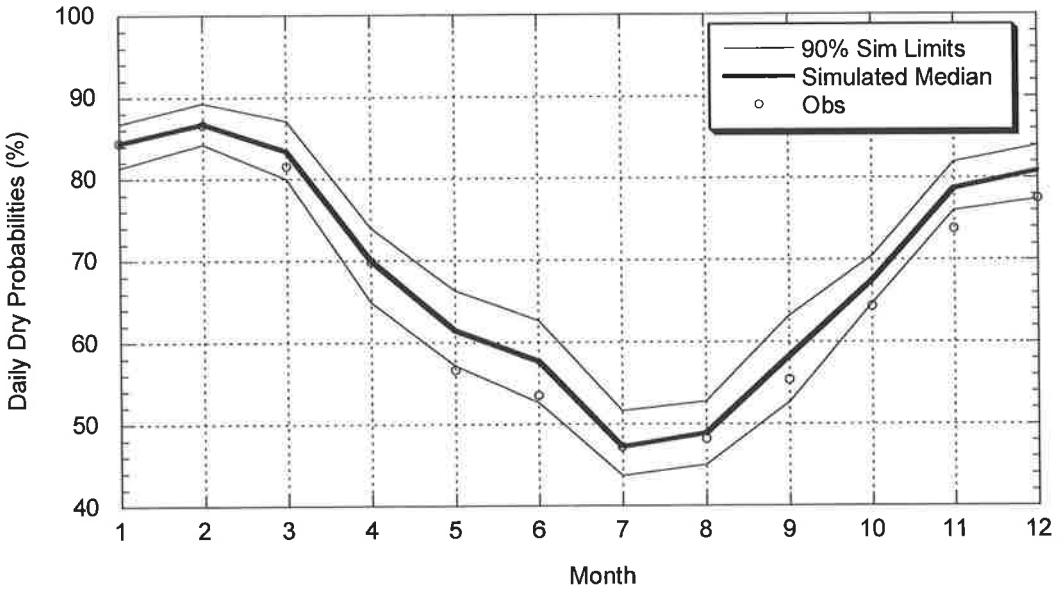

Figure B.1.7: Comparison between Observed and Simulated Daily Dry Probabilities (Adelaide)

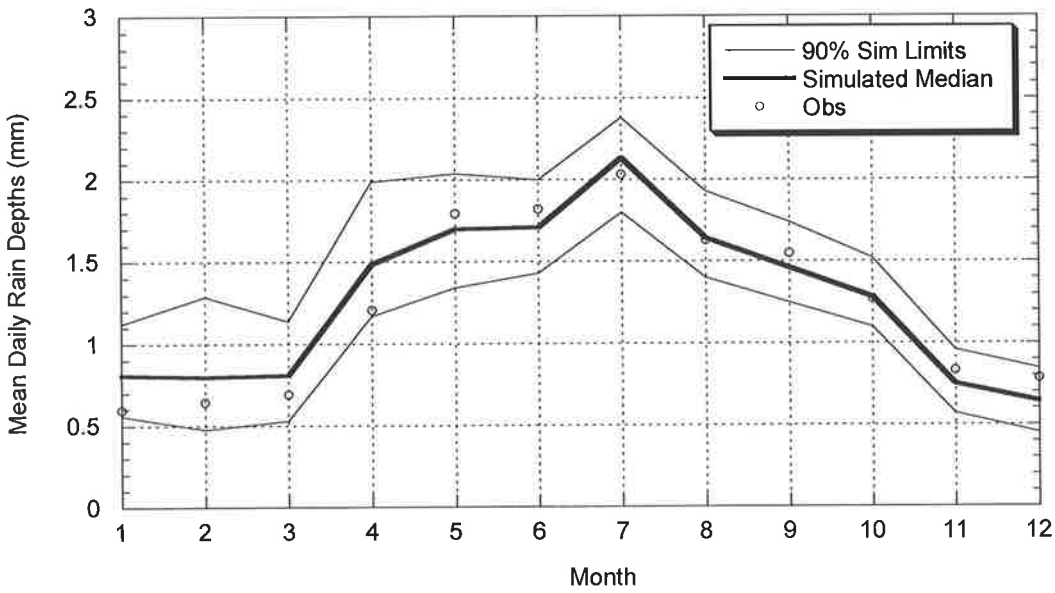

Figure B.1.8: Comparison between Observed and Simulated Daily Mean Depth (Adelaide)

B.1.3 Simulated and Observed Annual and Monthly Rainfall

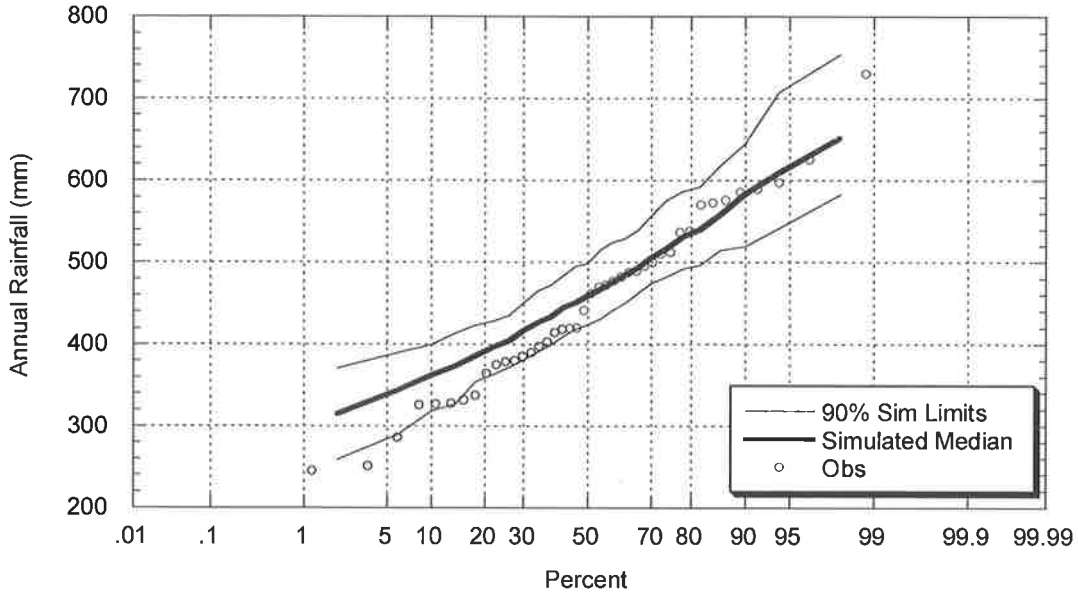

Figure B.1.9: Comparison between Observed and Simulated Annual Rainfall (Adelaide)

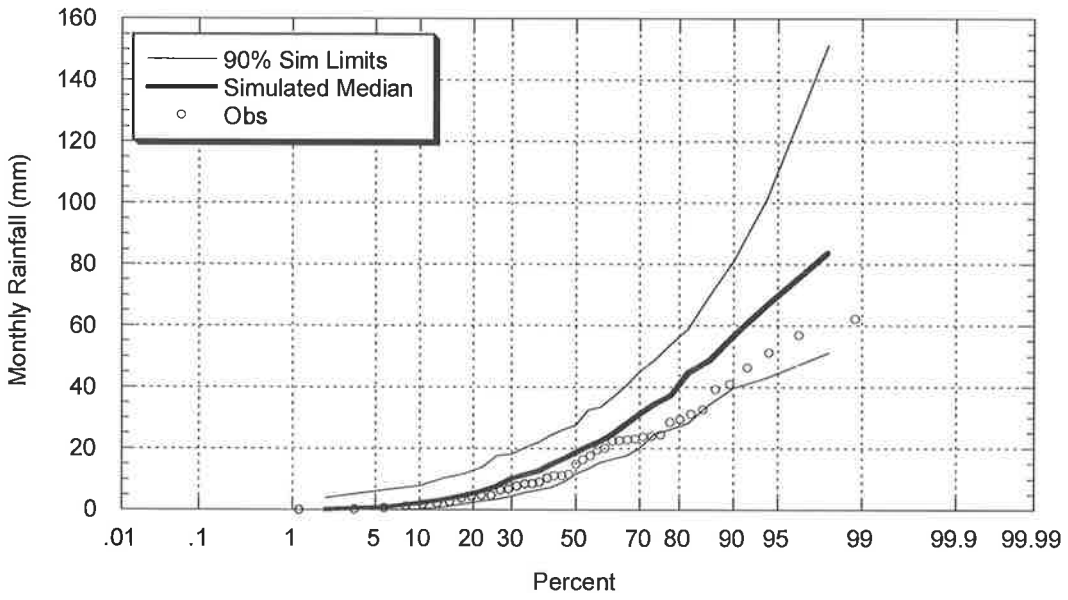

Figure B.1.10: Comparison between Observed and Simulated January Rainfall (Adelaide)

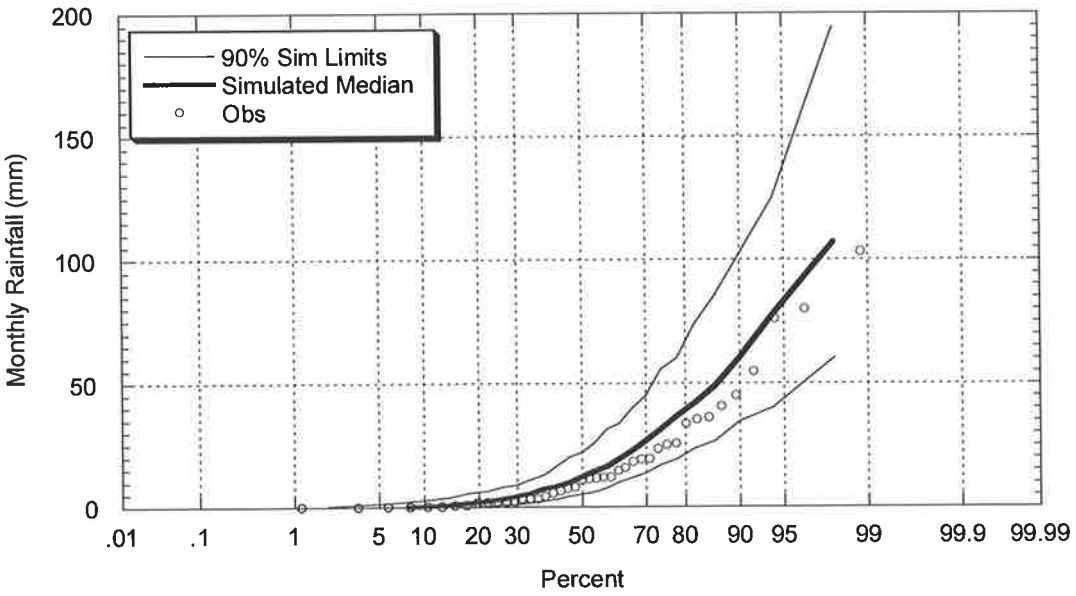

Figure B.1.11: Comparison between Observed and Simulated February Rainfall (Adelaide)

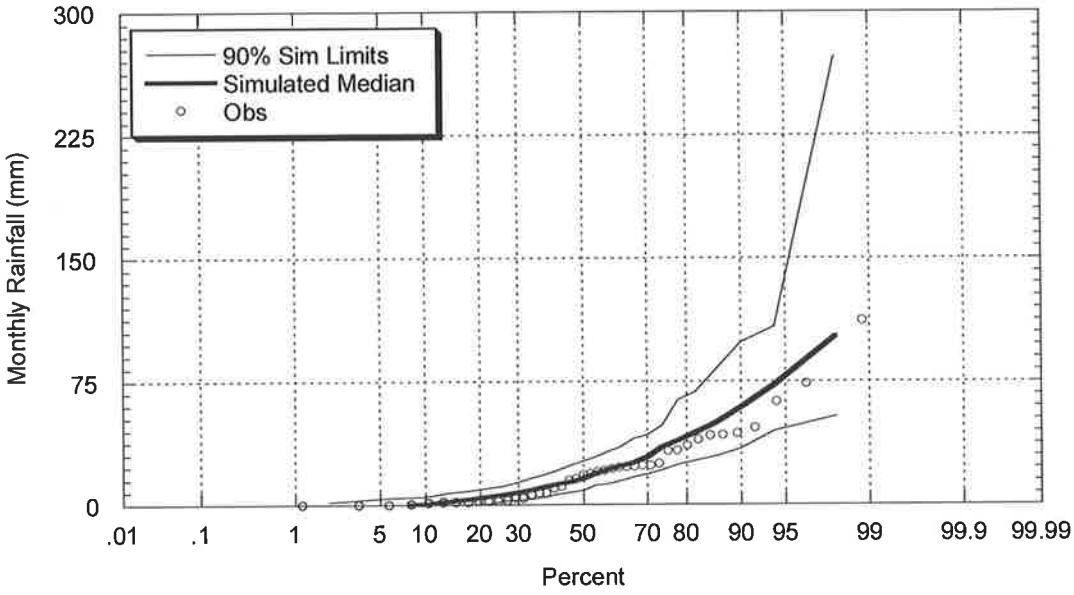

Figure B.1.12: Comparison between Observed and Simulated March Rainfall (Adelaide)

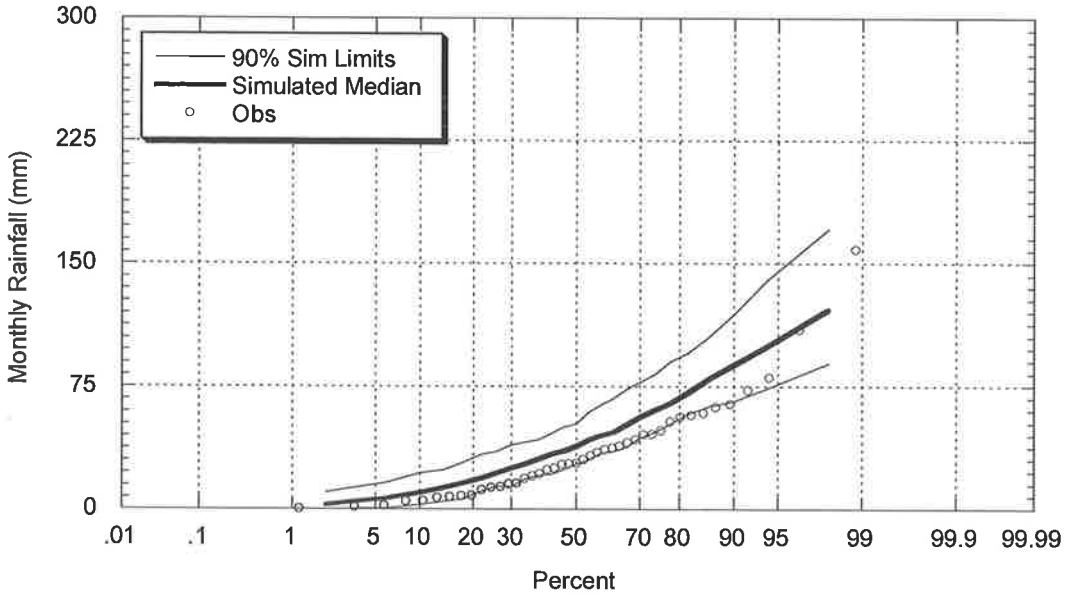

Figure B.1.13: Comparison between Observed and Simulated April Rainfall (Adelaide)

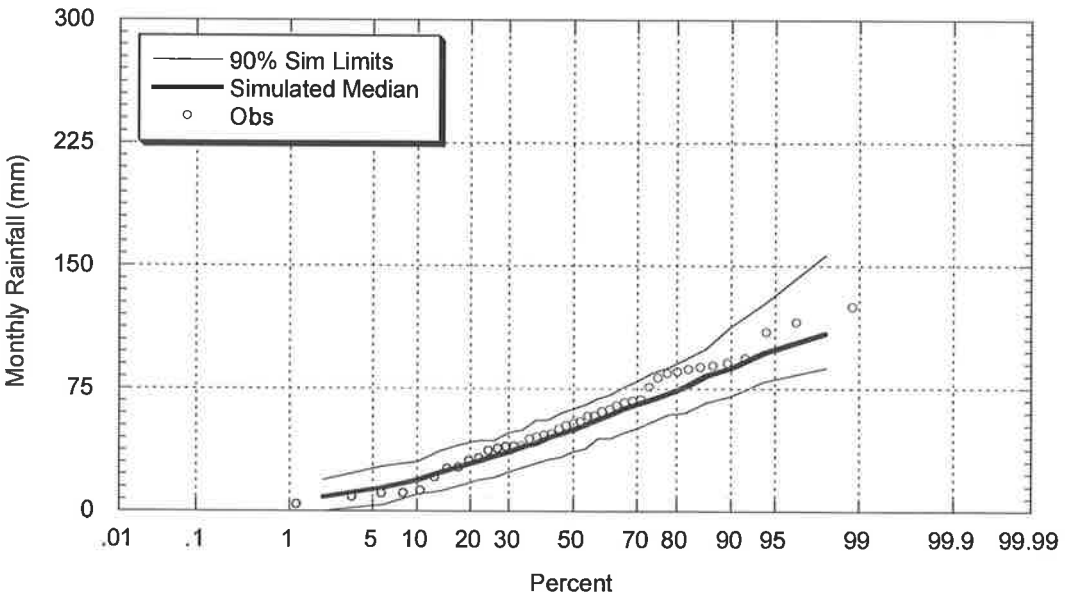

Figure B.1.14: Comparison between Observed and Simulated May Rainfall (Adelaide)

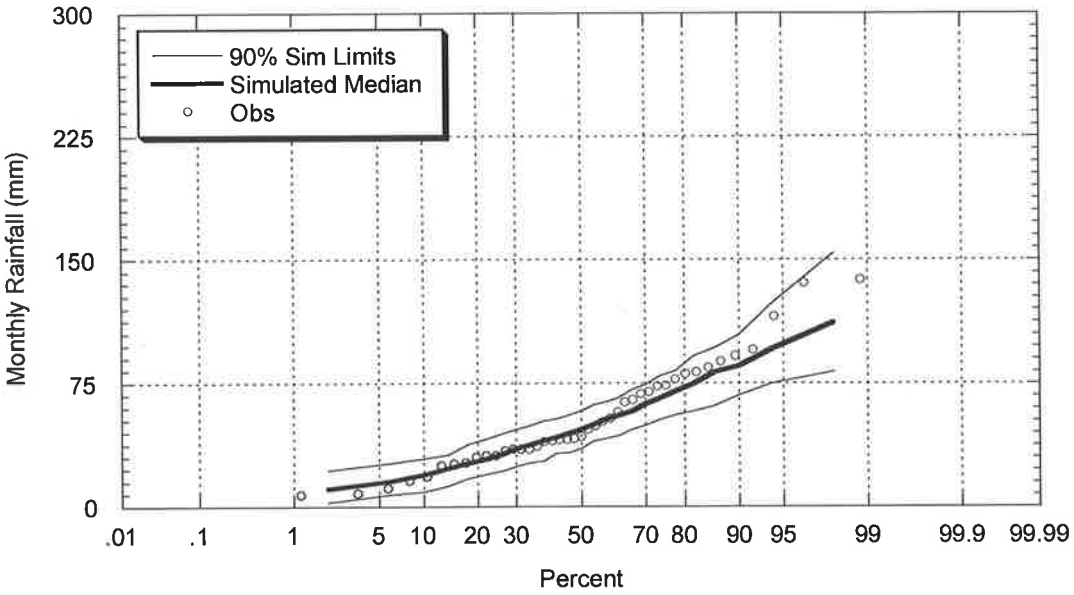

Figure B.1.15: Comparison between Observed and Simulated June Rainfall (Adelaide)

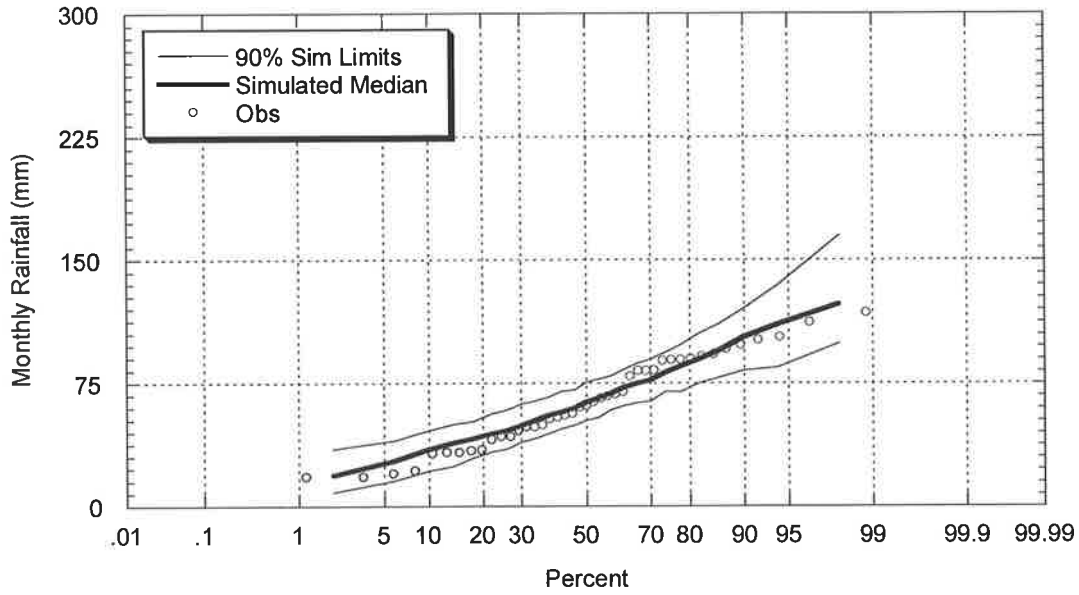

Figure B.1.16: Comparison between Observed and Simulated July Rainfall (Adelaide)

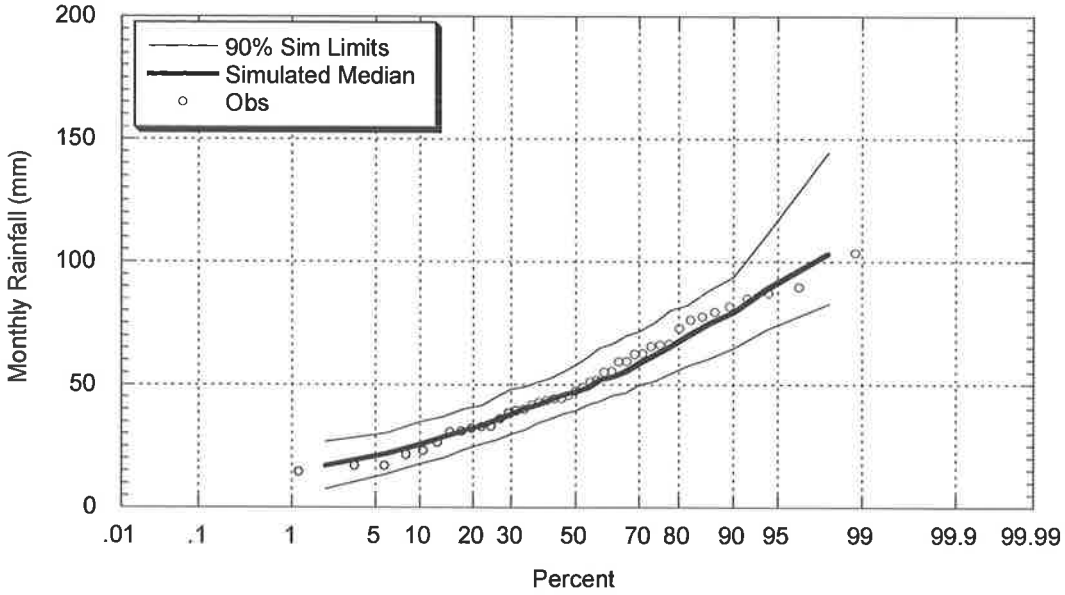

Figure B.1.17: Comparison between Observed and Simulated August Rainfall (Adelaide)

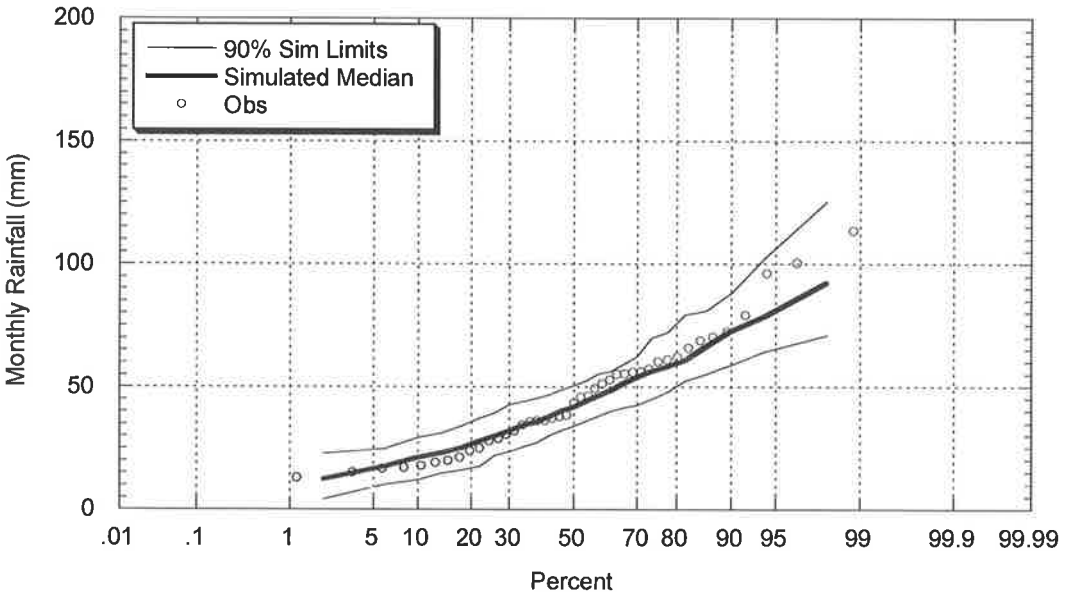

Figure B.1.18: Comparison between Observed and Simulated September Rainfall (Adelaide)

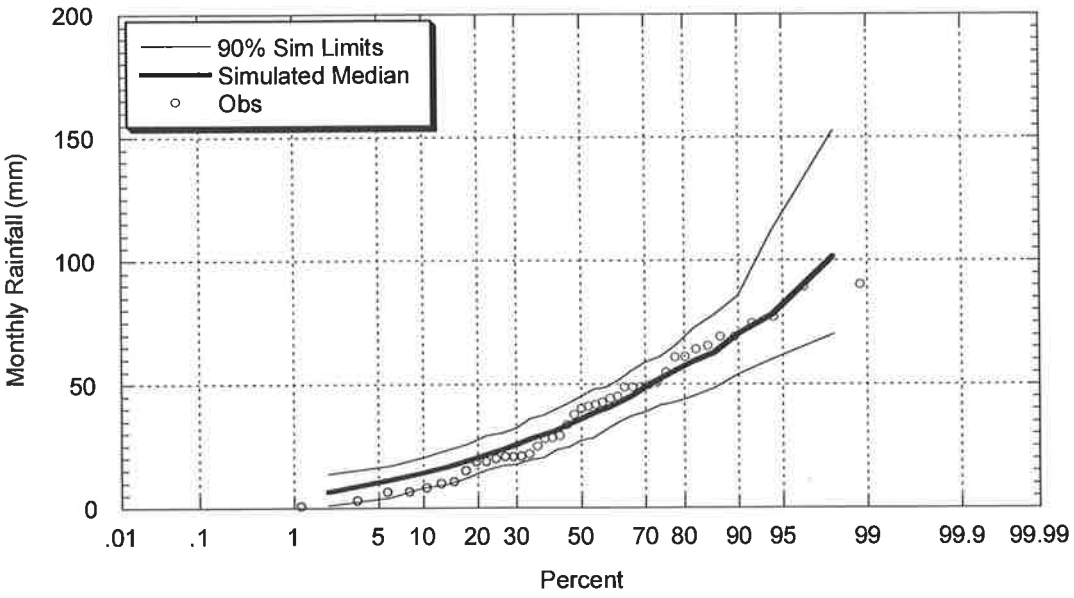

Figure B.1.19: Comparison between Observed and Simulated October Rainfall (Adelaide)

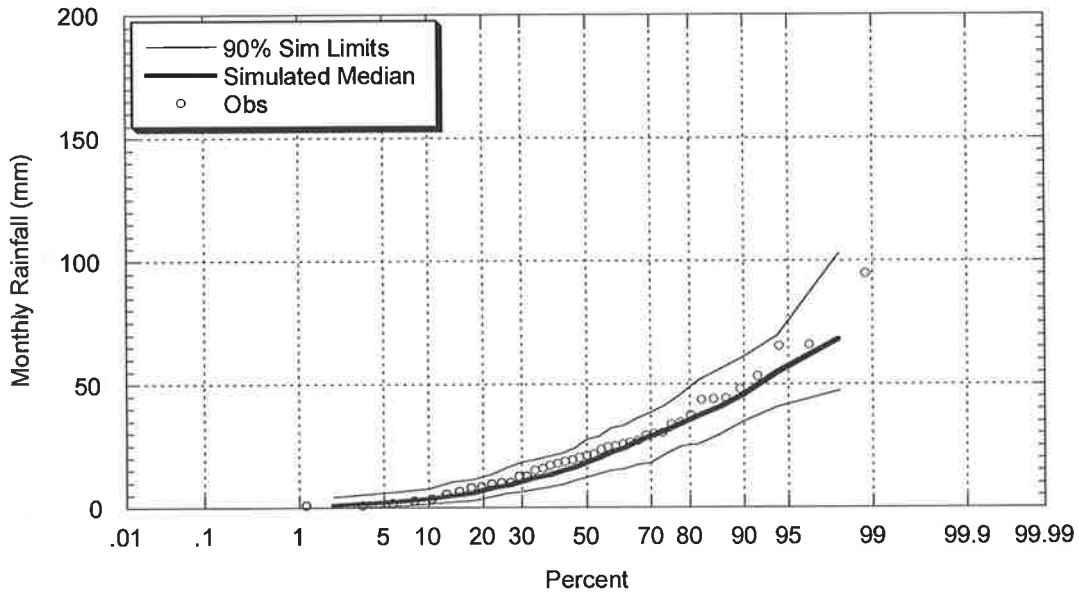

Figure B.1.20: Comparison between Observed and Simulated November Rainfall (Adelaide)

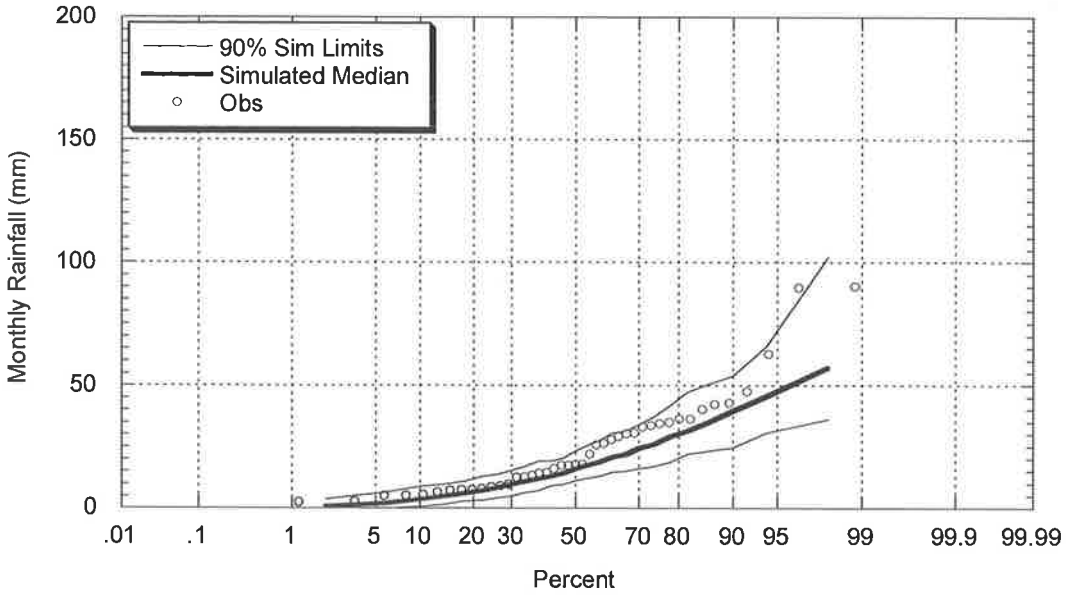

Figure B.1.21: Comparison between Observed and Simulated December Rainfall (Adelaide)

B.1.4 Simulated and Observed Annual Intensity – Frequency - Duration Curves

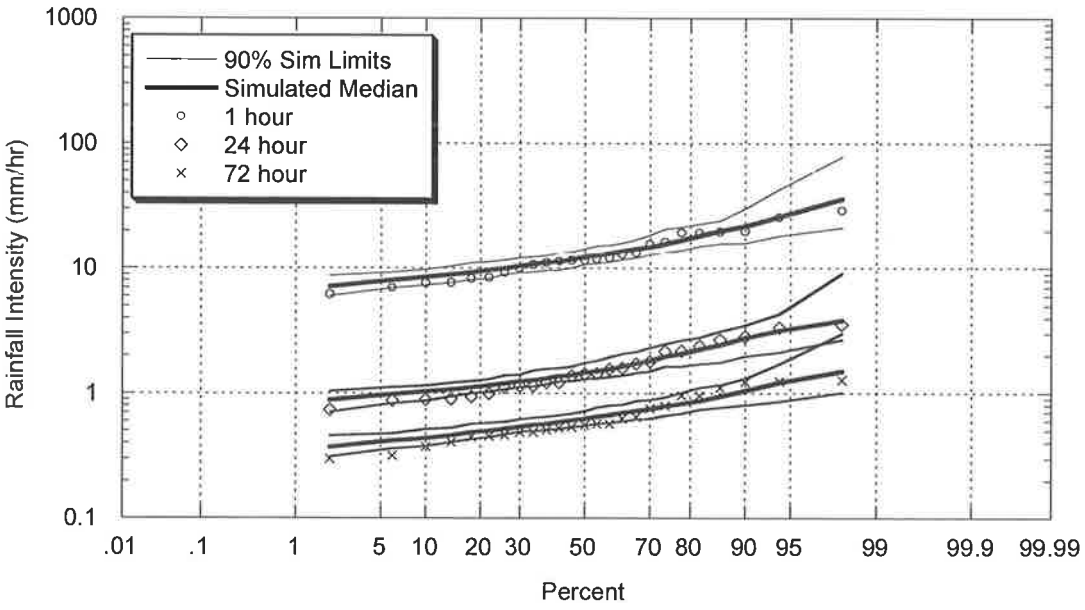

Figure B.1.22: Comparison between Observed and Simulated Annual Intensity Frequency Duration Relationship (Adelaide)

B.2 Brisbane, Queensland (BOM# 40214)

B.2.1 Simulated and Observed Storm Event Characteristics

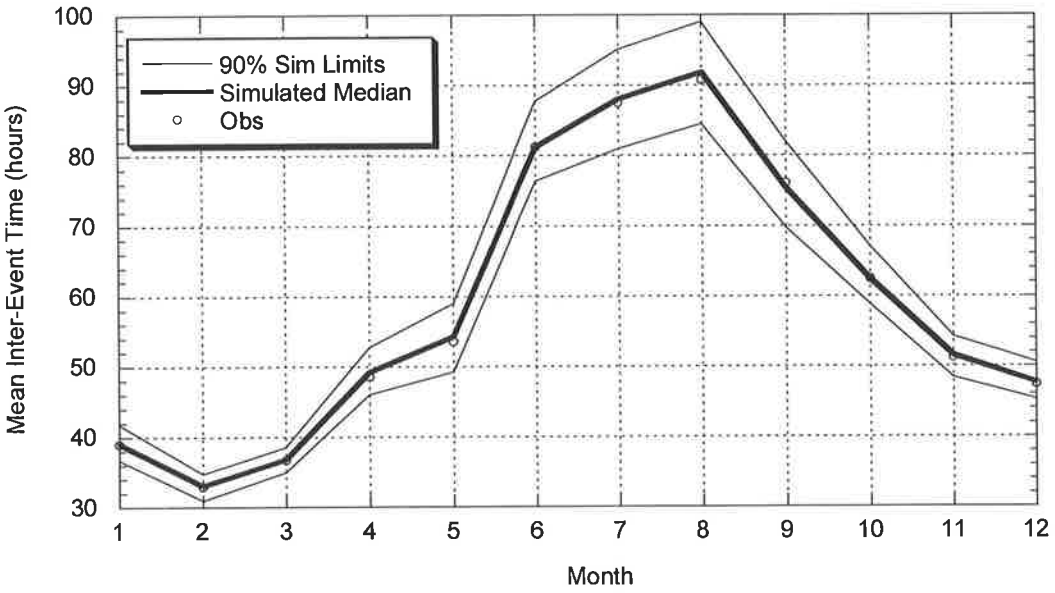

Figure B.2.1: Comparison between Observed and Simulated Mean Inter-Event Times (Brisbane)

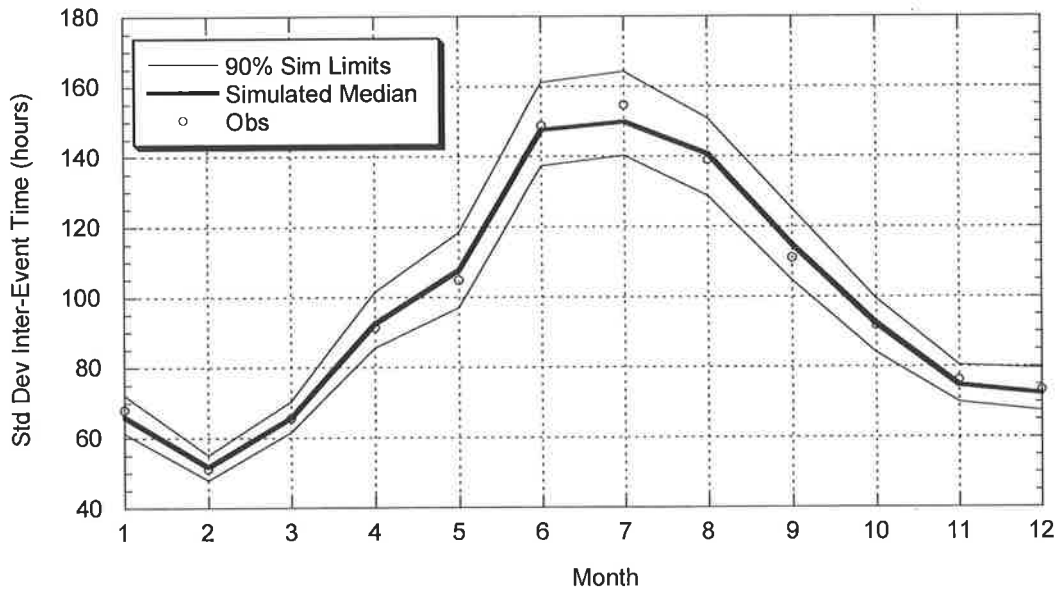

Figure B.2.2: Comparison between Observed and Simulated Standard Deviation of Inter-Event Times (Brisbane)

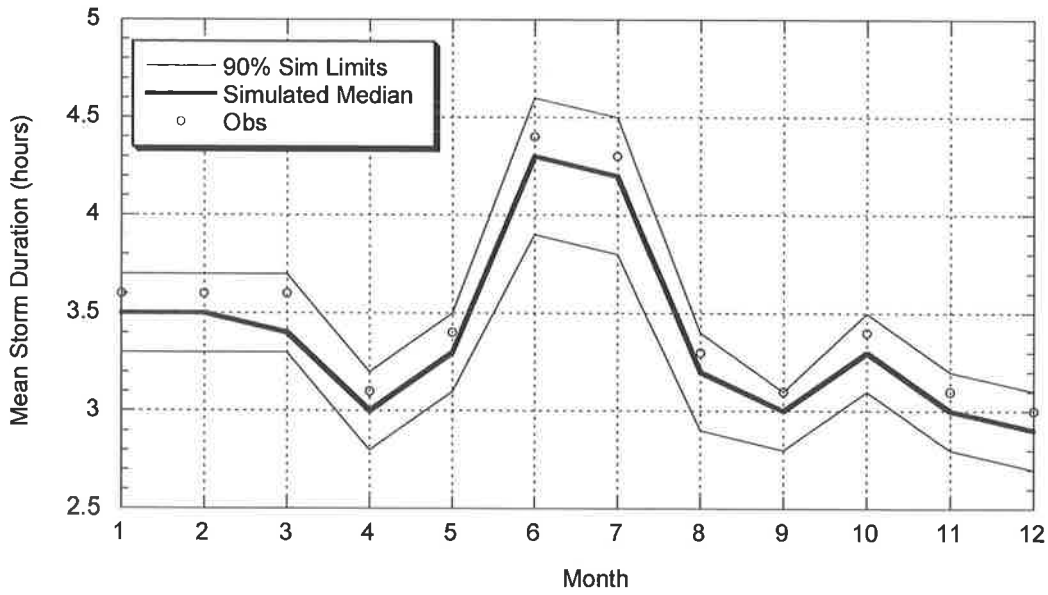

Figure B.2.3: Comparison between Observed and Simulated Mean of Event Storm Durations (Brisbane)

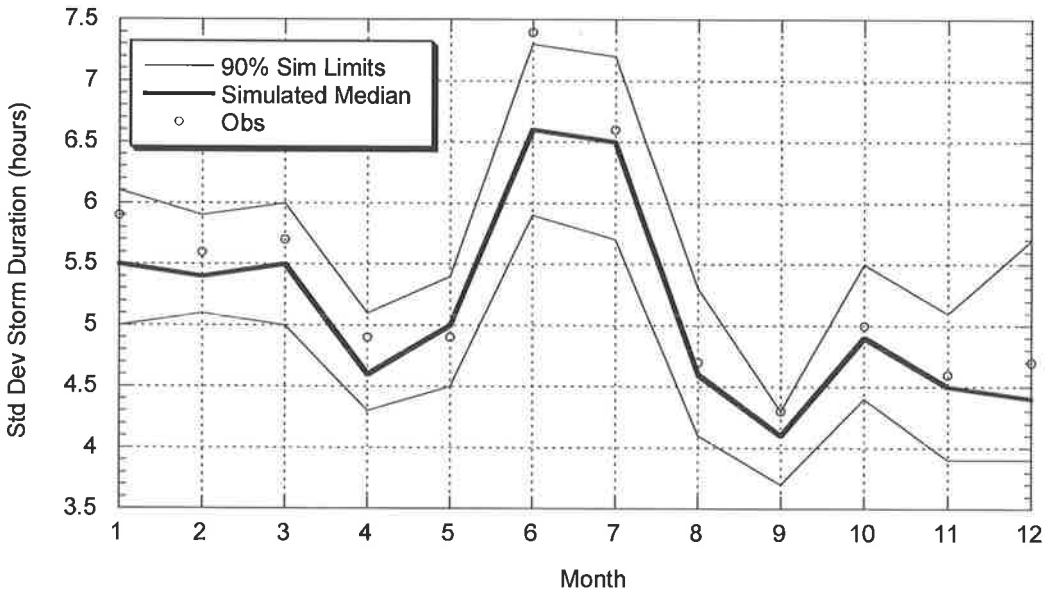

Figure B.2.4: Comparison between Observed and Simulated Standard Deviation of Event Storm Durations (Brisbane)

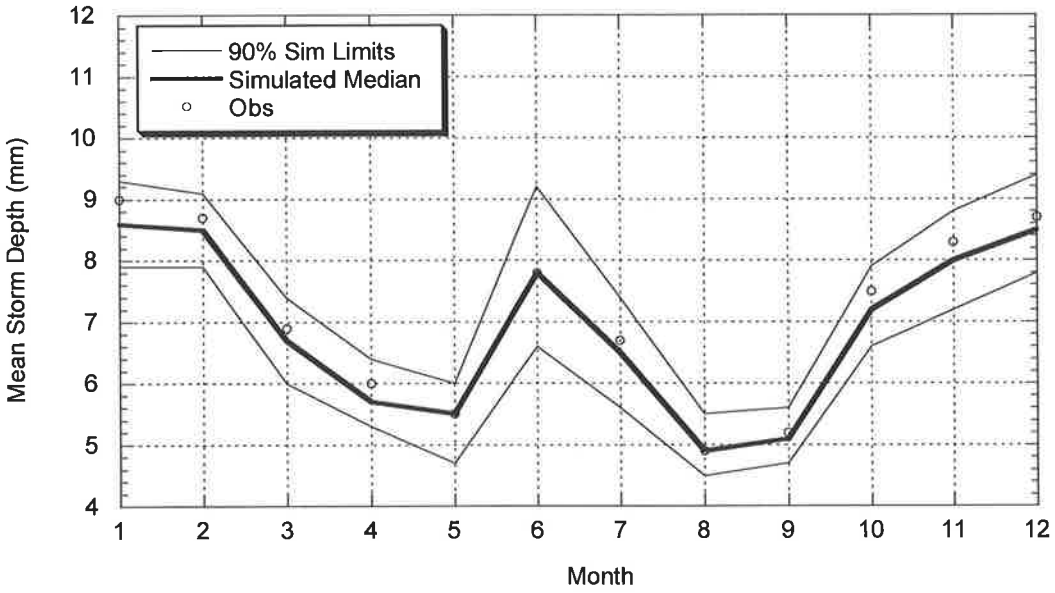

Figure B.2.5: Comparison between Observed and Simulated Average of Event Depths (Brisbane)

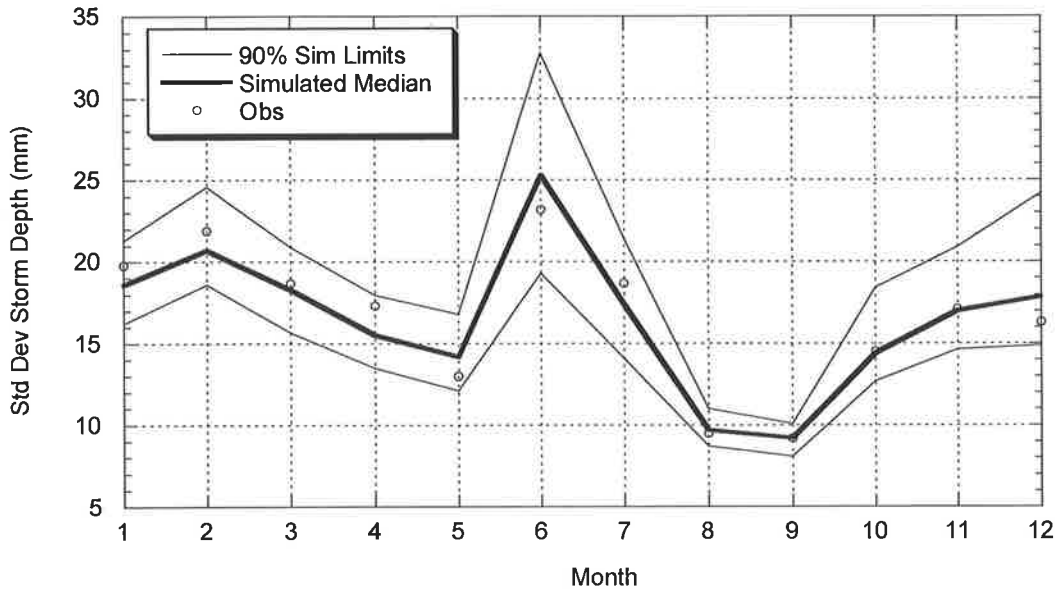

Figure B.2.6: Comparison between Observed and Simulated Standard Deviation of Event Depths (Brisbane)

B.2.2 Simulated and Observed Daily Statistics

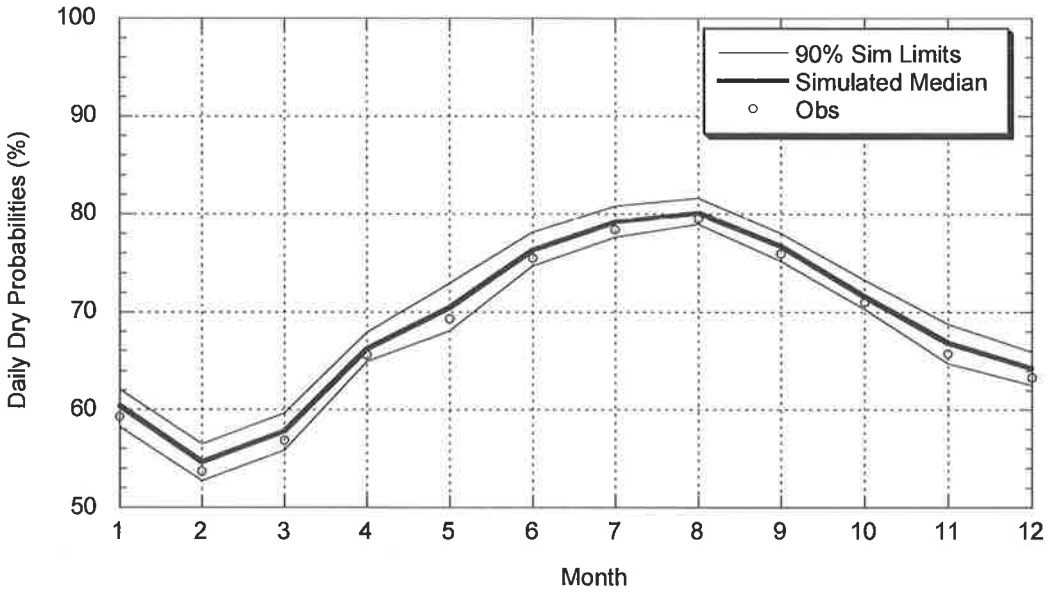

Figure B.2.7: Comparison between Observed and Simulated Daily Dry Probabilities (Brisbane)

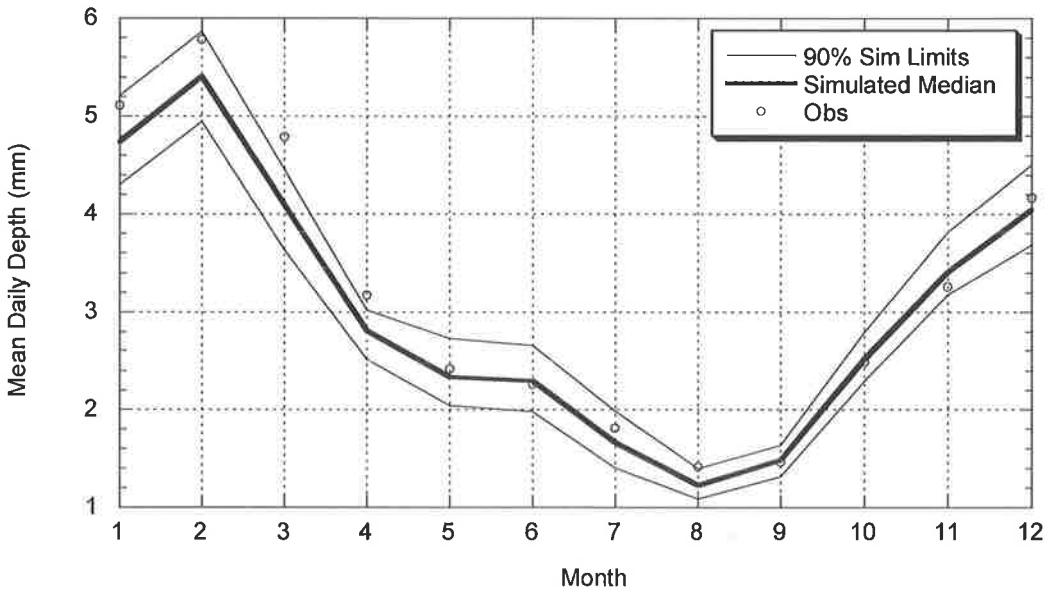

Figure B.2.8: Comparison between Observed and Simulated Daily Mean Depth (Brisbane)

B.2.3 Simulated and Observed Annual and Monthly Rainfall

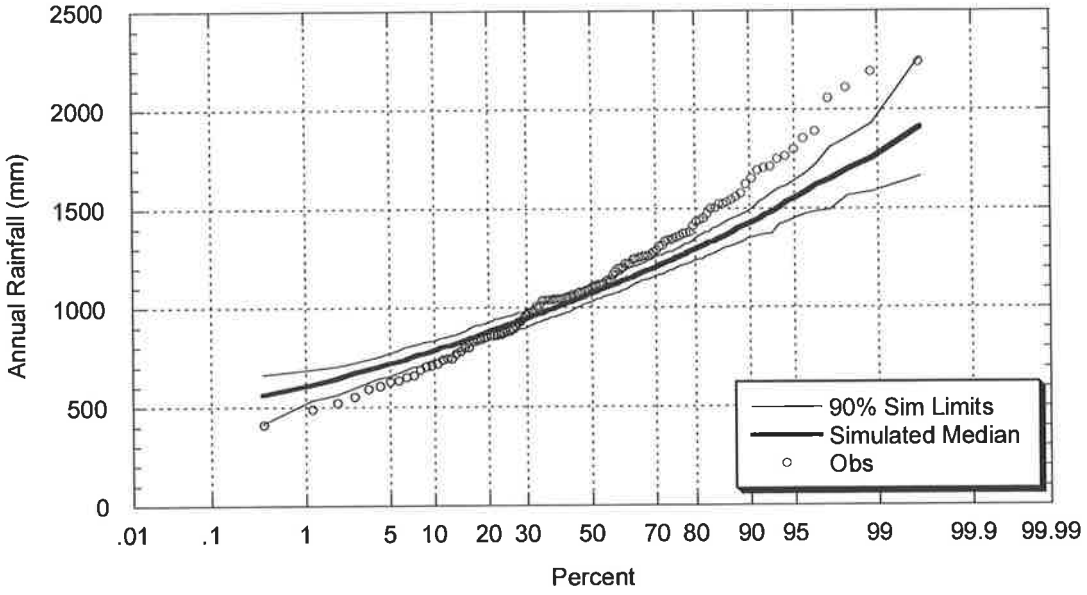

Figure B.2.9: Comparison between Observed and Simulated Annual Rainfall (Brisbane)

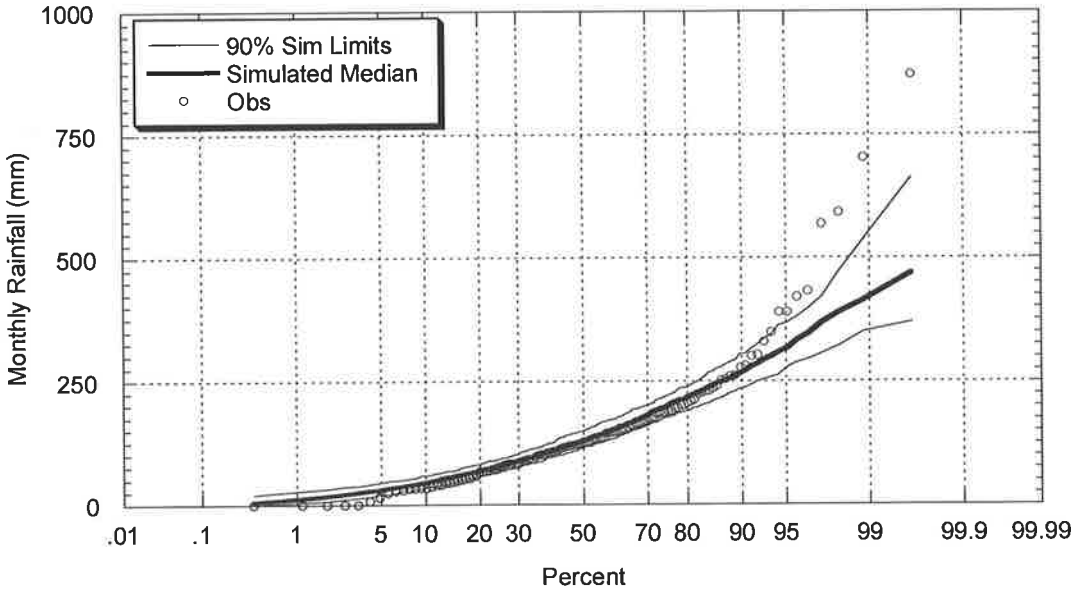

Figure B.2.10: Comparison between Observed and Simulated January Rainfall (Brisbane)

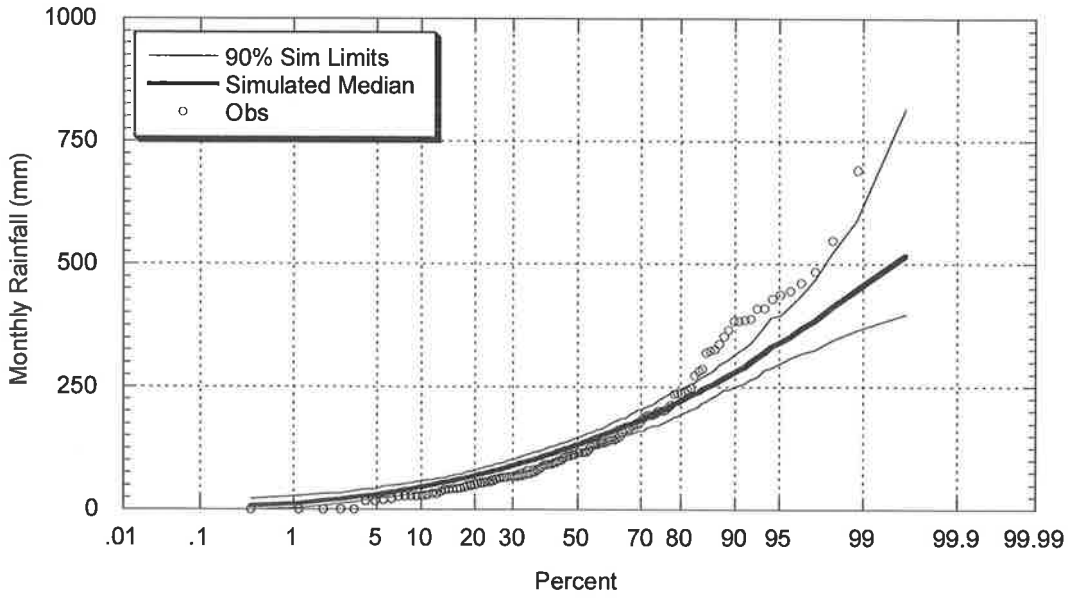

Figure B.2.11: Comparison between Observed and Simulated February Rainfall (Brisbane)

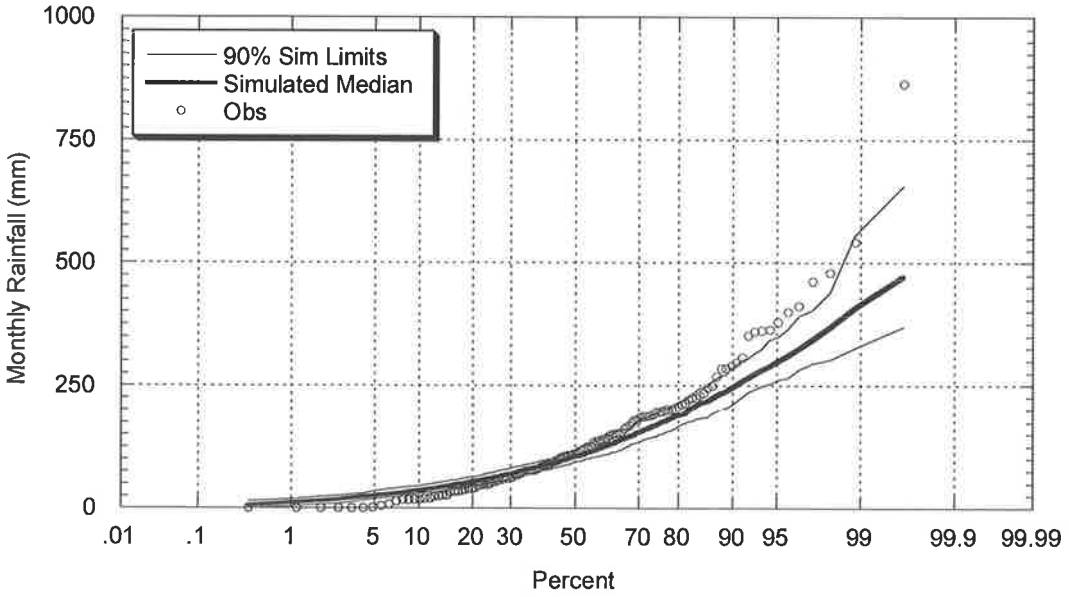

Figure B.2.12: Comparison between Observed and Simulated March Rainfall (Brisbane)

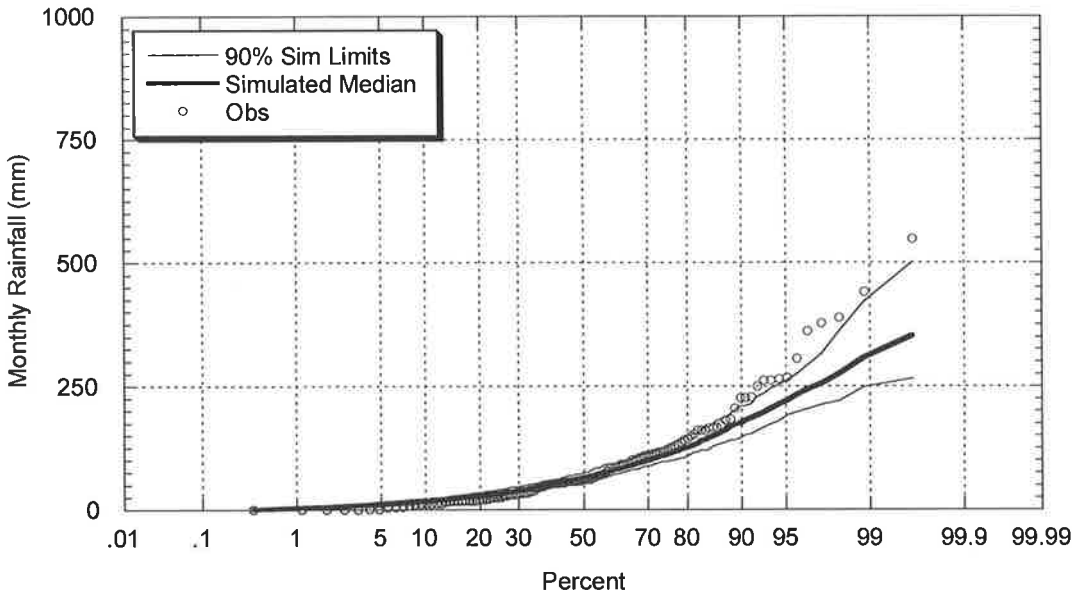

Figure B.2.13: Comparison between Observed and Simulated April Rainfall (Brisbane)

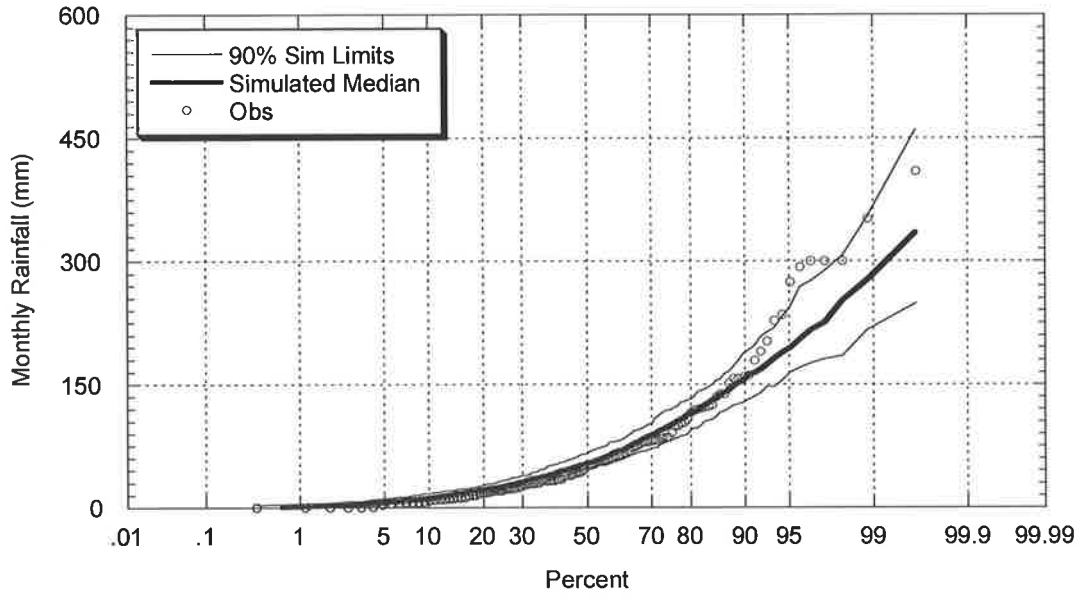

Figure B.2.14: Comparison between Observed and Simulated May Rainfall (Brisbane)

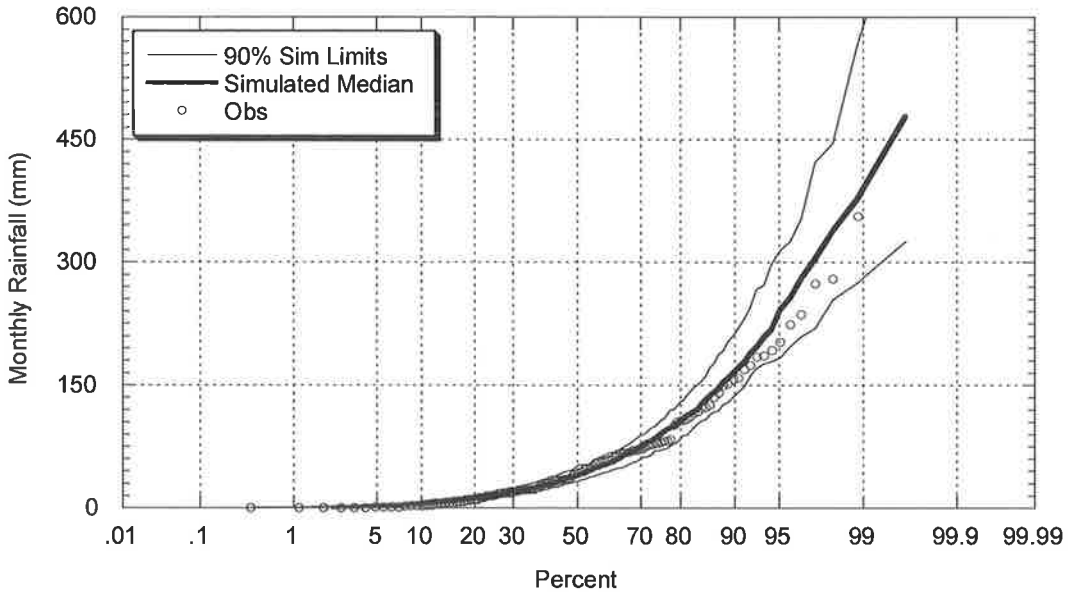

Figure B.2.15: Comparison between Observed and Simulated June Rainfall (Brisbane)

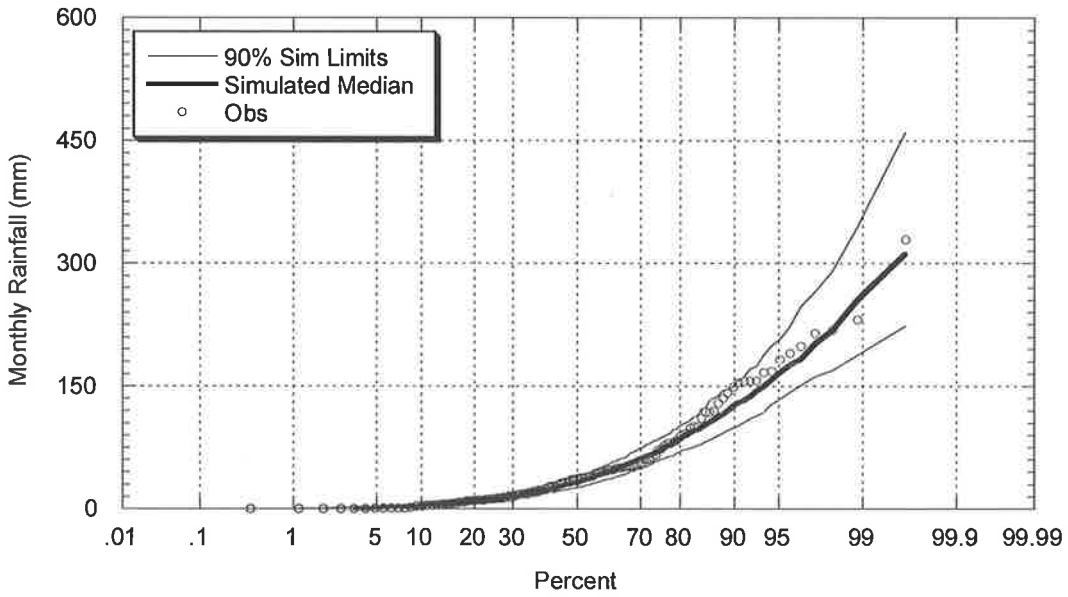

Figure B.2.16: Comparison between Observed and Simulated July Rainfall (Brisbane)

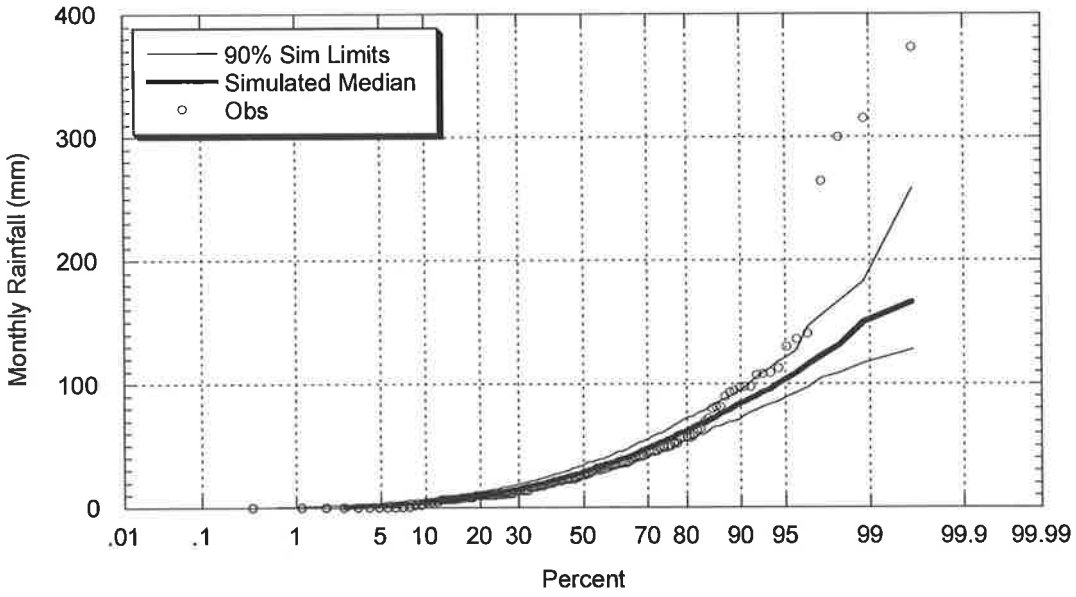

Figure B.2.17: Comparison between Observed and Simulated August Rainfall (Brisbane)

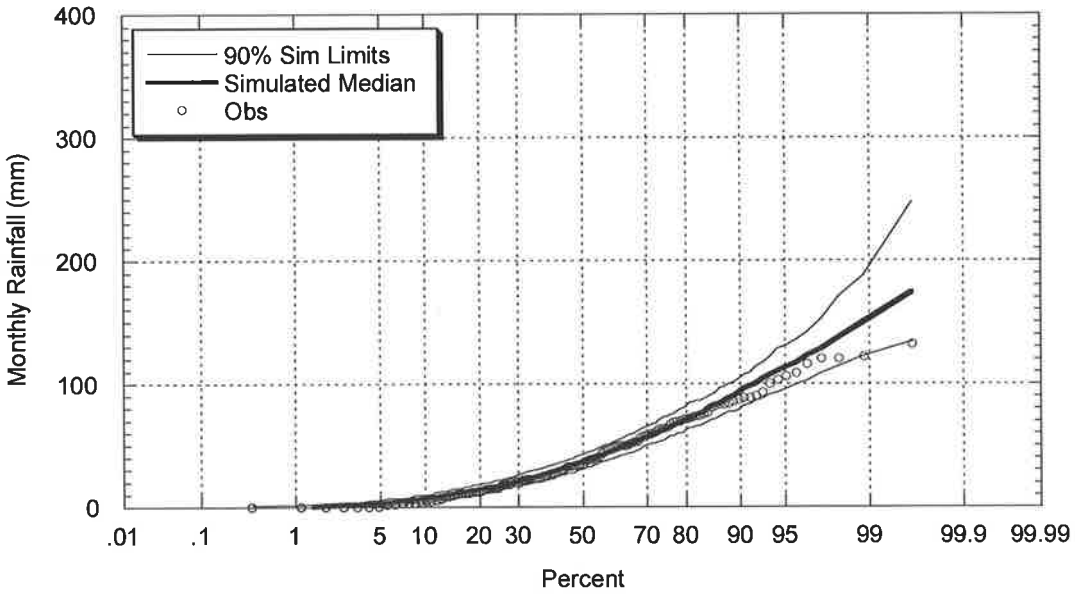

Figure B.2.18: Comparison between Observed and Simulated September Rainfall (Brisbane)

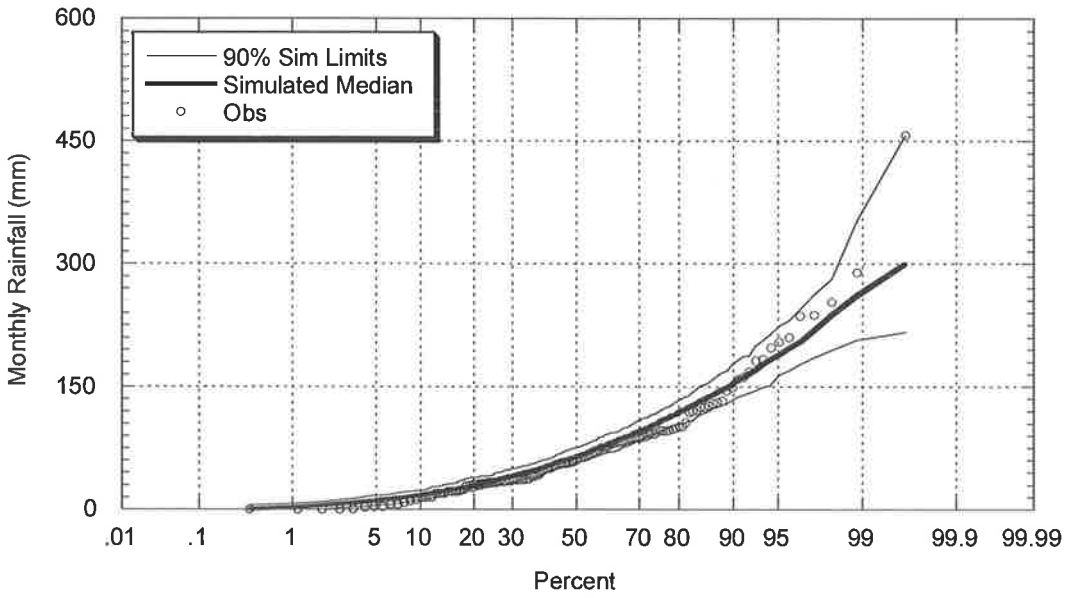

Figure B.2.19: Comparison between Observed and Simulated October Rainfall (Brisbane)

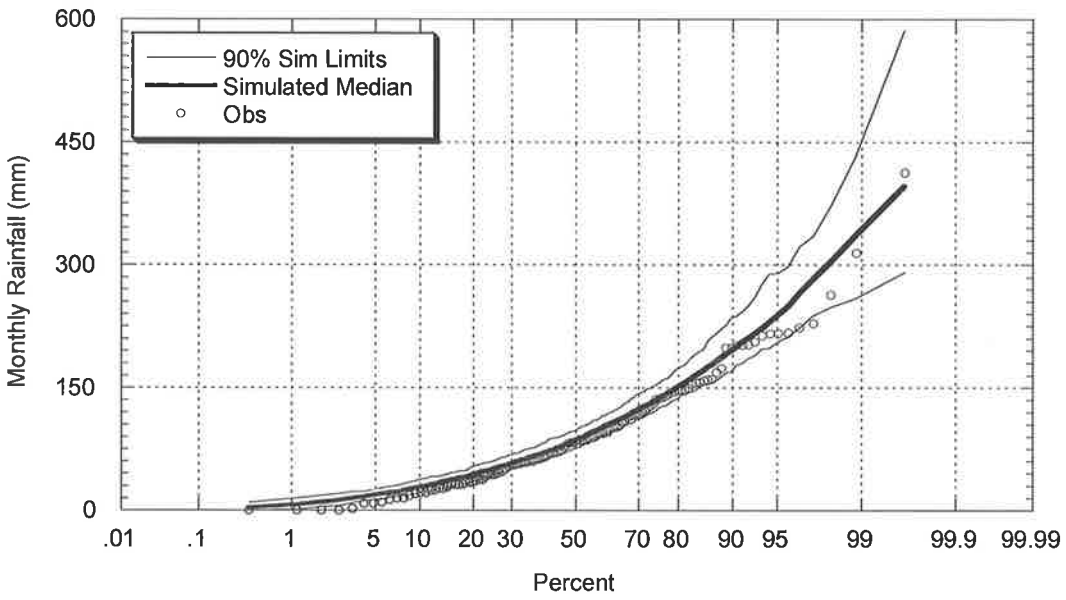

Figure B.2.20: Comparison between Observed and Simulated November Rainfall (Brisbane)

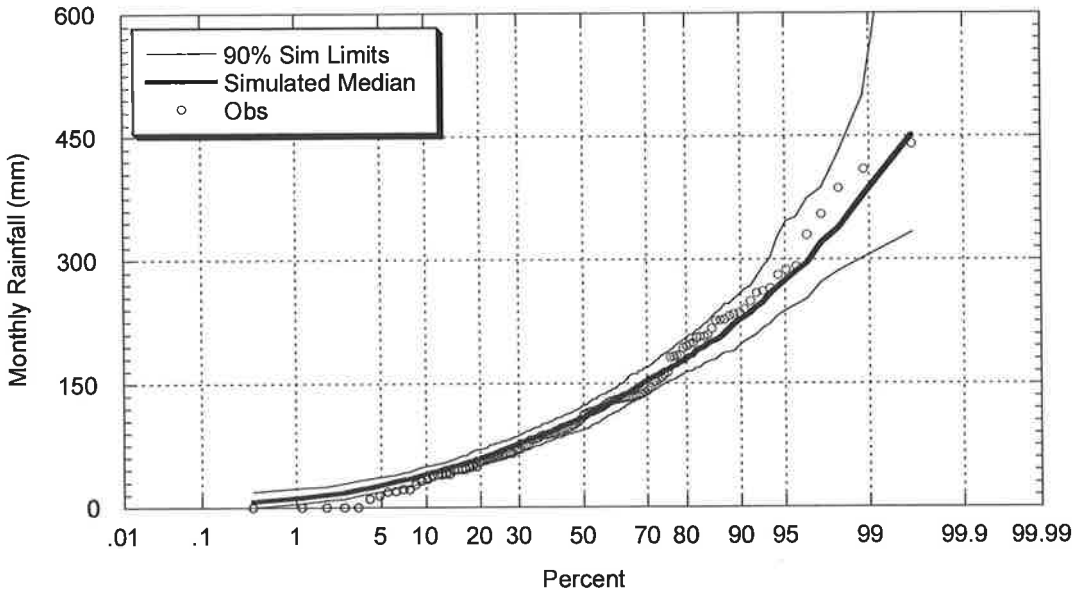

Figure B.2.21: Comparison between Observed and Simulated December Rainfall (Brisbane)

B.2.4 Simulated and Observed Annual Intensity – Frequency - Duration Curves

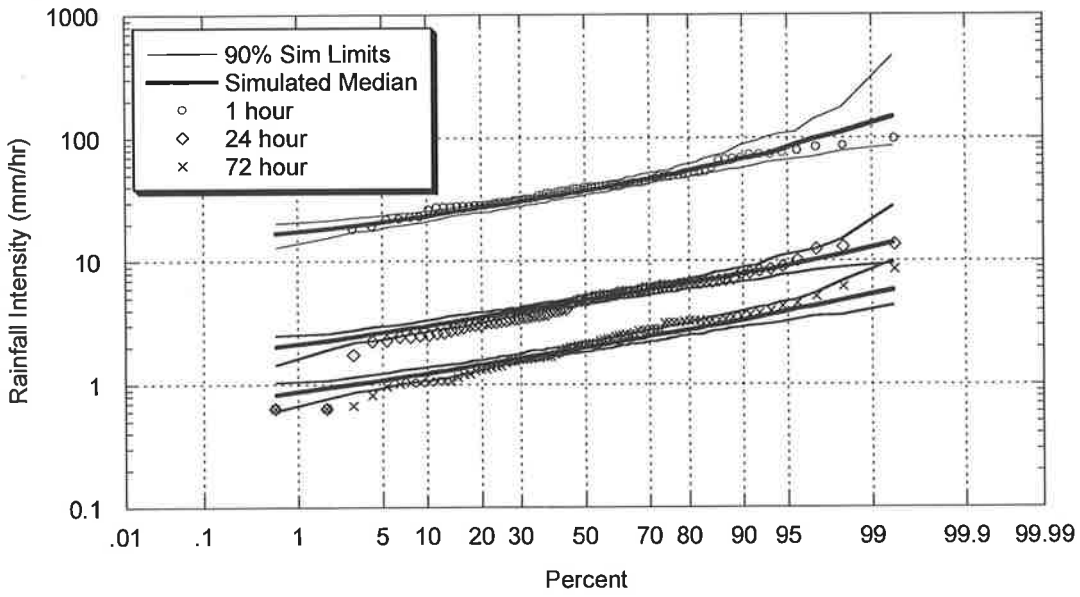

Figure B.2.22: Comparison between Observed and Simulated Annual Intensity Frequency Duration Relationship (Brisbane)

B.3 Melbourne, Victoria (BOM# 86071)

B.3.1 Simulated and Observed Storm Event Characteristics

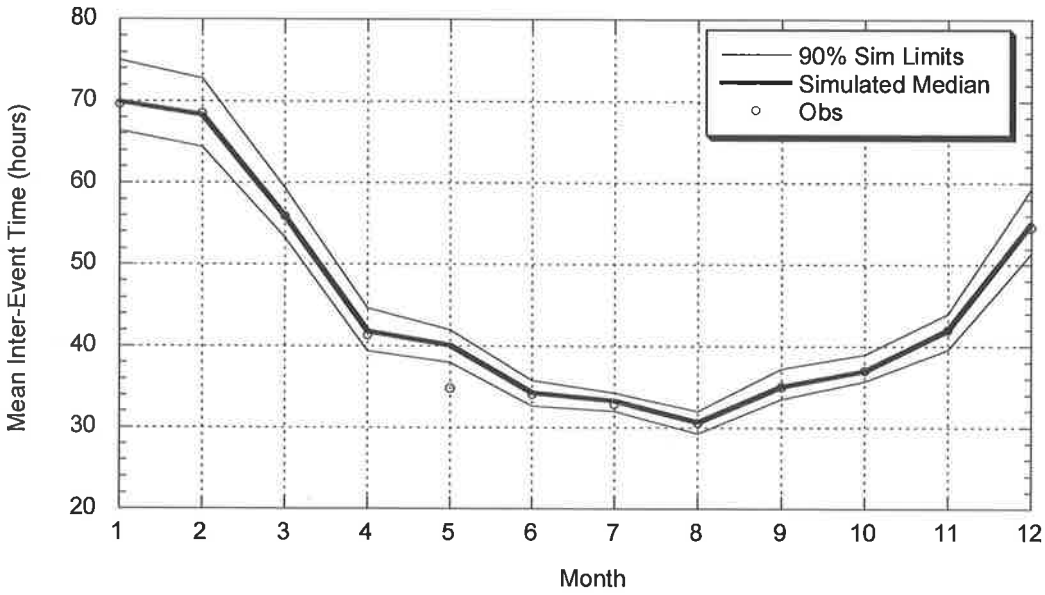

Figure B.3.1: Comparison between Observed and Simulated Mean Inter-Event Times (Melbourne)

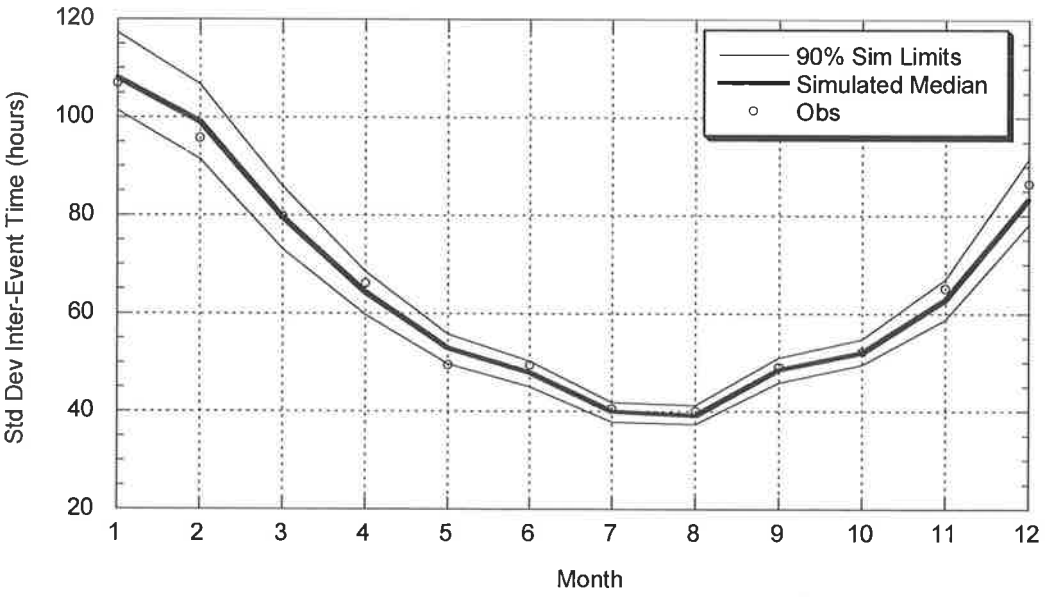

Figure B.3.2: Comparison between Observed and Simulated Standard Deviation of Inter-Event Times (Melbourne)

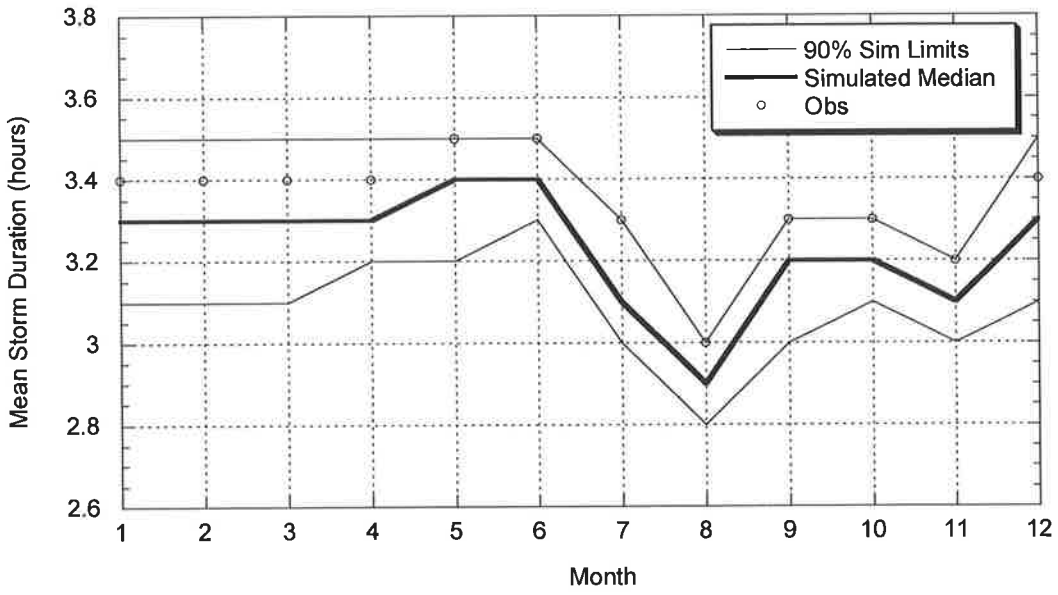

Figure B.3.3: Comparison between Observed and Simulated Mean of Event Storm Durations (Melbourne)

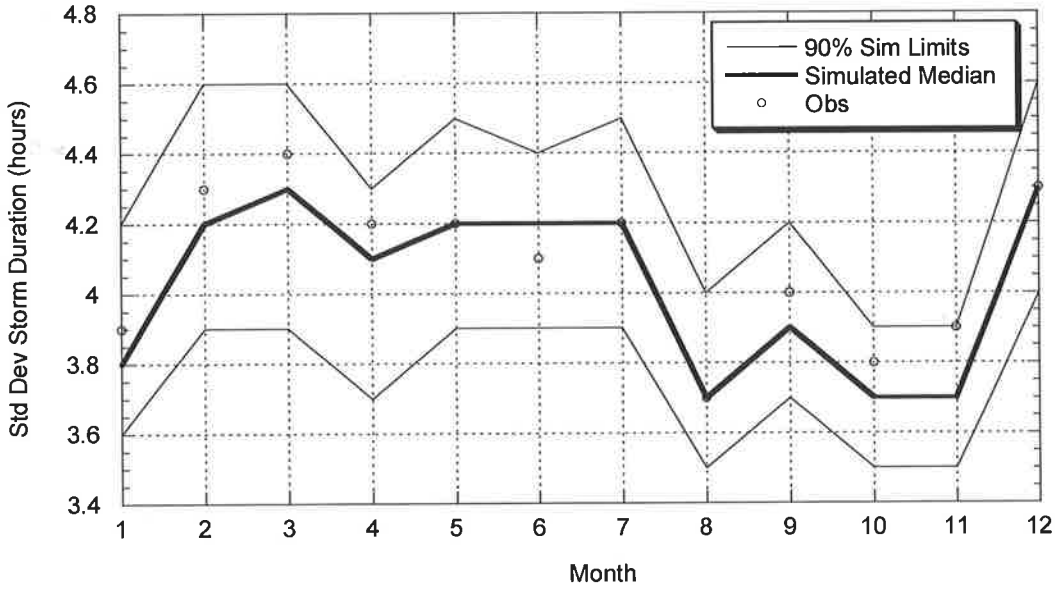

Figure B.3.4: Comparison between Observed and Simulated Standard Deviation of Event Storm Durations (Melbourne)

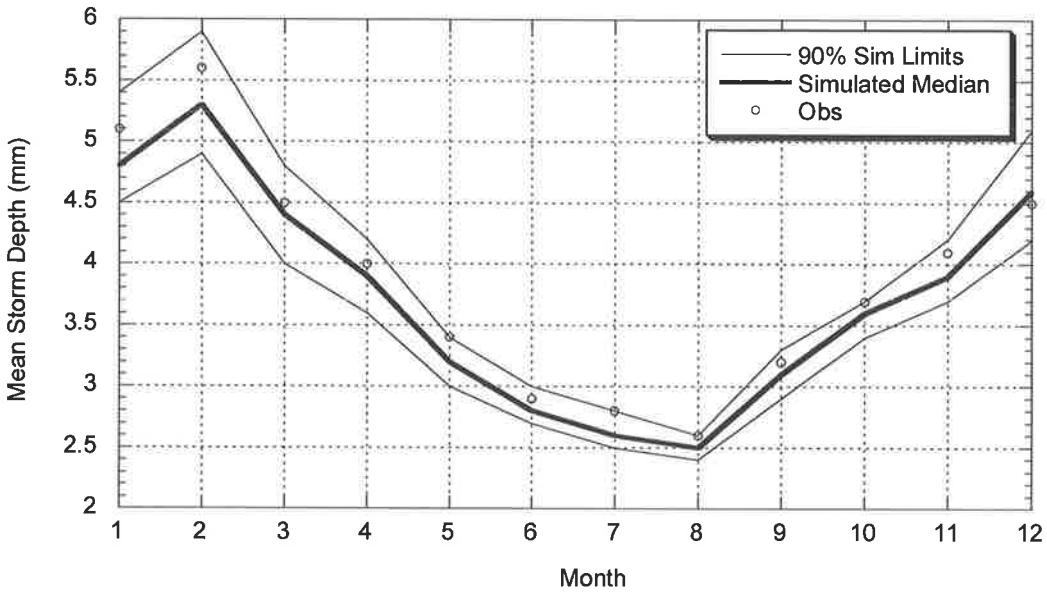

Figure B.3.5: Comparison between Observed and Simulated Average of Event Depths (Melbourne)

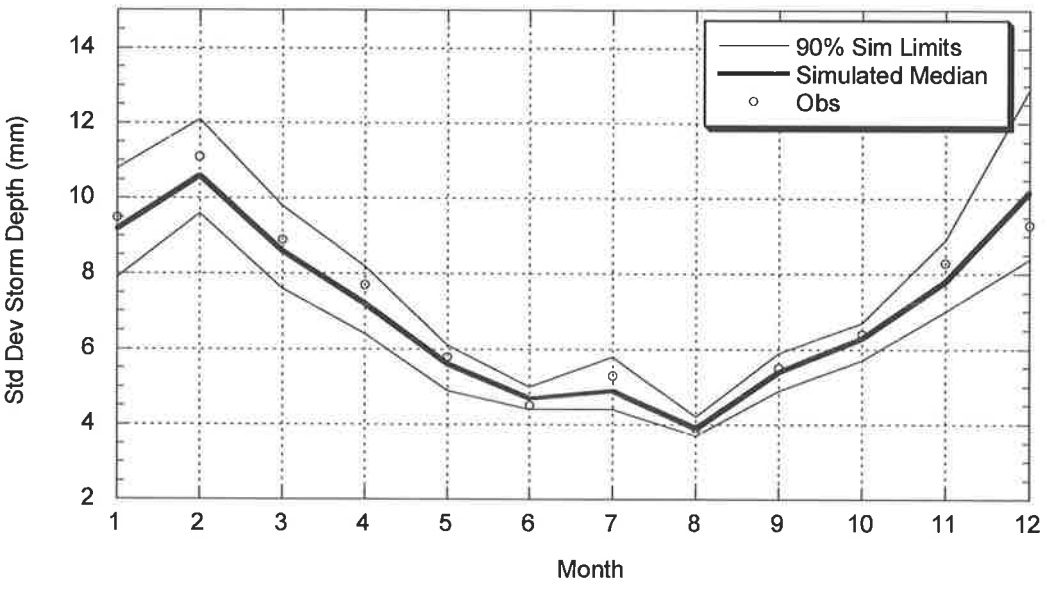

Figure B.3.6: Comparison between Observed and Simulated Standard Deviation of Event Depths (Melbourne)

B.3.2 Simulated and Observed Daily Statistics

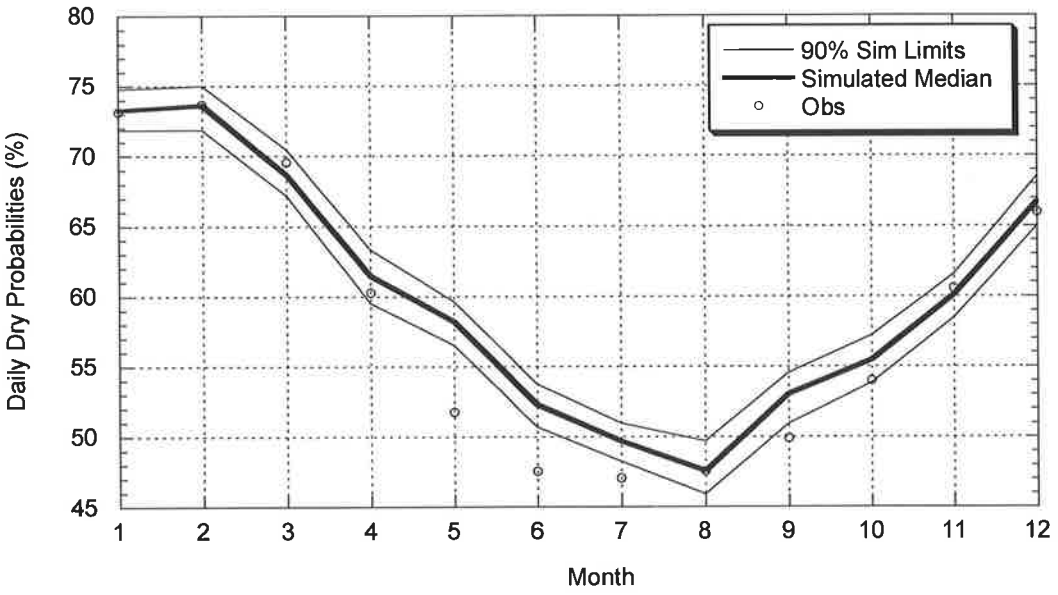

Figure B.3.7: Comparison between Observed and Simulated Daily Dry Probabilities (Melbourne)

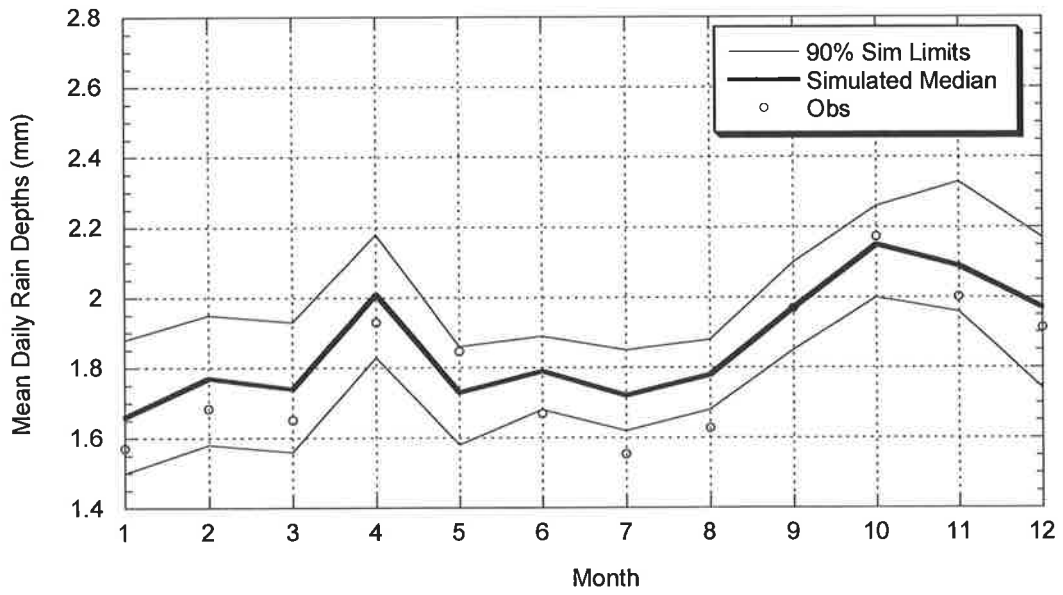

Figure B.3.8: Comparison between Observed and Simulated Daily Mean Depth (Melbourne)

B.3.3 Simulated and Observed Annual and Monthly Rainfall

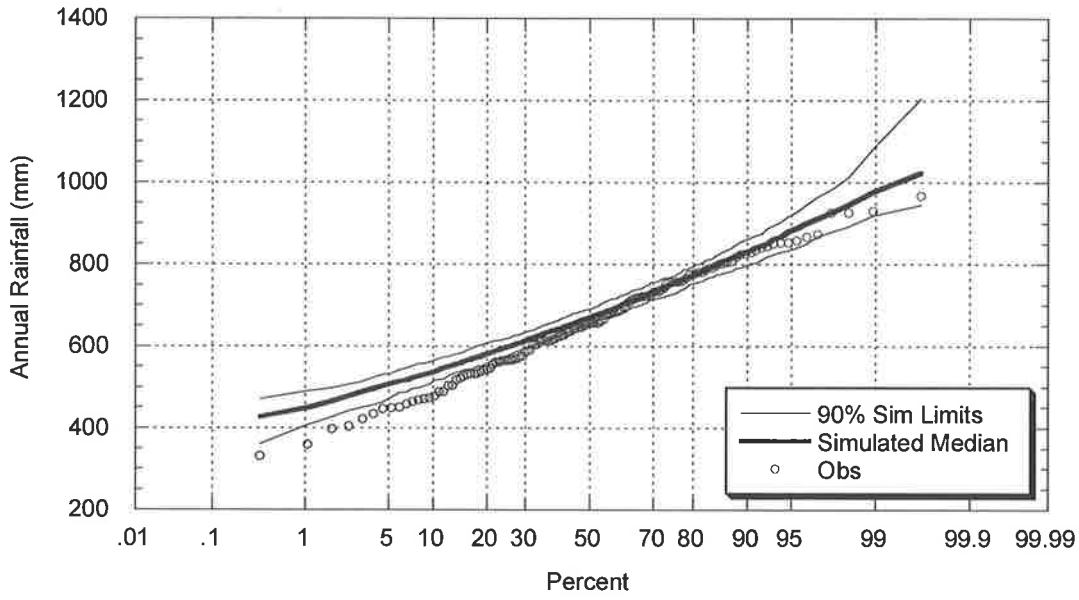

Figure B.3.9: Comparison between Observed and Simulated Annual Rainfall (Melbourne)

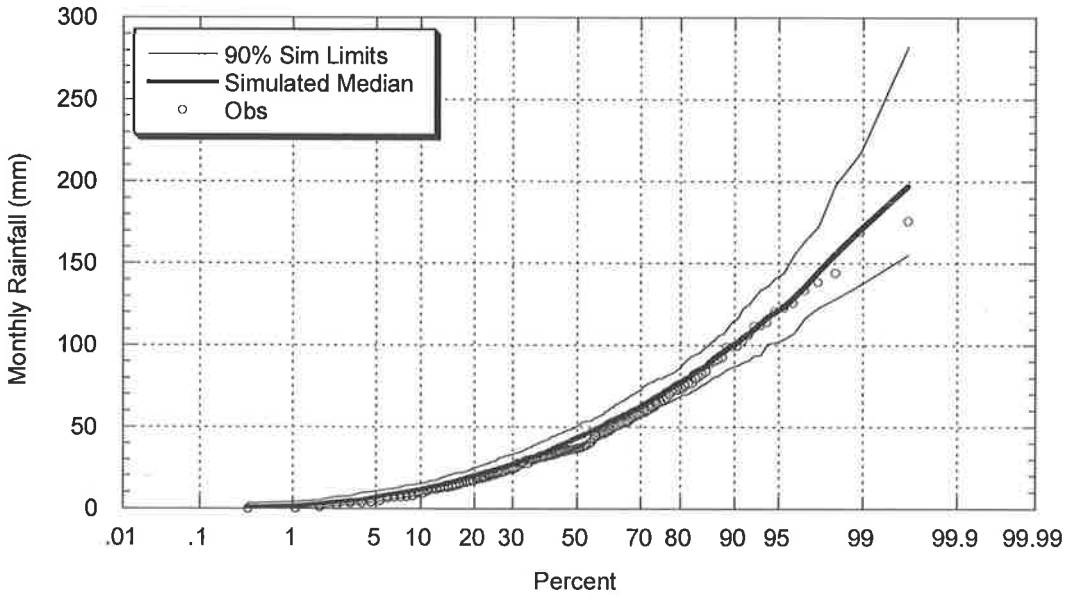

Figure B.3.10: Comparison between Observed and Simulated January Rainfall (Melbourne)

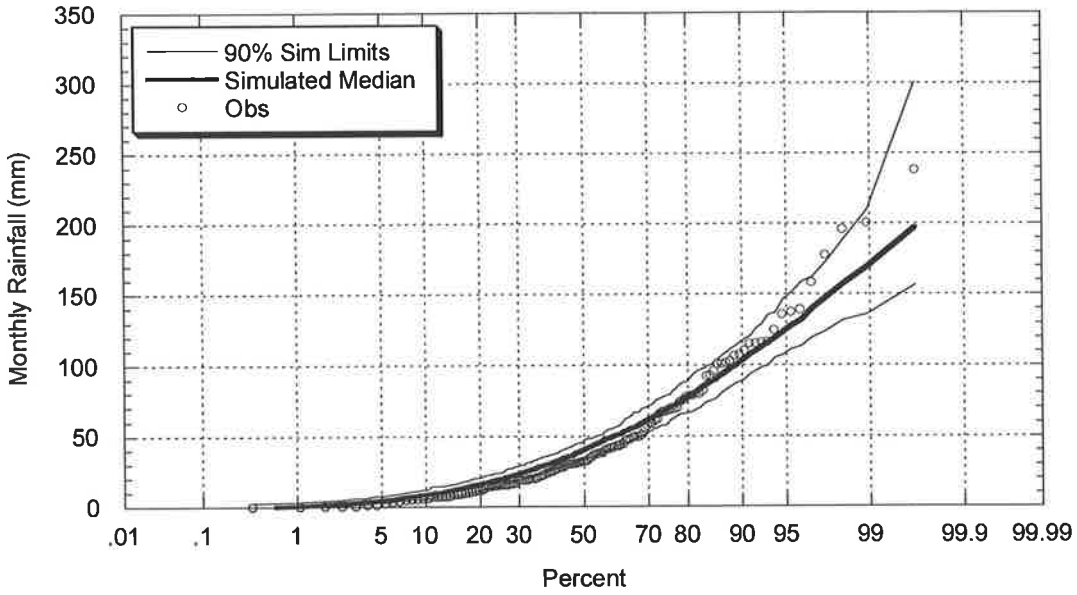

Figure B.3.11: Comparison between Observed and Simulated February Rainfall (Melbourne)

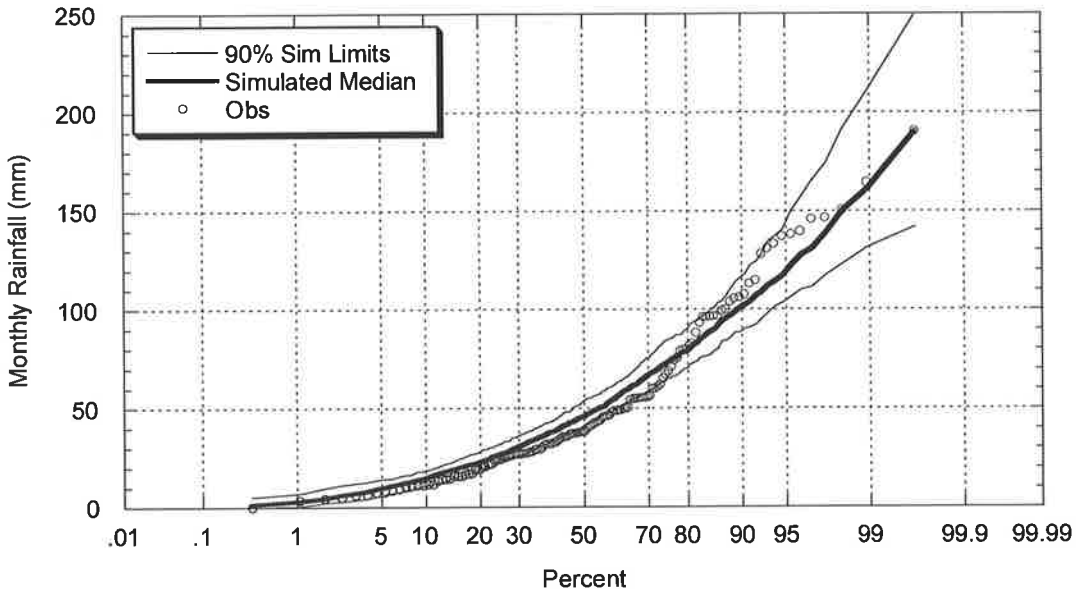

Figure B.3.12: Comparison between Observed and Simulated March Rainfall (Melbourne)

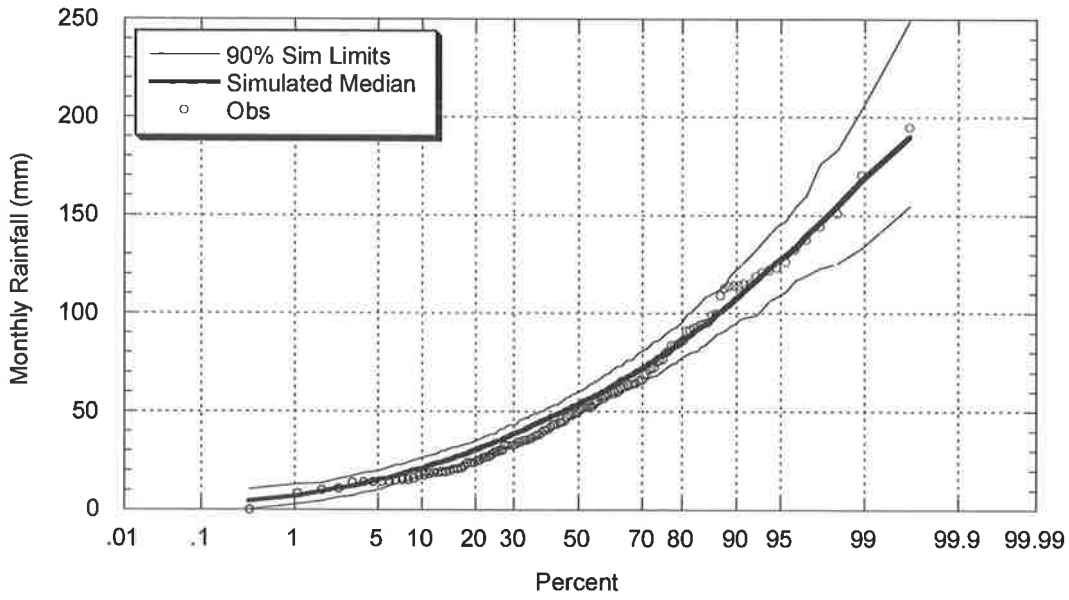

Figure B.3.13: Comparison between Observed and Simulated April Rainfall (Melbourne)

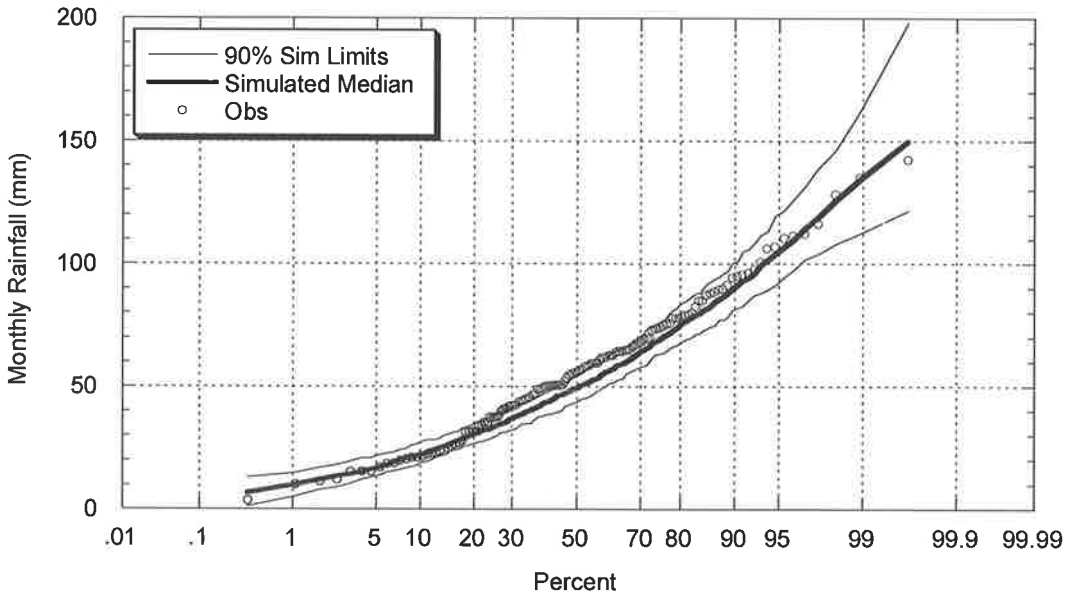

Figure B.3.14: Comparison between Observed and Simulated May Rainfall (Melbourne)

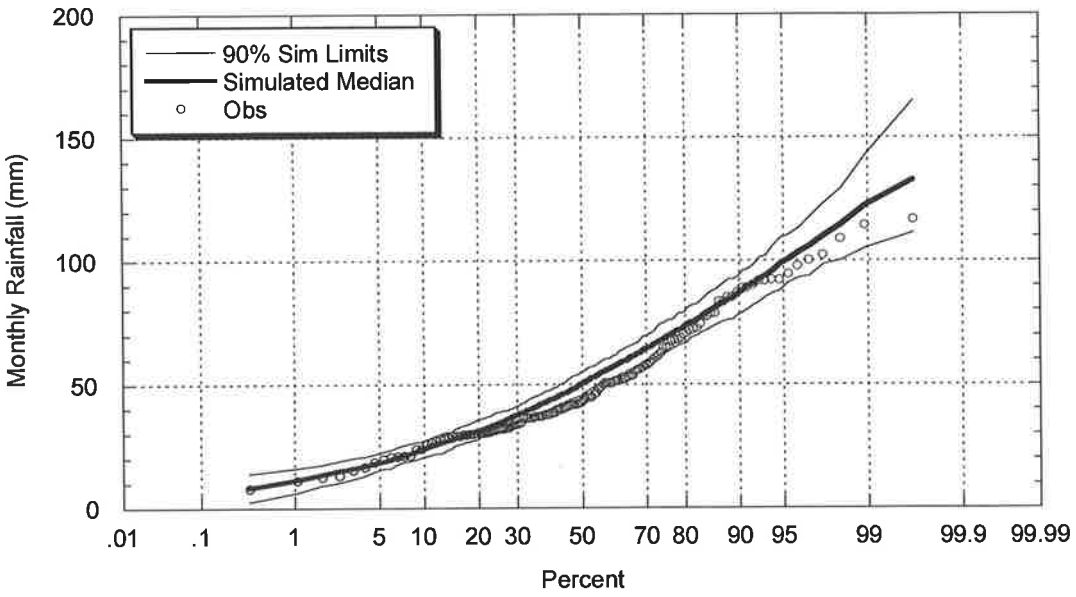

Figure B.3.15: Comparison between Observed and Simulated June Rainfall (Melbourne)

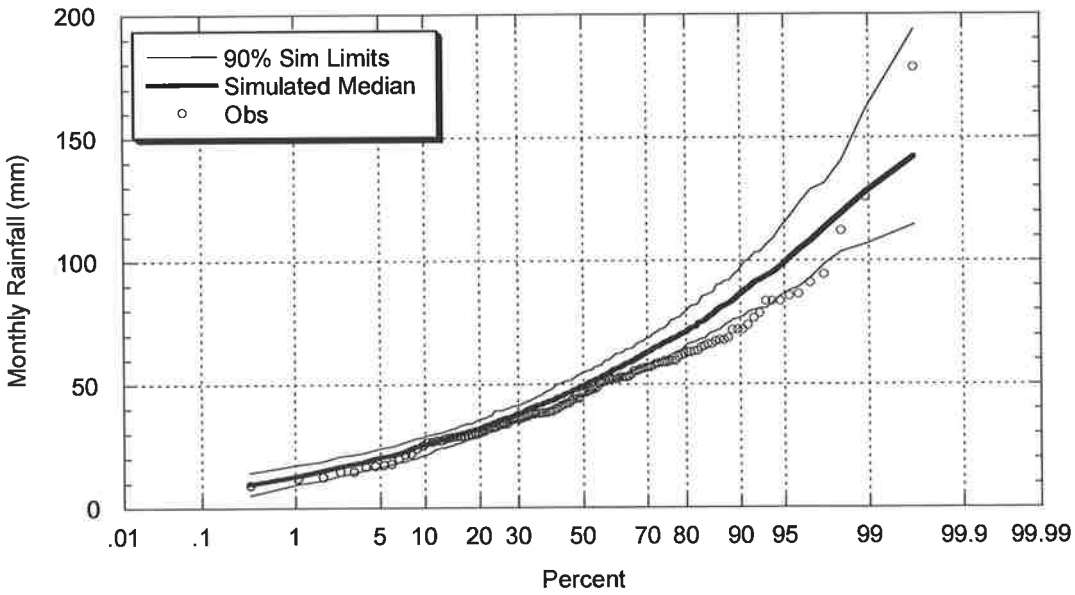

Figure B.3.16: Comparison between Observed and Simulated July Rainfall (Melbourne)

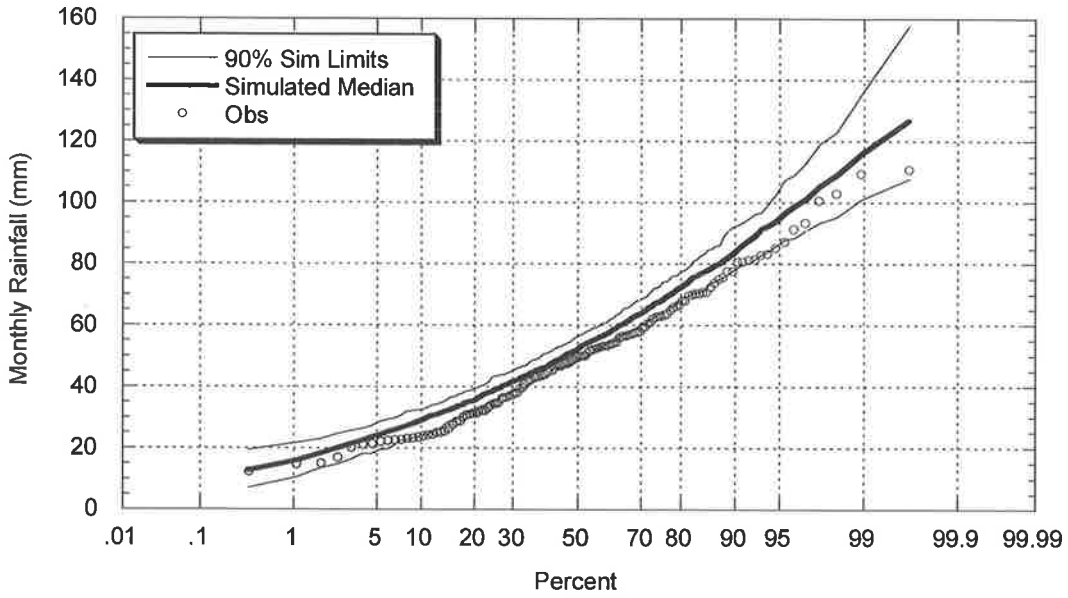

Figure B.3.17: Comparison between Observed and Simulated August Rainfall (Melbourne)

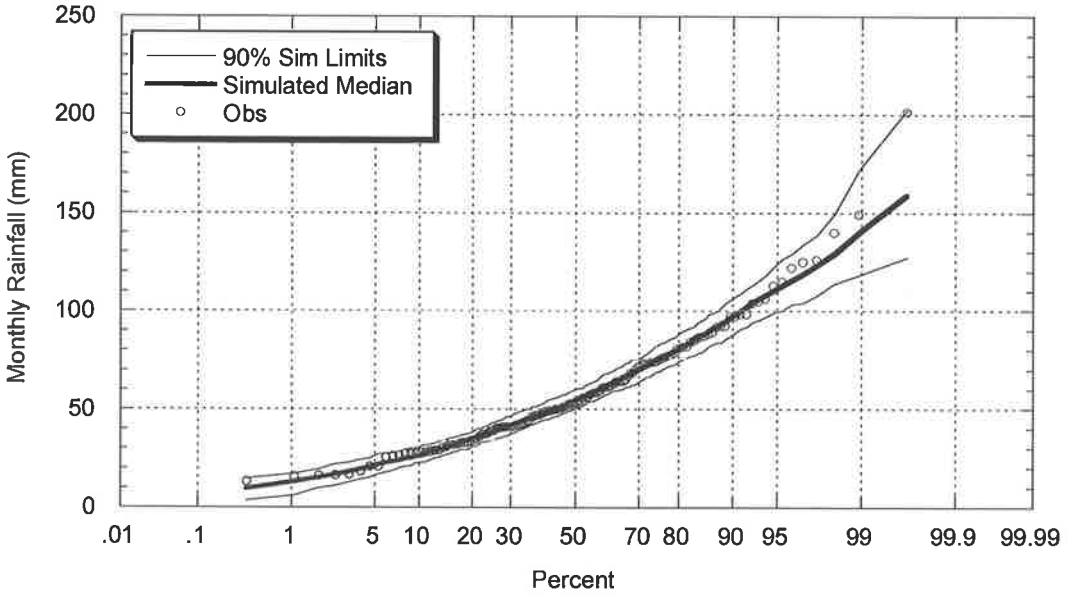

Figure B.3.18: Comparison between Observed and Simulated September Rainfall (Melbourne)

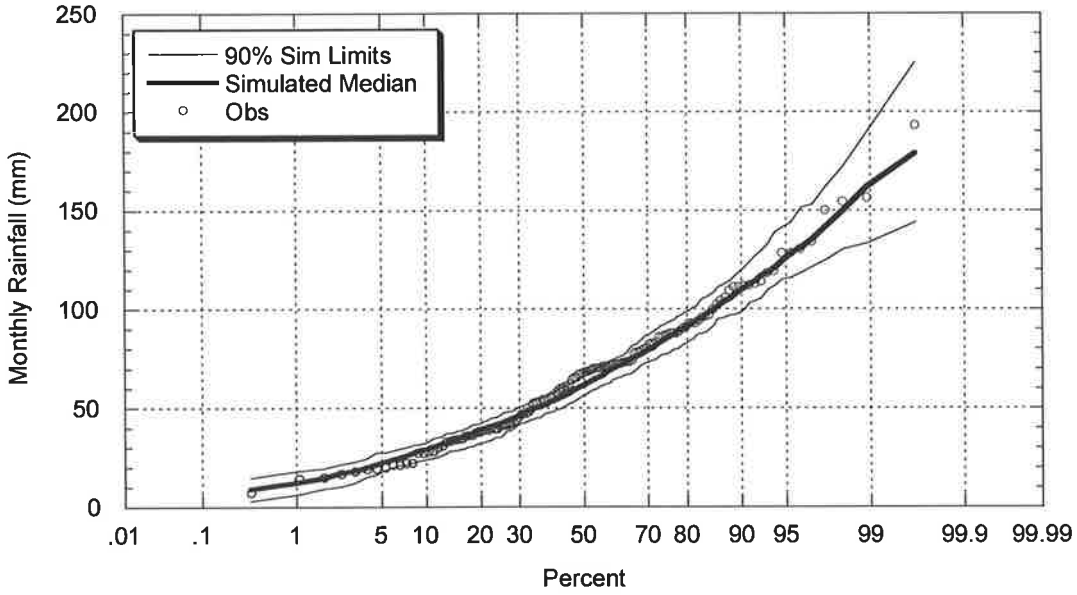

Figure B.3.19: Comparison between Observed and Simulated October Rainfall (Melbourne)

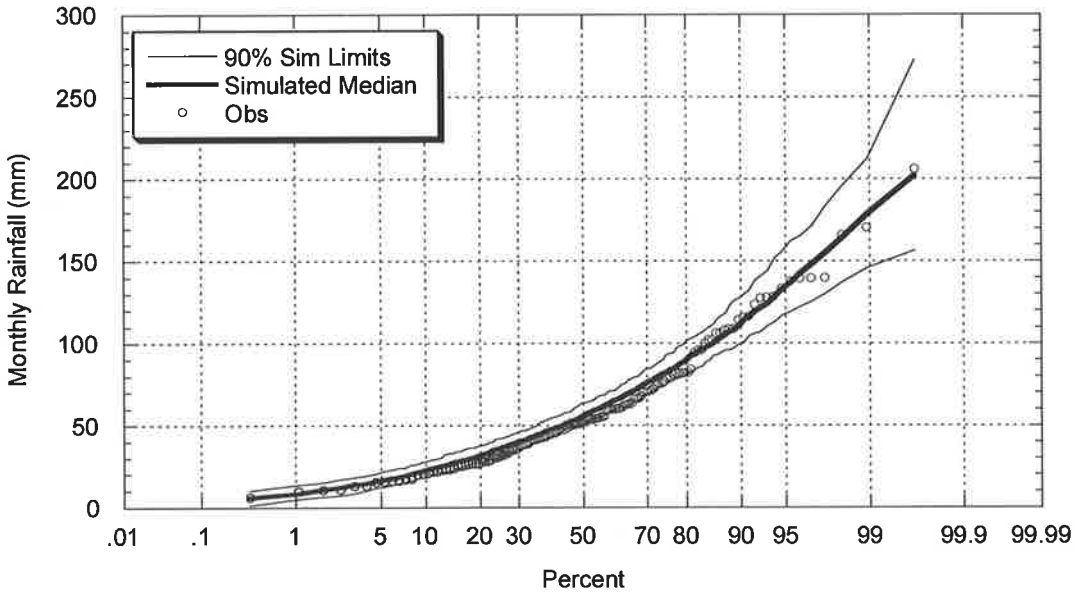

Figure B.3.20: Comparison between Observed and Simulated November Rainfall (Melbourne)

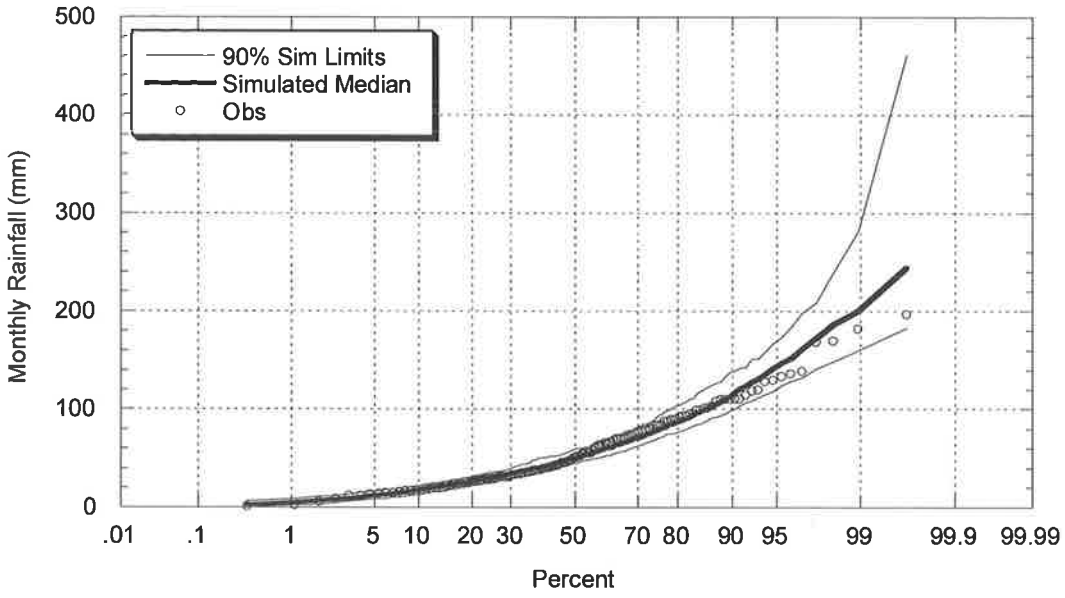

Figure B.3.21: Comparison between Observed and Simulated December Rainfall (Melbourne)

B.3.4 Simulated and Observed Annual Intensity – Frequency - Duration Curves

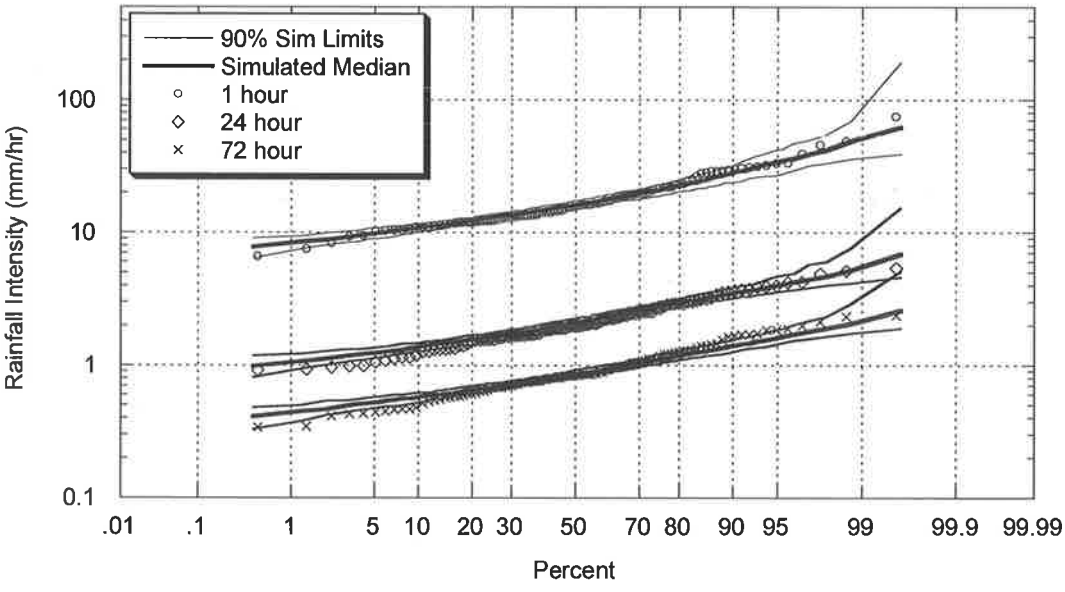

Figure B.3.22: Comparison between Observed and Simulated Annual Intensity Frequency Duration Relationship (Melbourne)

B.4 Perth, Western Australia (BOM# 9034)

B.4.1 Simulated and Observed Storm Event Characteristics

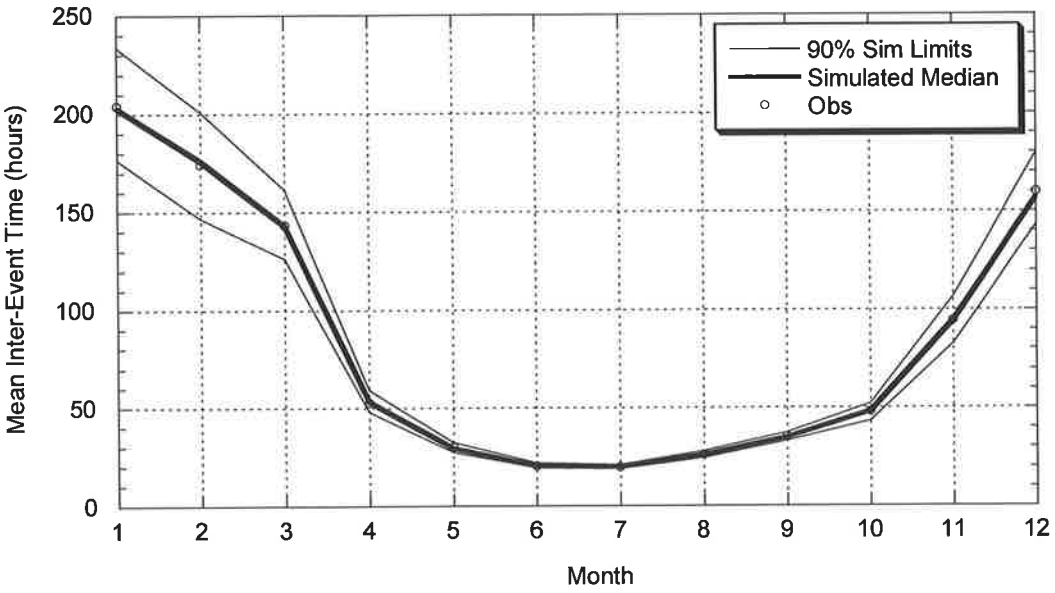

Figure B.4.1: Comparison between Observed and Simulated Mean of Inter-Event Times (Perth)

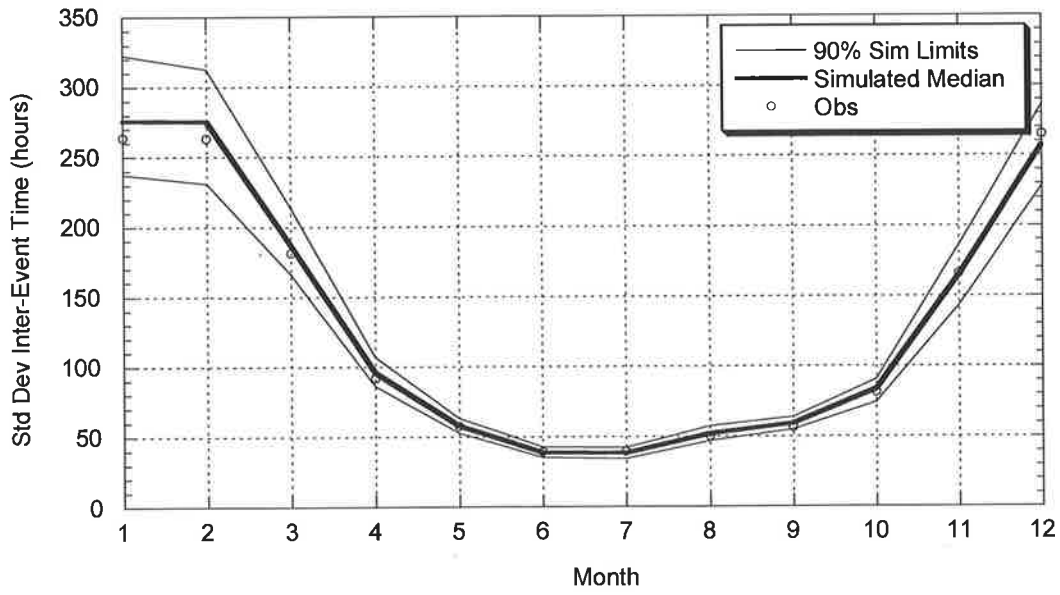

Figure B.4.2: Comparison between Observed and Simulated Standard Deviation of Inter-Event Times (Perth)

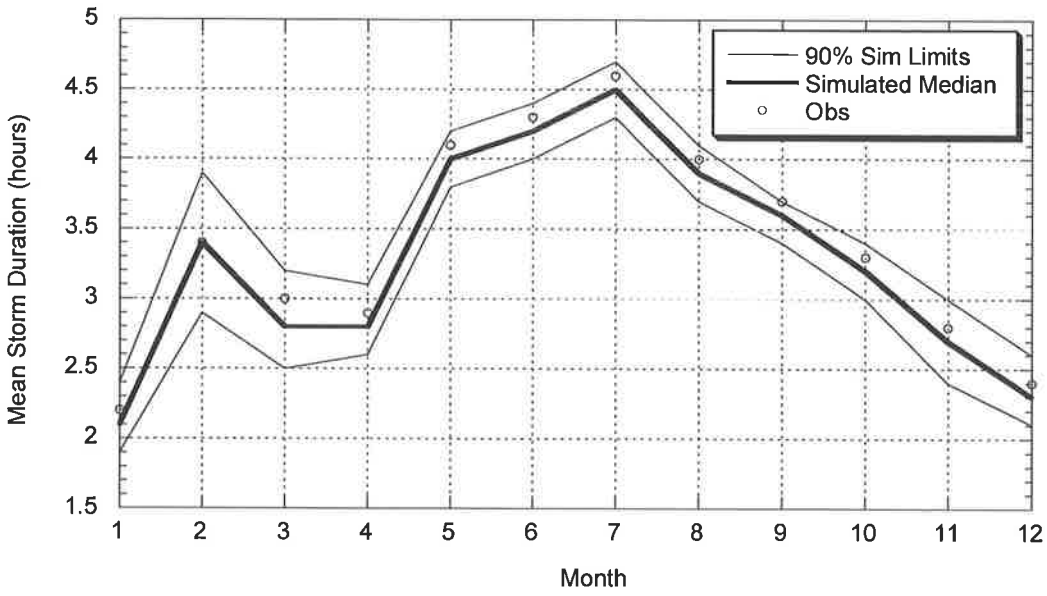

Figure B.4.3: Comparison between Observed and Simulated Mean of Event Storm Durations (Perth)

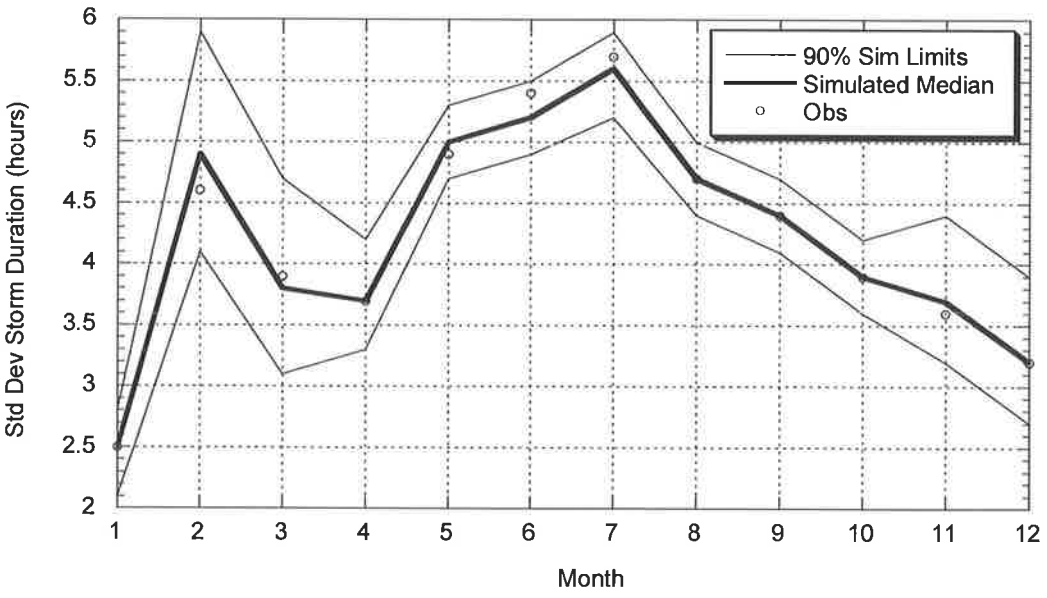

Figure B.4.4: Comparison between Observed and Simulated Standard Deviation of Event Storm Durations (Perth)

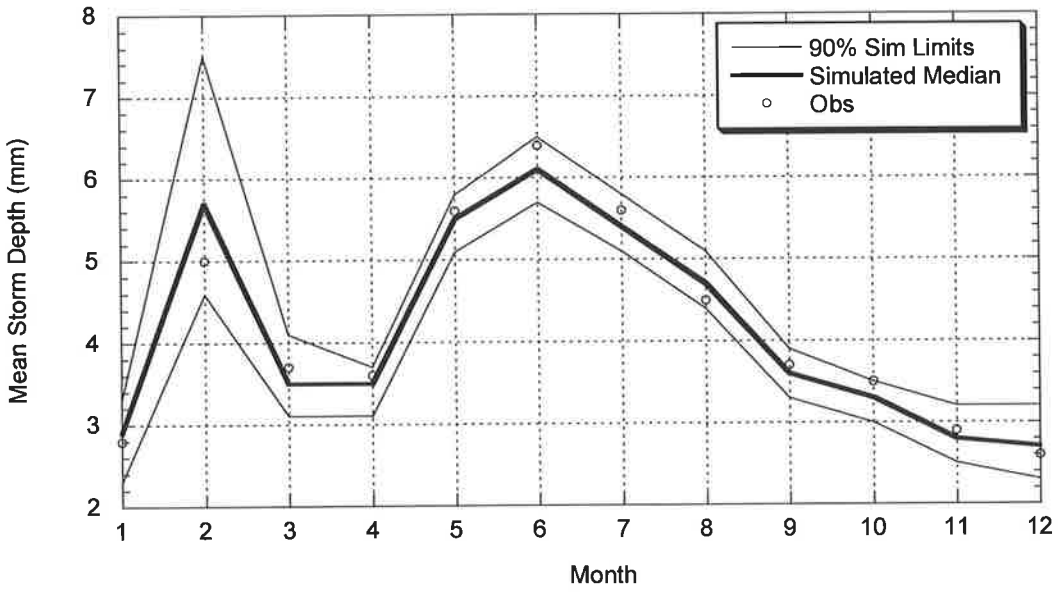

Figure B.4.5: Comparison between Observed and Simulated Average of Event Depths (Perth)

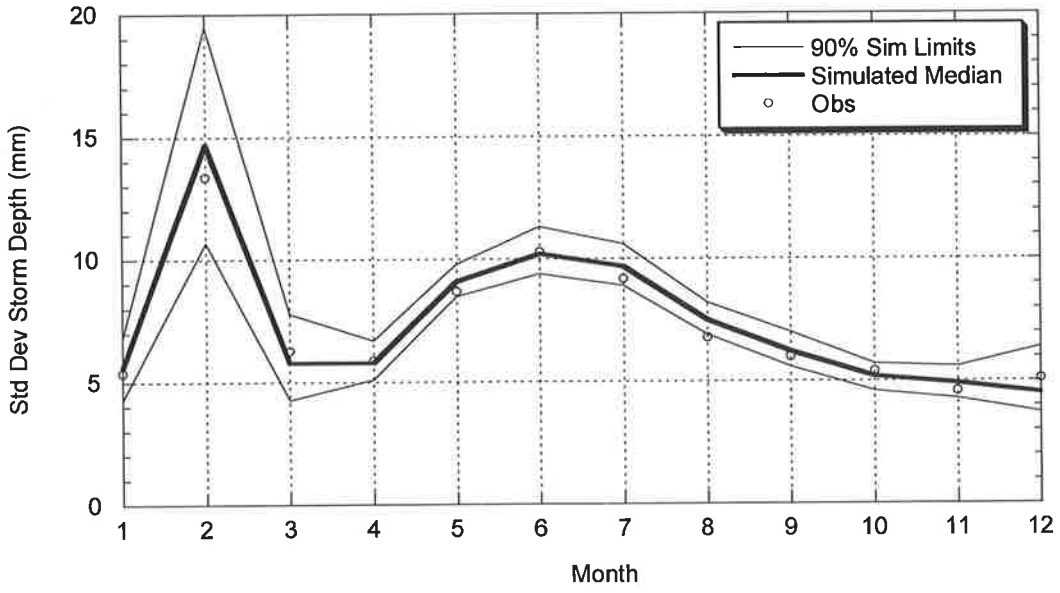

Figure B.4.6: Comparison between Observed and Simulated Standard Deviation of Event Depths (Perth)

B.4.2 Simulated and Observed Daily Statistics

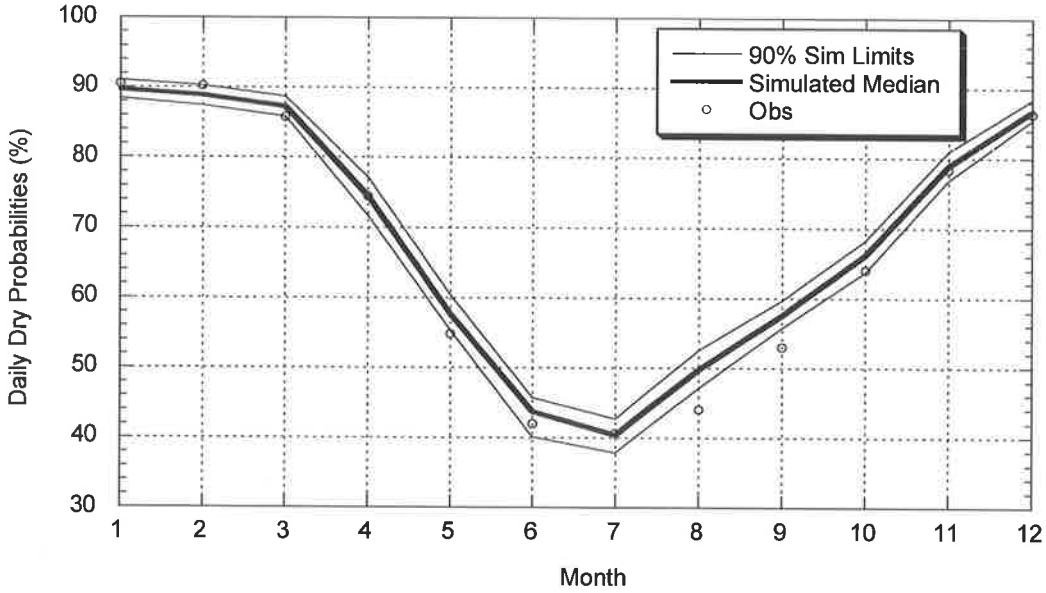

Figure B.4.7: Comparison between Observed and Simulated Daily Dry Probabilities (Perth)

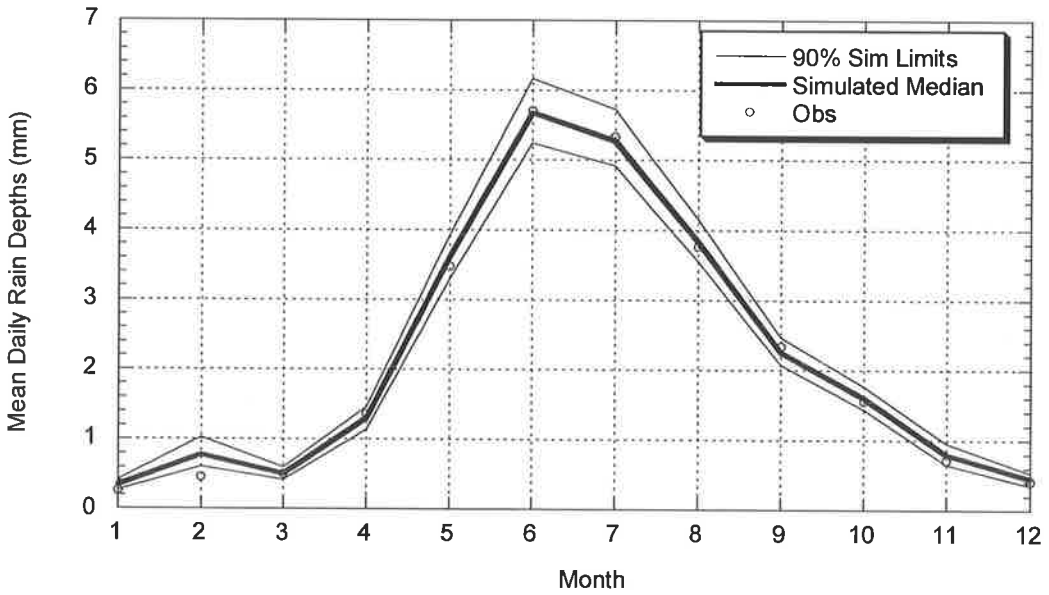

Figure B.4.8: Comparison between Observed and Simulated Daily Mean Depth (Perth)

B.4.3 Simulated and Observed Annual and Monthly Rainfall

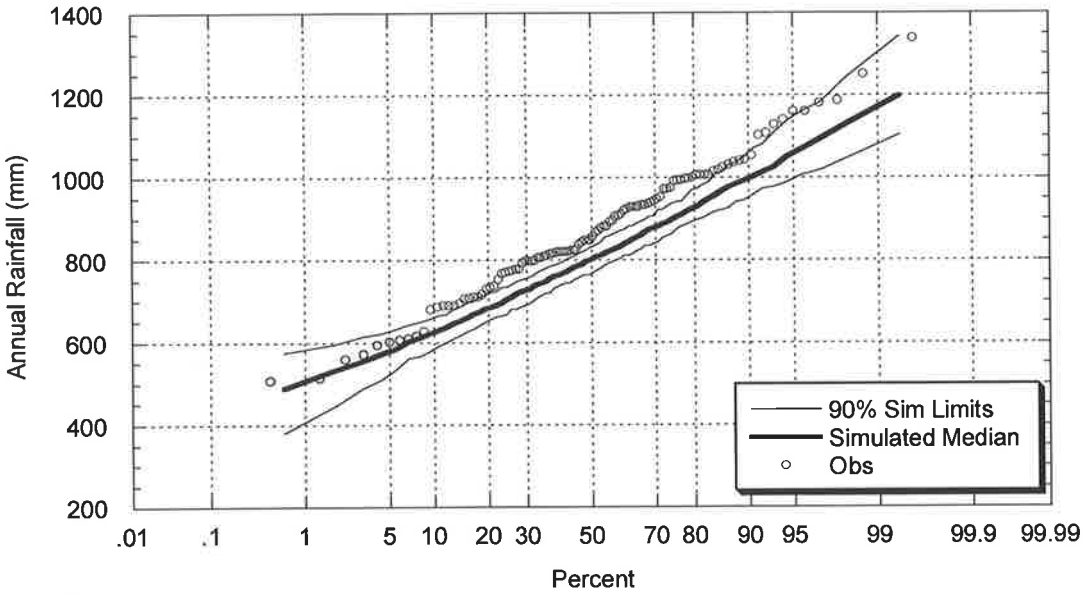

Figure B.4.9: Comparison between Observed and Simulated Annual Rainfall (Perth)

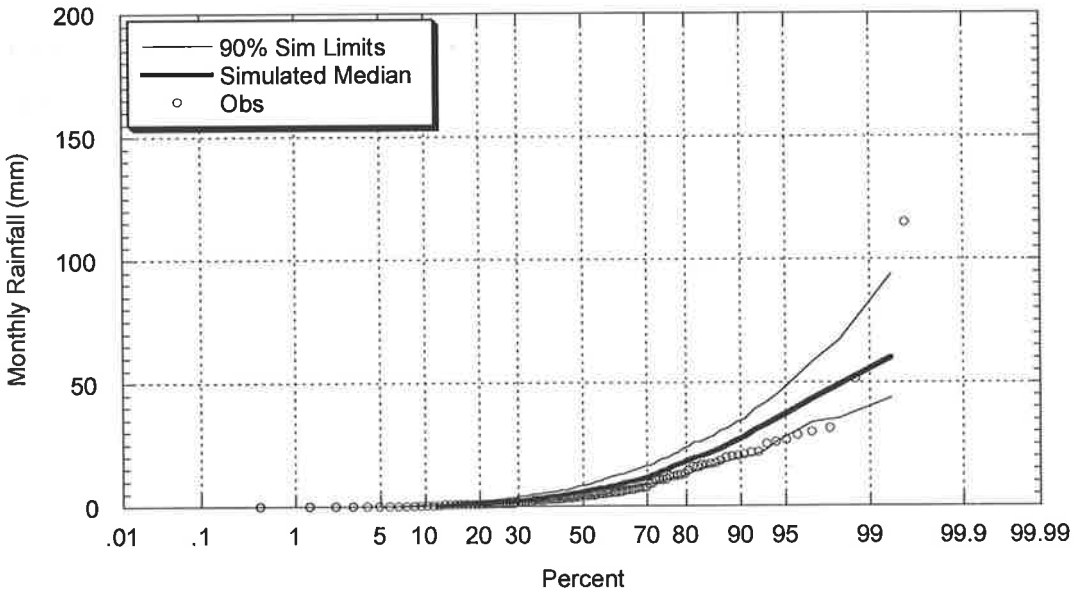

Figure B.4.10: Comparison between Observed and Simulated January Rainfall (Perth)

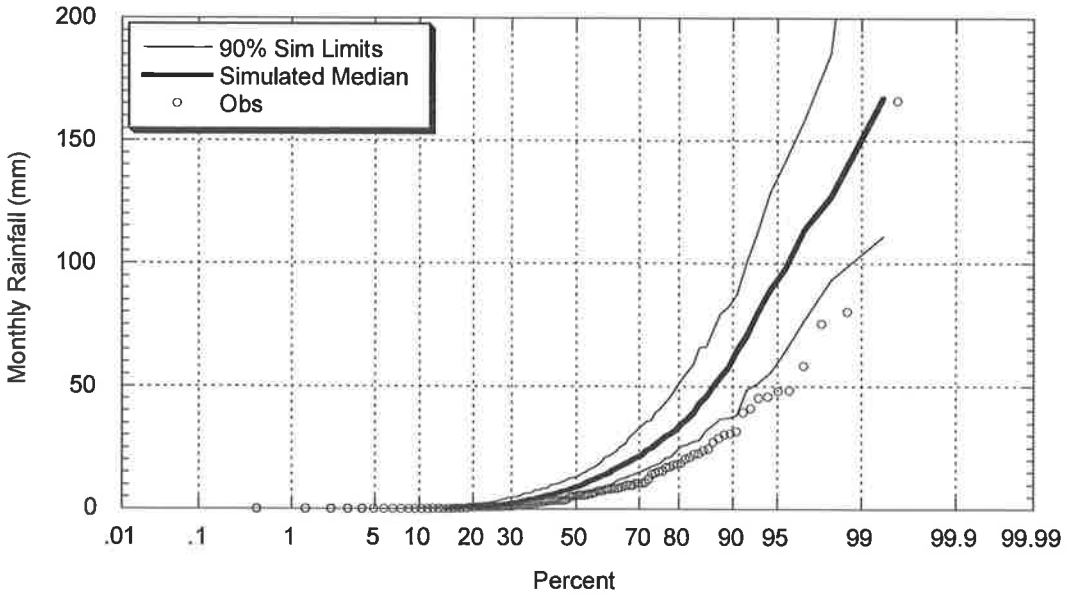

Figure B.4.11: Comparison between Observed and Simulated February Rainfall (Perth)

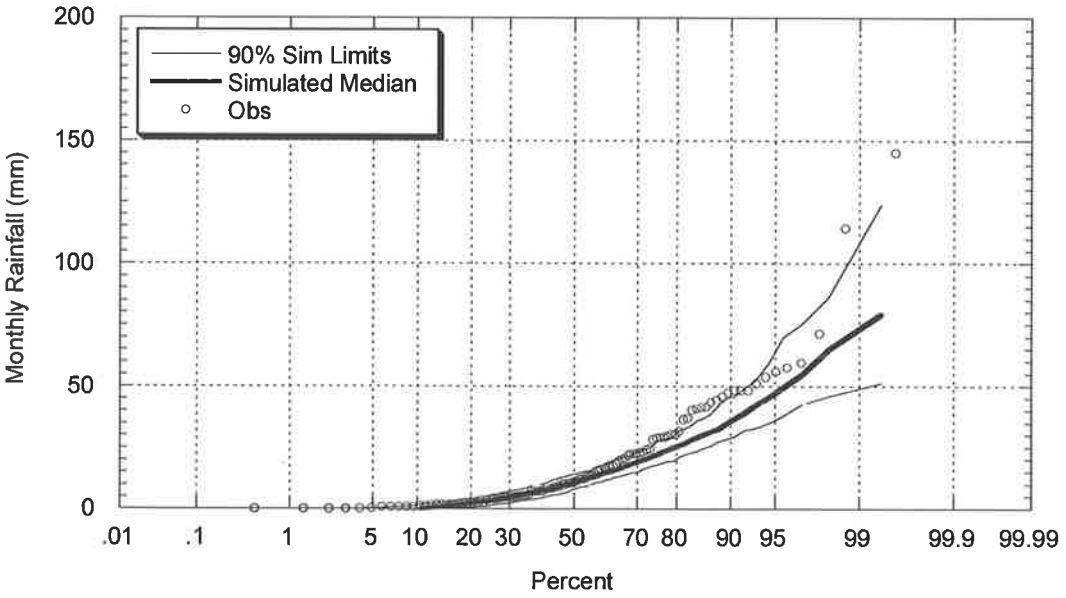

Figure B.4.12: Comparison between Observed and Simulated March Rainfall (Perth)

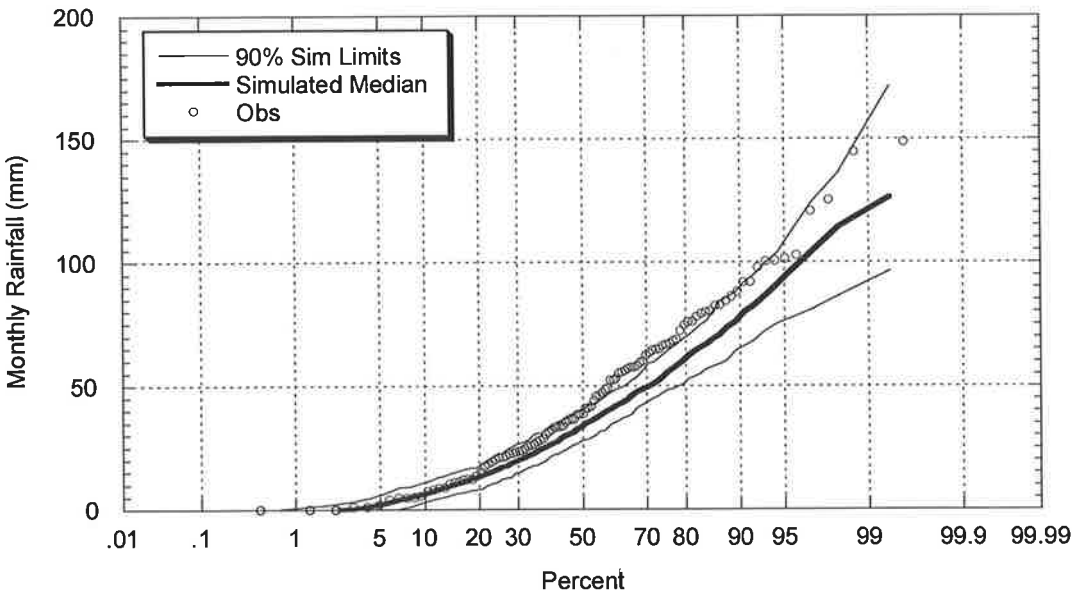

Figure B.4.13: Comparison between Observed and Simulated April Rainfall (Perth)

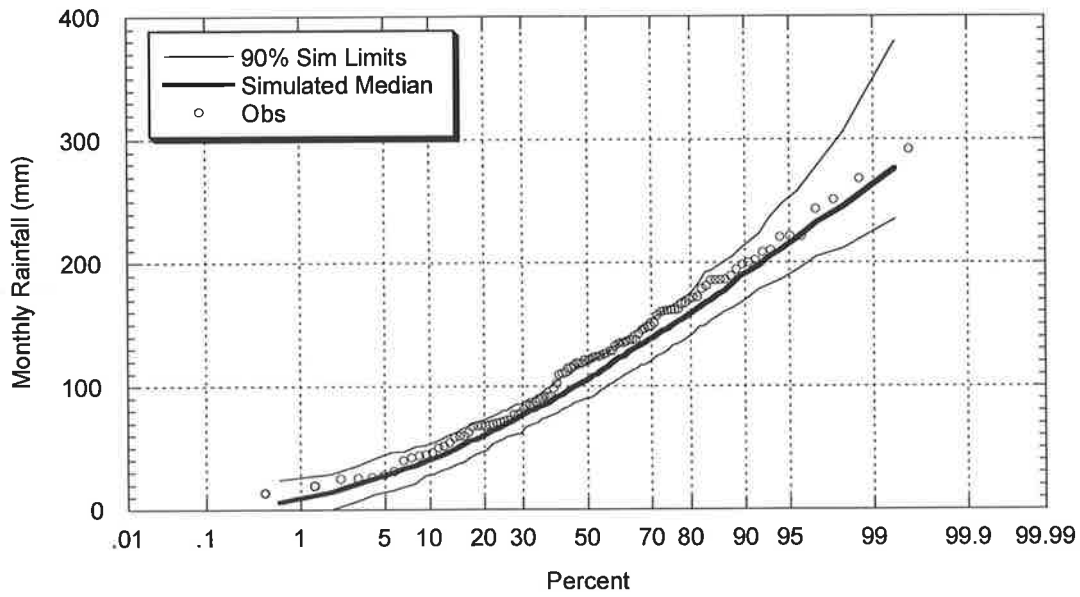

Figure B.4.14: Comparison between Observed and Simulated May Rainfall (Perth)

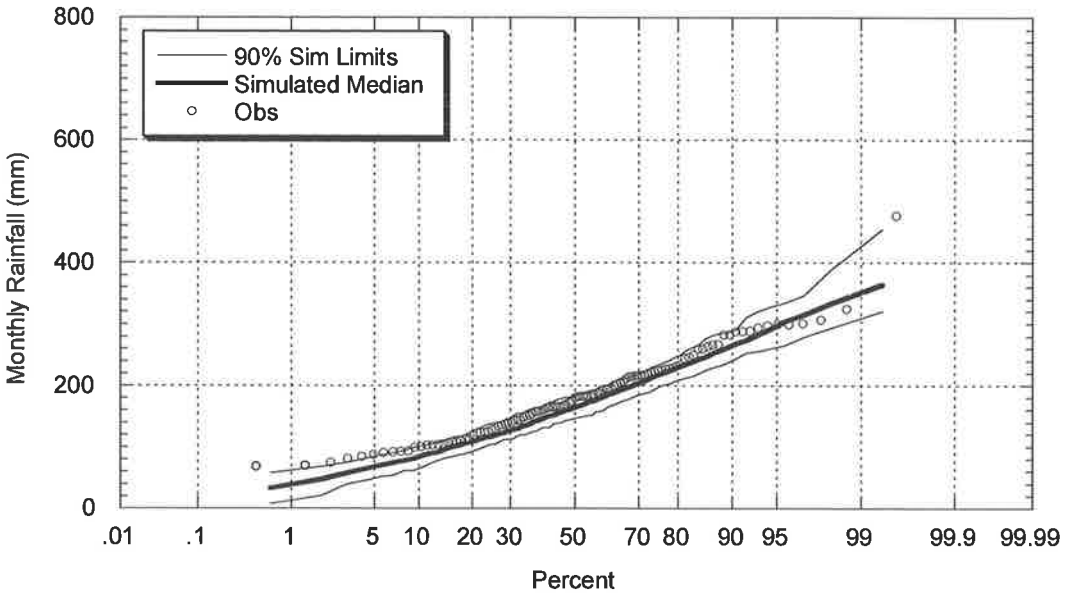

Figure B.4.15: Comparison between Observed and Simulated June Rainfall (Perth)

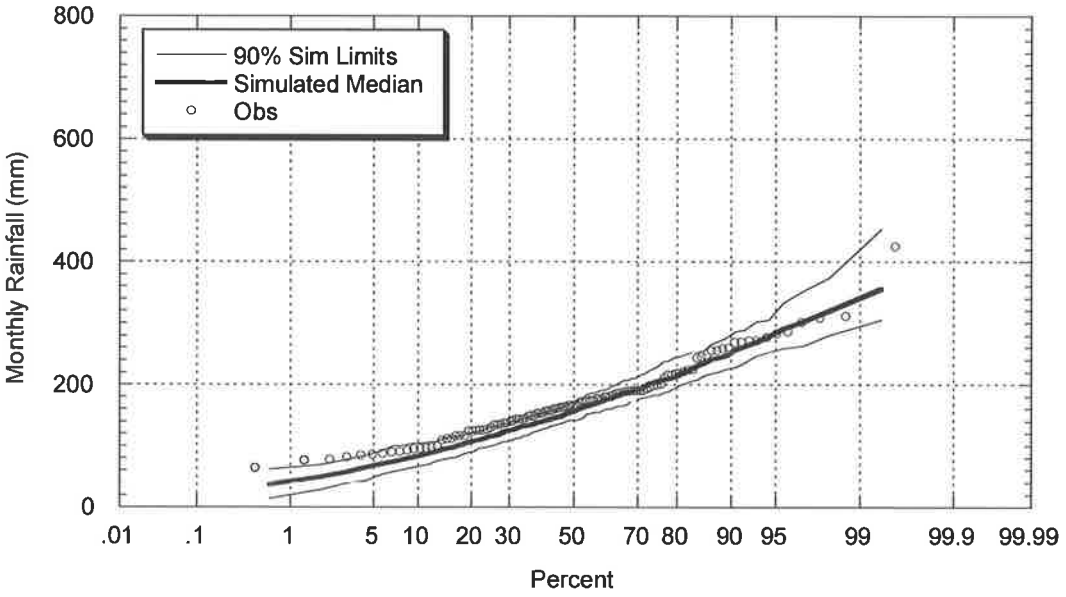

Figure B.4.16: Comparison between Observed and Simulated July Rainfall (Perth)

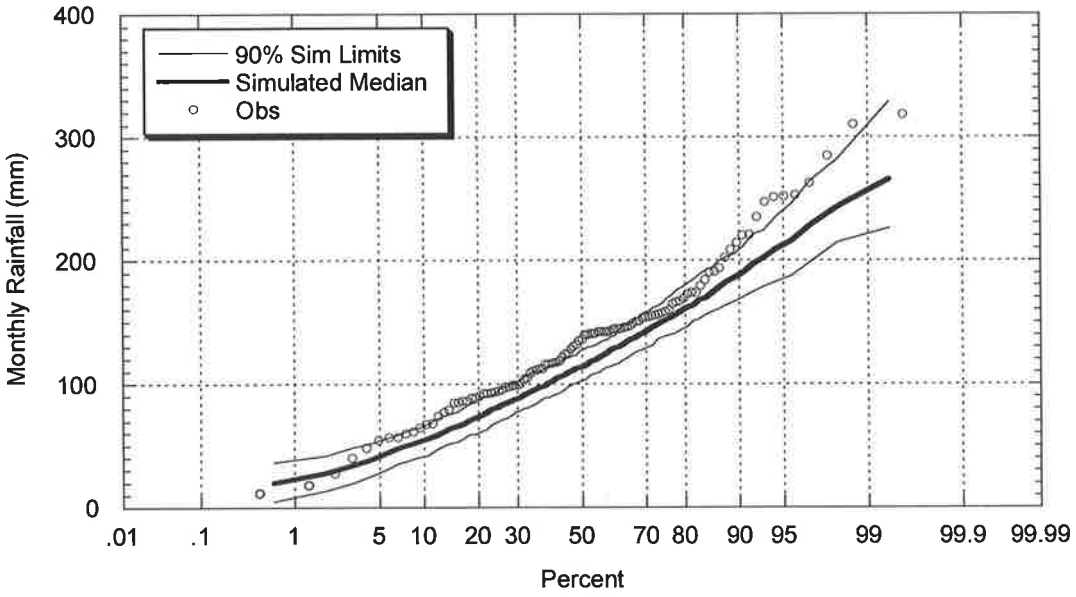

Figure B.4.17: Comparison between Observed and Simulated August Rainfall (Perth)

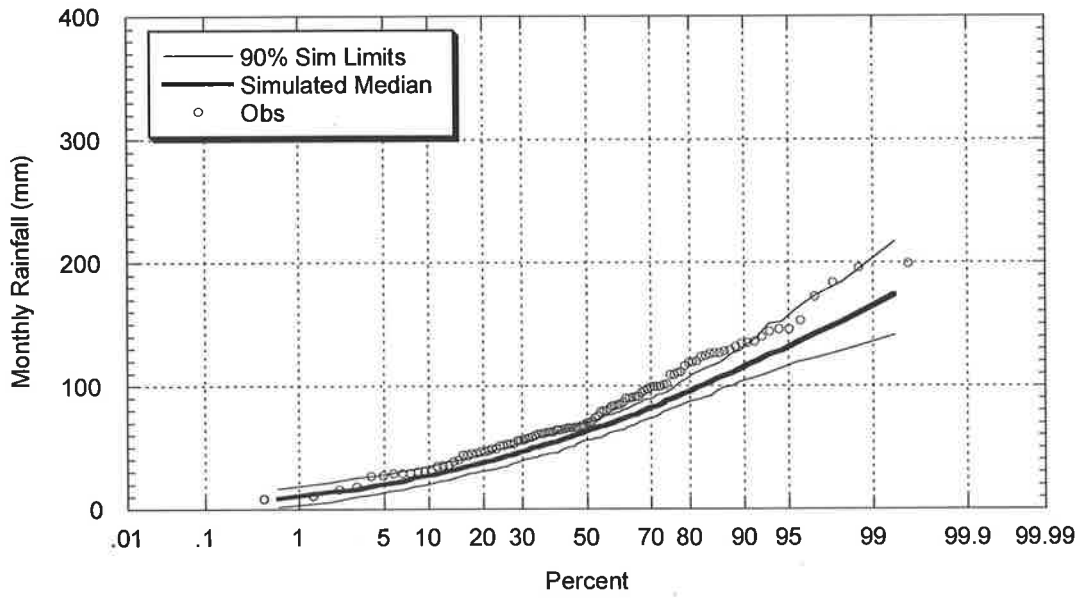

Figure B.4.18: Comparison between Observed and Simulated September Rainfall (Perth)

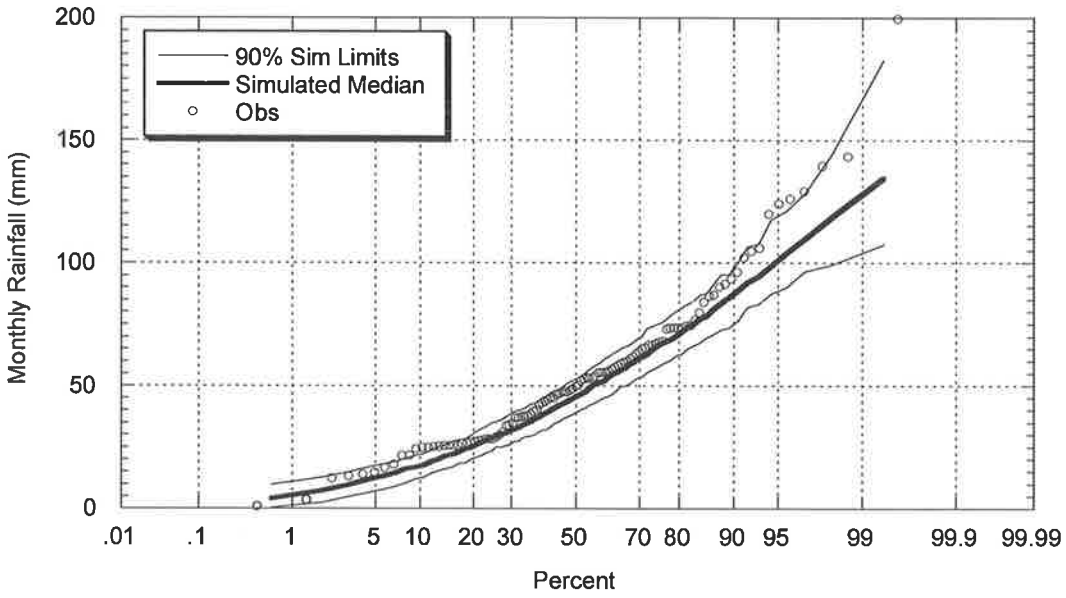

Figure B.4.19: Comparison between Observed and Simulated October Rainfall (Perth)

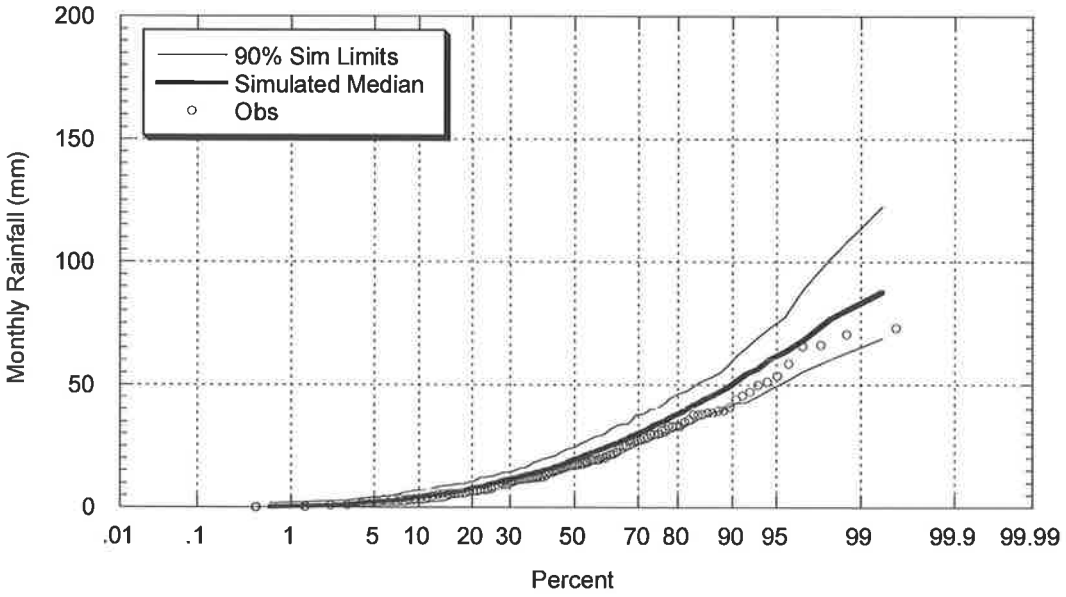

Figure B.4.20 Comparison between Observed and Simulated November Rainfall (Perth)

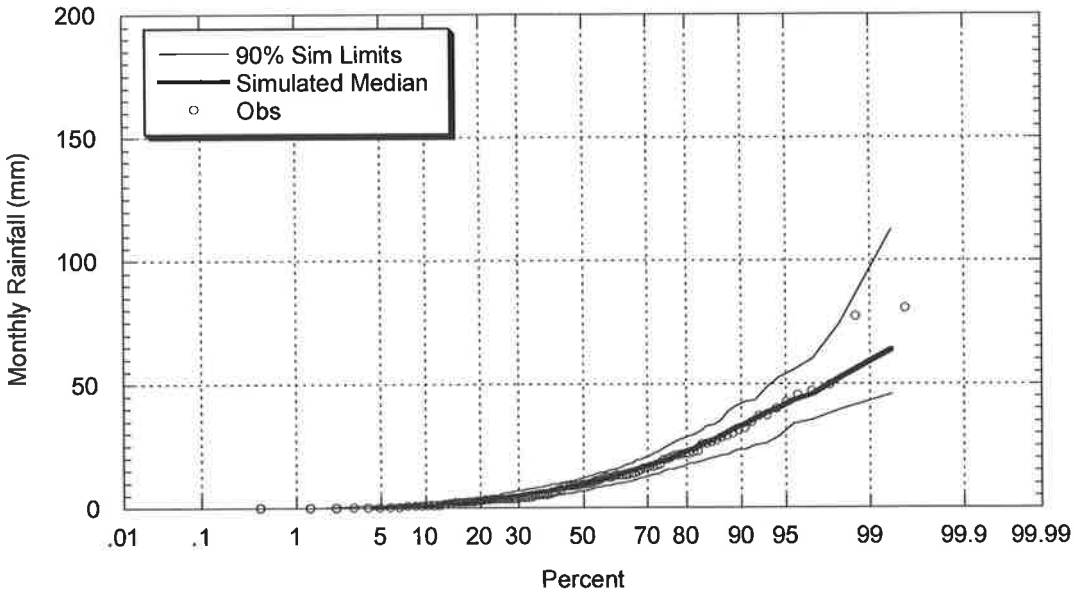

Figure B.4.21 Comparison between Observed and Simulated December Rainfall (Perth)

B.4.4 Simulated and Observed Annual Intensity – Frequency - Duration Curves

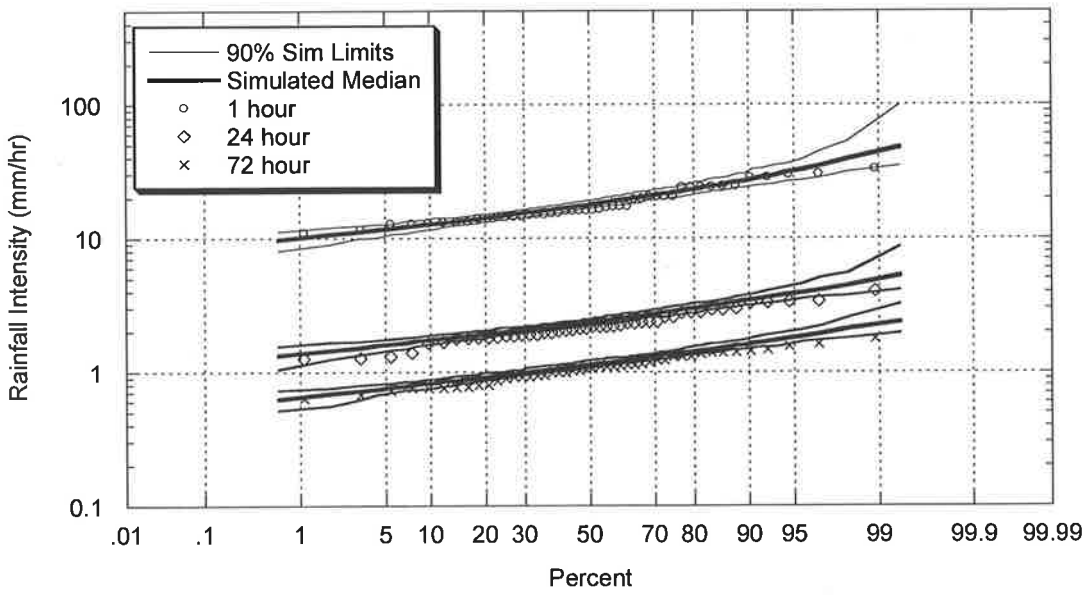

Figure B.4.22: Comparison between Observed and Simulated Annual Intensity Frequency Duration Relationship Perth)

B.5 Sydney, New South Wales (BOM# 66062)

B.5.1 Simulated and Observed Storm Event Characteristics

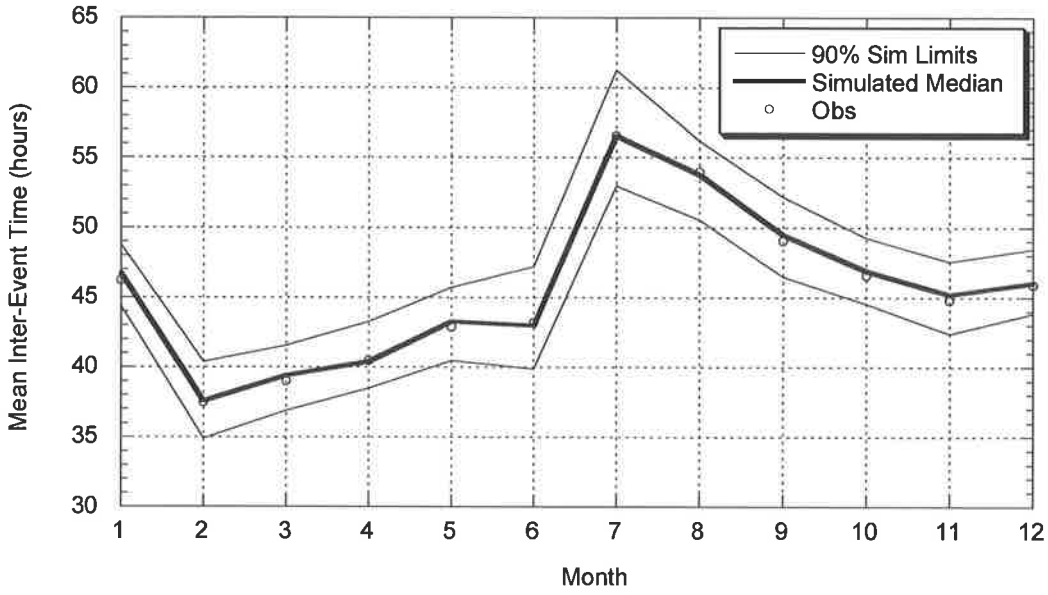

Figure B.5.1: Comparison between Observed and Simulated Mean of Inter-Event Times (Sydney)

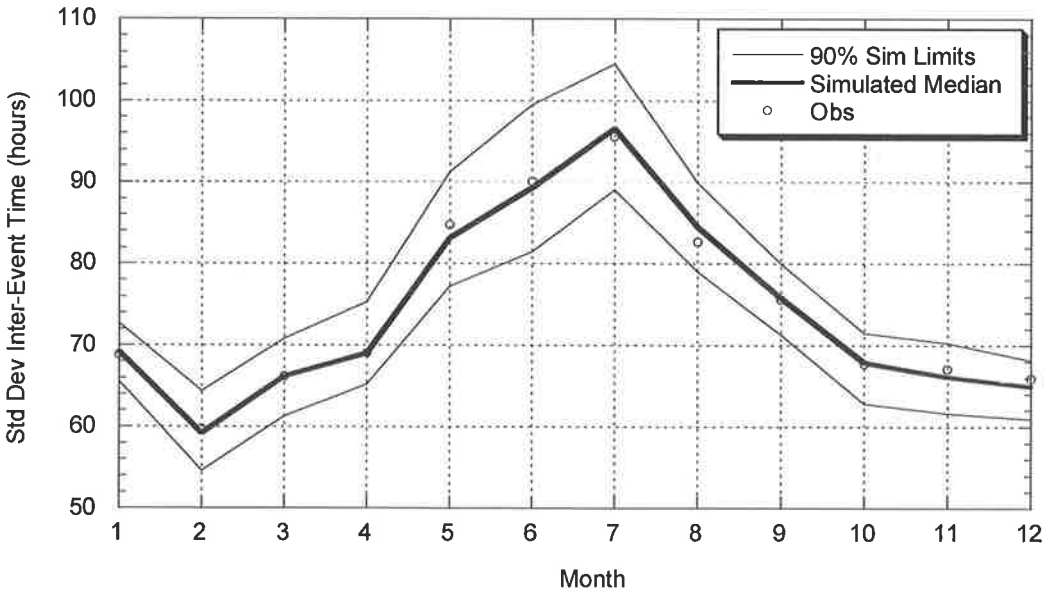

Figure B.5.2: Comparison between Observed and Simulated Standard Deviation of Inter-Event Times (Sydney)

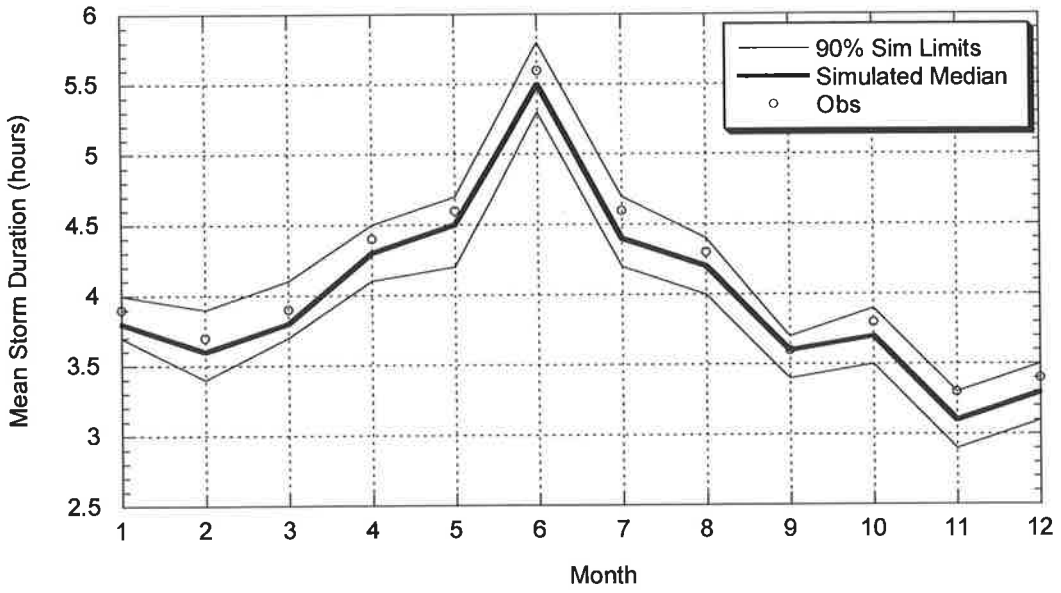

Figure B.5.3: Comparison between Observed and Simulated Mean of Event Storm Durations (Sydney)

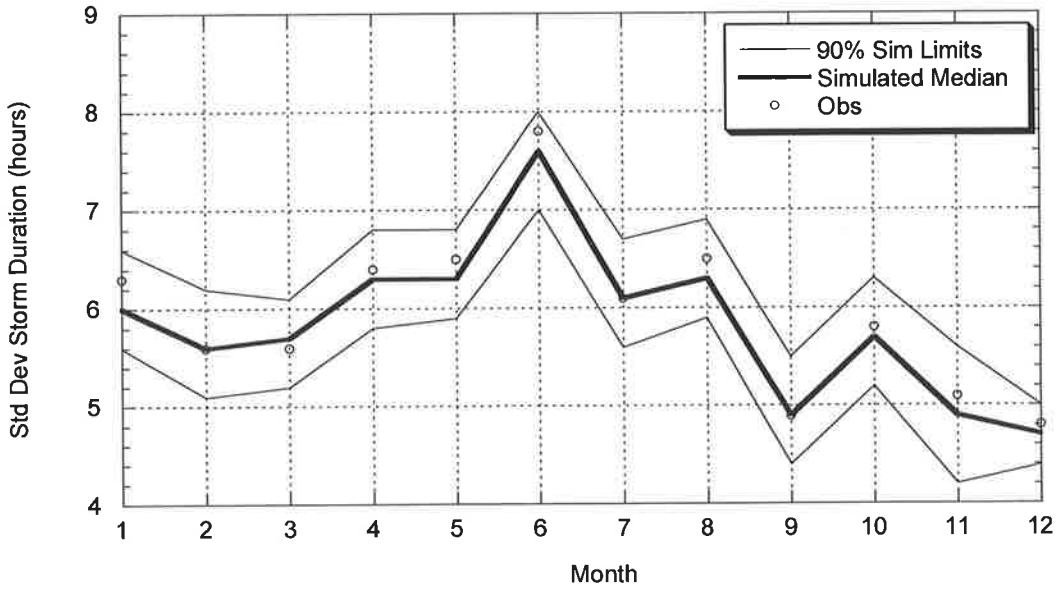

Figure B.5.4: Comparison between Observed and Simulated Standard Deviation of Event Storm Durations (Sydney)

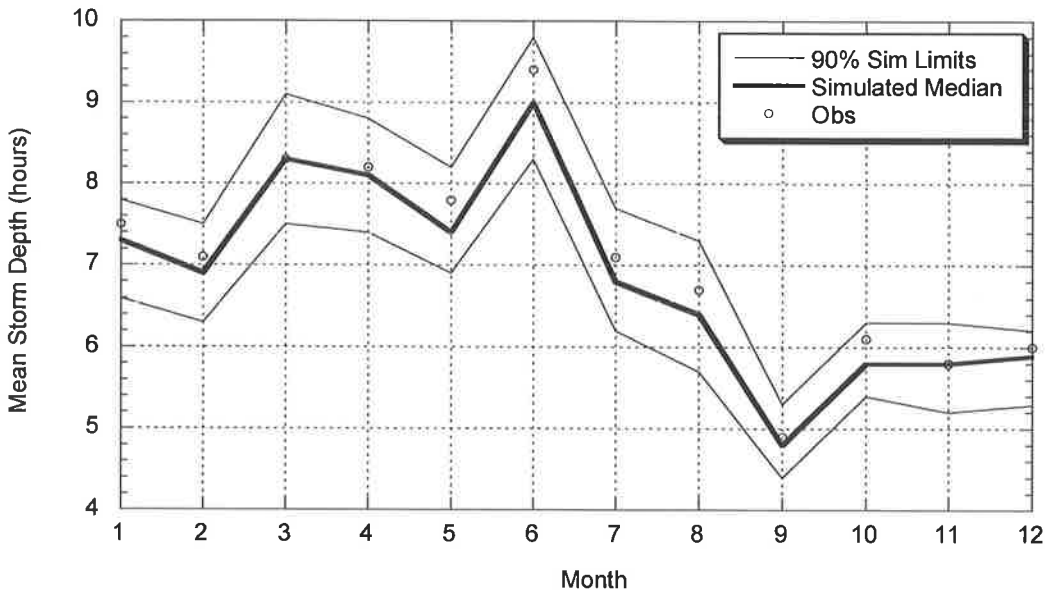

Figure B.5.5: Comparison between Observed and Simulated Average of Event Depths (Sydney)

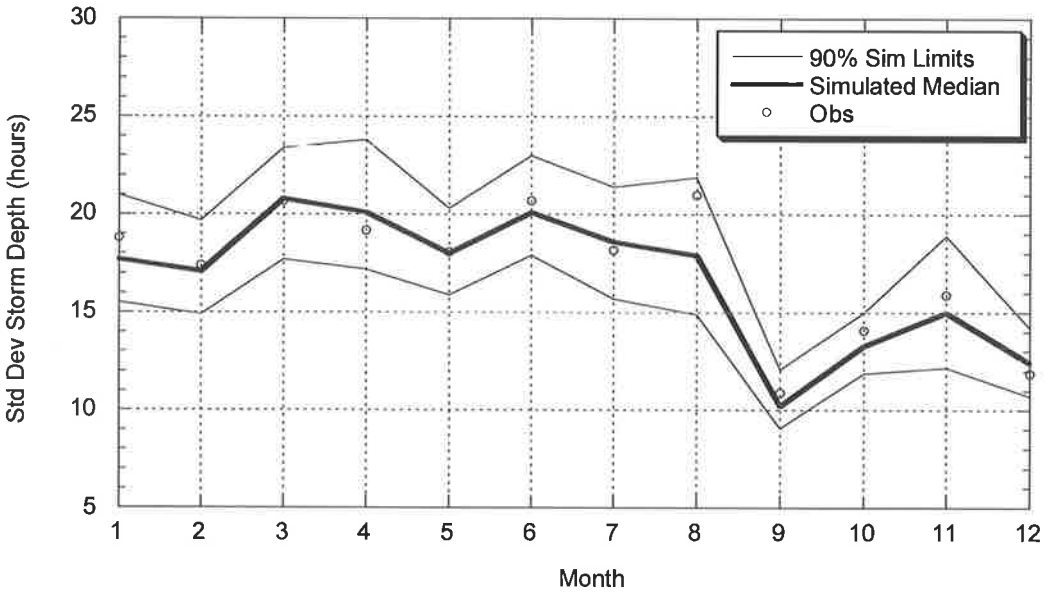

Figure B.5.6: Comparison between Observed and Simulated Standard Deviation of Event Depths (Sydney)

B.5.2 Simulated and Observed Daily Statistics

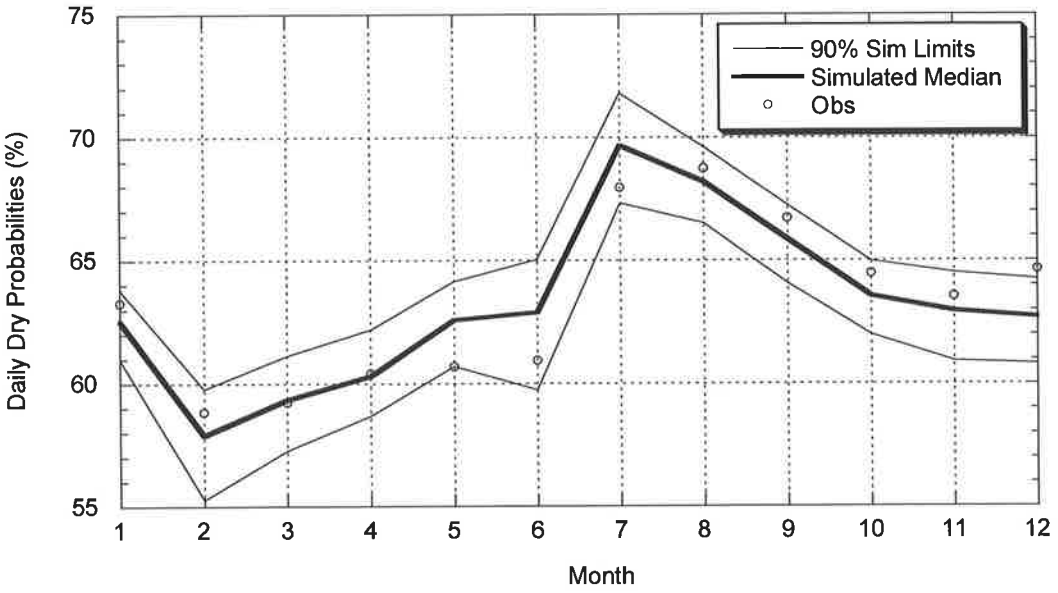

Figure B.5.7: Comparison between Observed and Simulated Daily Dry Probabilities (Sydney)

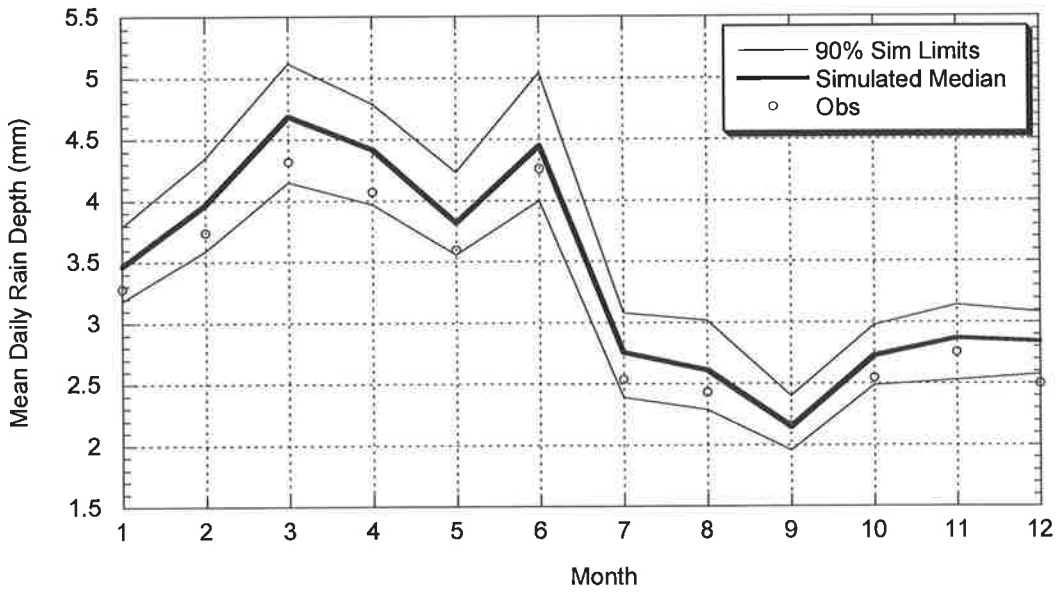

Figure B.5.8: Comparison between Observed and Simulated Daily Mean Depth (Sydney)

B.5.3 Simulated and Observed Annual and Monthly Rainfall

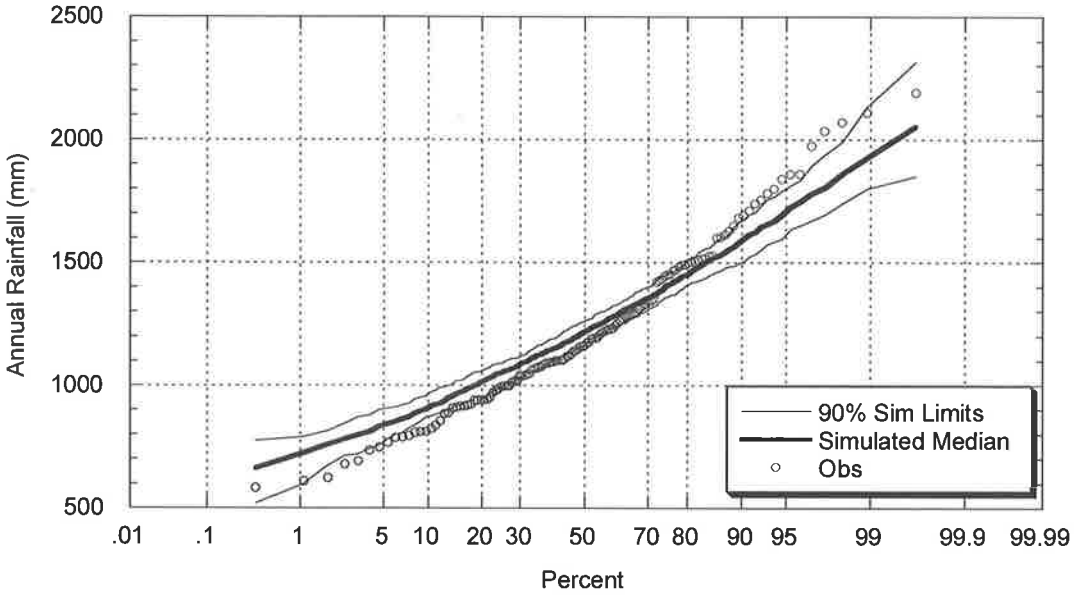

Figure B.5.9: Comparison between Observed and Simulated Annual Rainfall (Sydney)

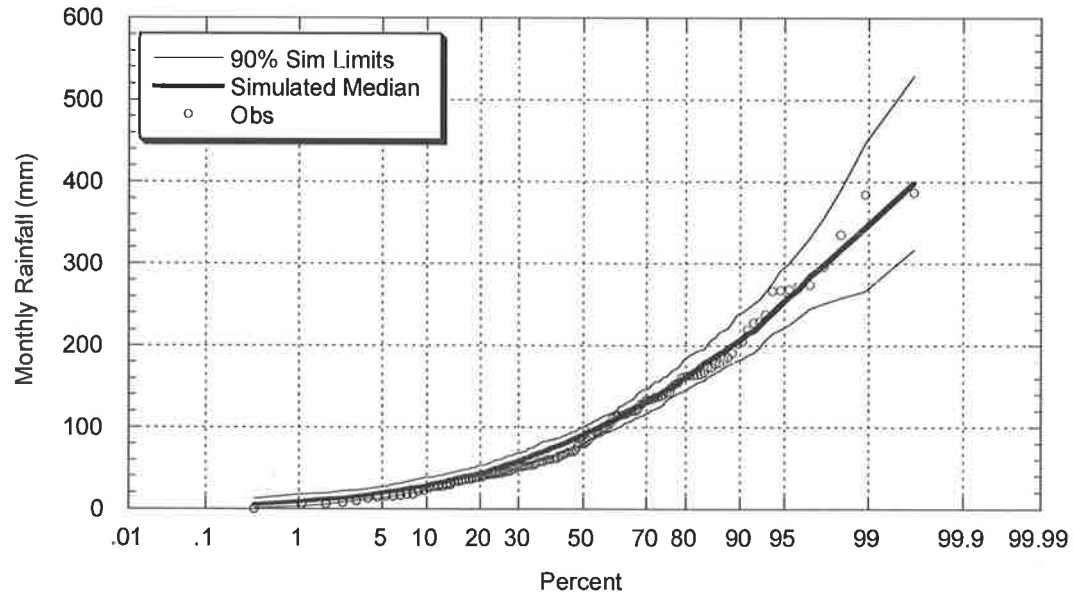

Figure B.5.10: Comparison between Observed and Simulated January Rainfall (Sydney)

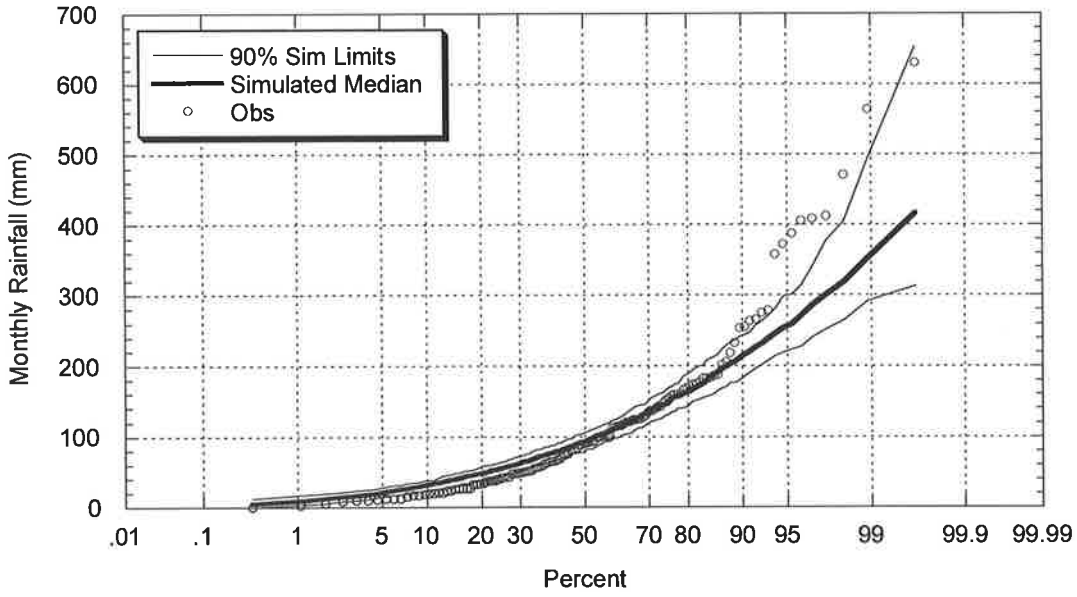

Figure B.5.11: Comparison between Observed and Simulated February Rainfall (Sydney)

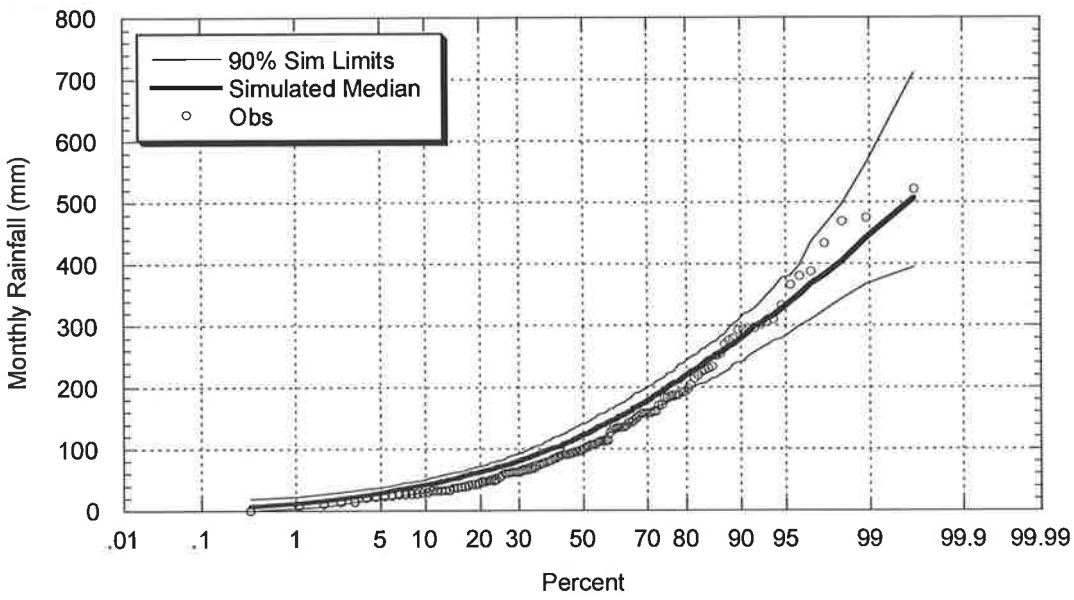

Figure B.5.12: Comparison between Observed and Simulated March Rainfall (Sydney)

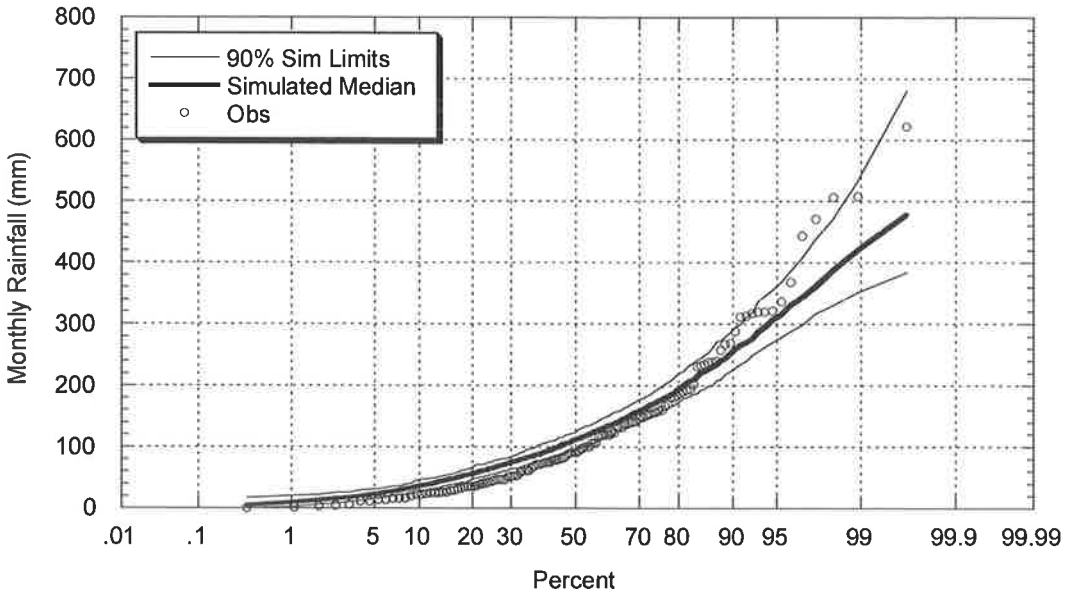

Figure B.5.13: Comparison between Observed and Simulated April Rainfall (Sydney)

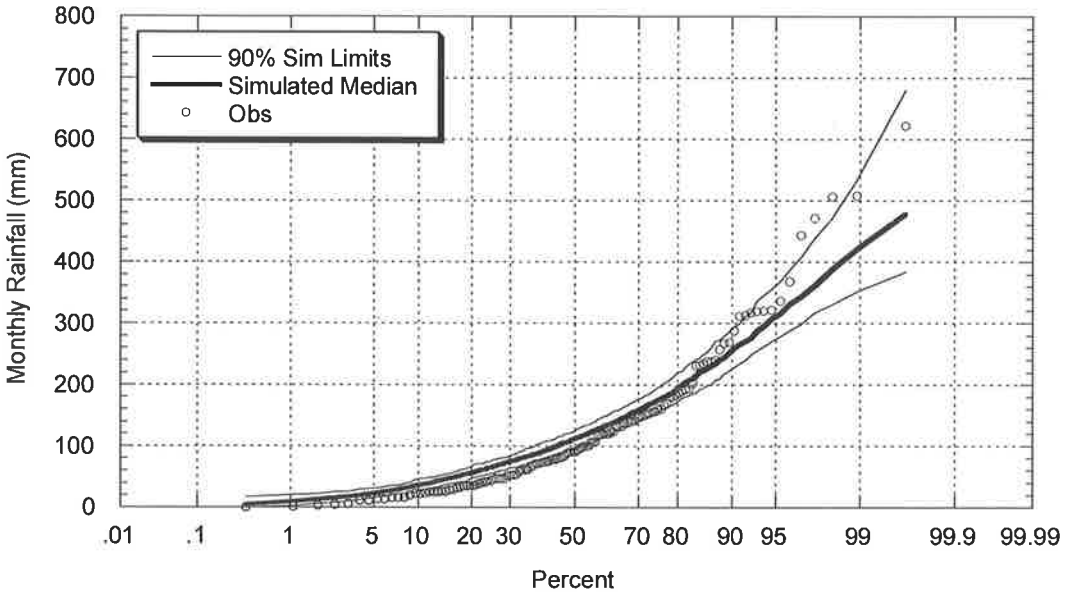

Figure B.5.14: Comparison between Observed and Simulated May Rainfall (Sydney)

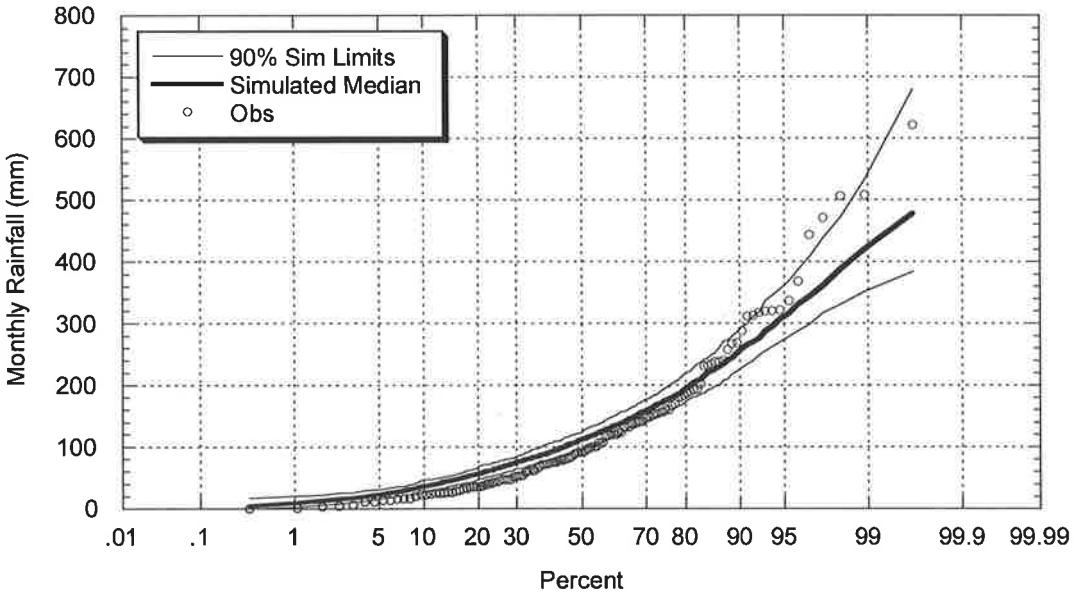

Figure B.5.15: Comparison between Observed and Simulated June Rainfall (Sydney)

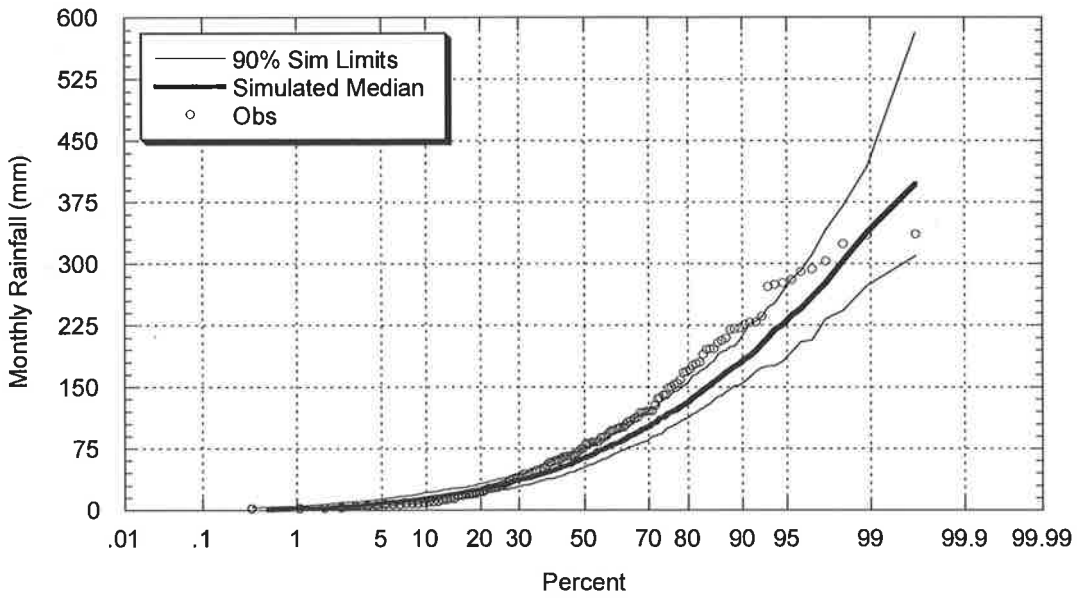

Figure B.5.16: Comparison between Observed and Simulated July Rainfall (Sydney)

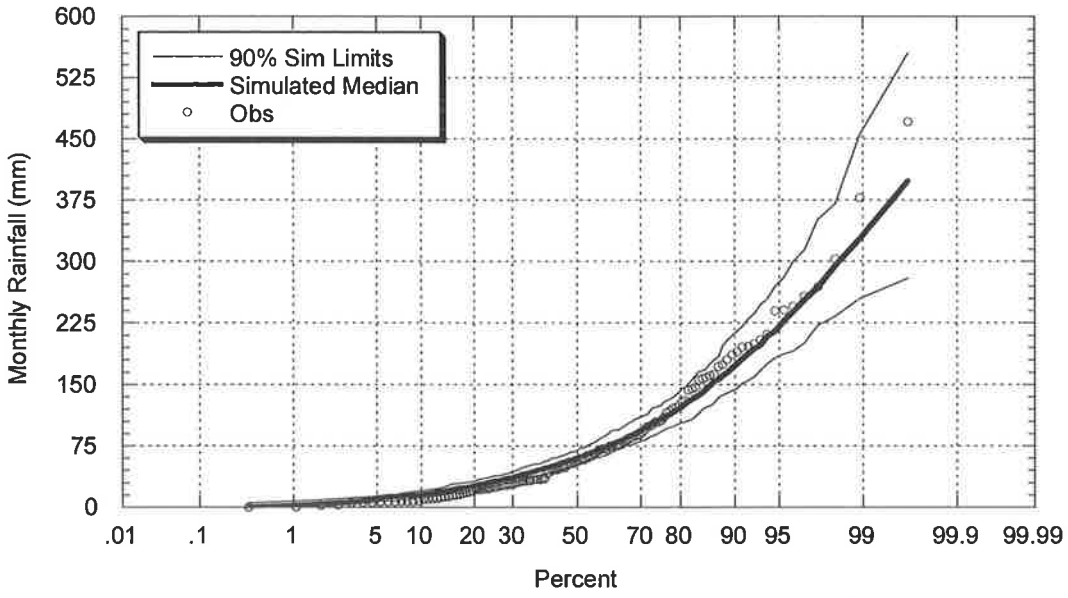

Figure B.5.17: Comparison between Observed and Simulated August Rainfall (Sydney)

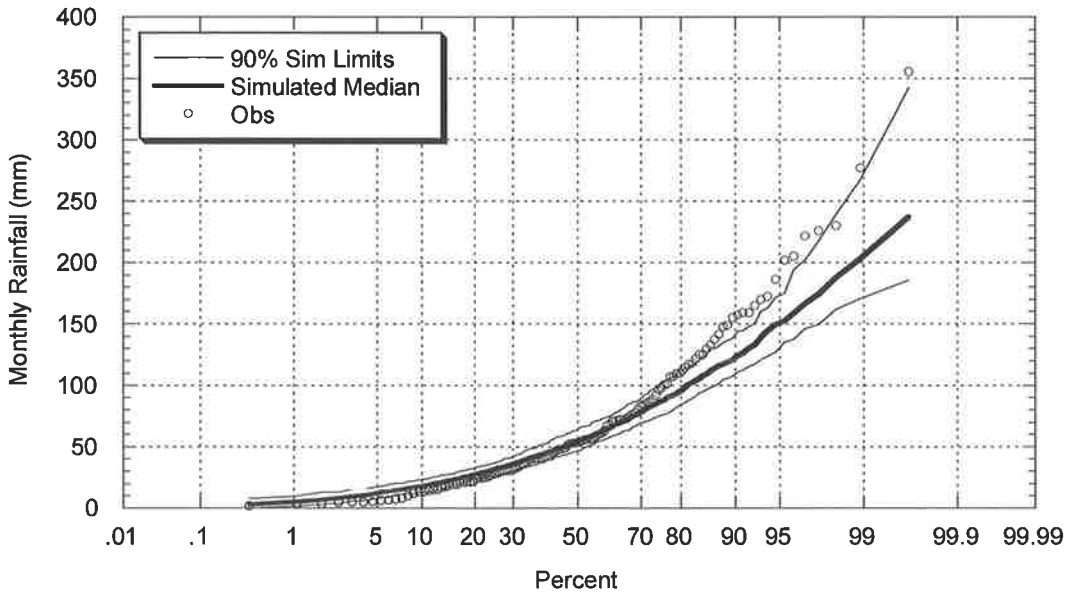

Figure B.5.18: Comparison between Observed and Simulated September Rainfall (Sydney)

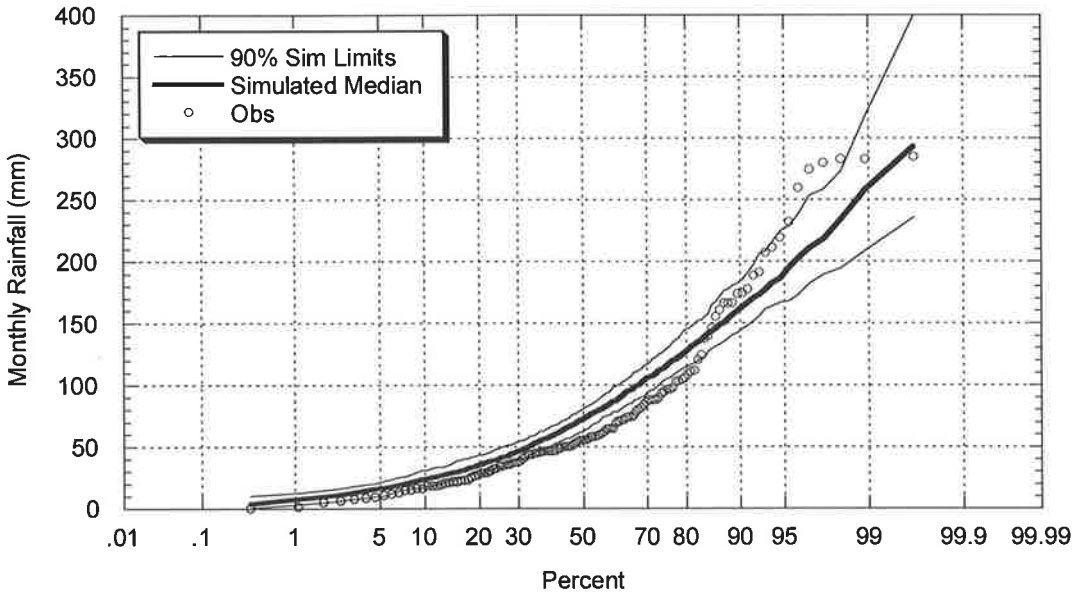

Figure B.5.19: Comparison between Observed and Simulated October Rainfall (Sydney)

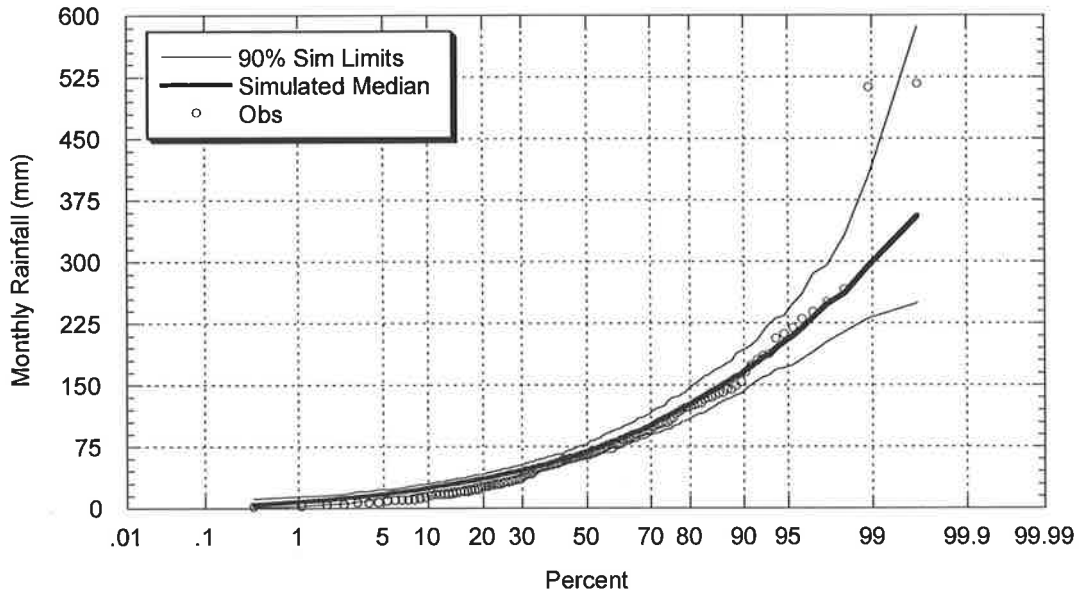

Figure B.5.20: Comparison between Observed and Simulated November Rainfall (Sydney)

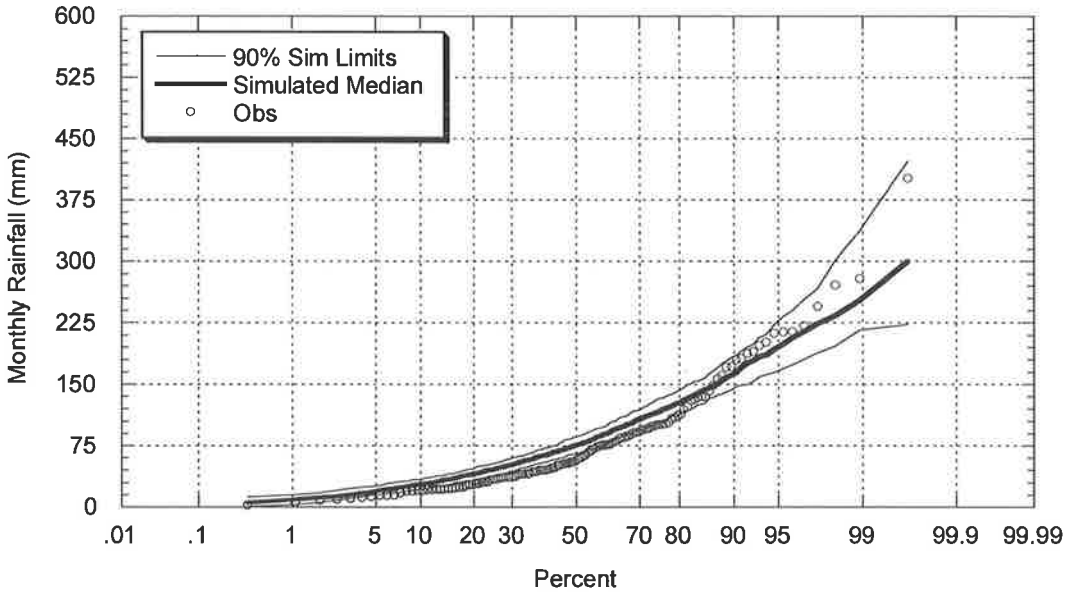

Figure B.5.21: Comparison between Observed and Simulated December Rainfall (Sydney)

B.5.4 Simulated and Observed Annual Intensity – Frequency - Duration Curves

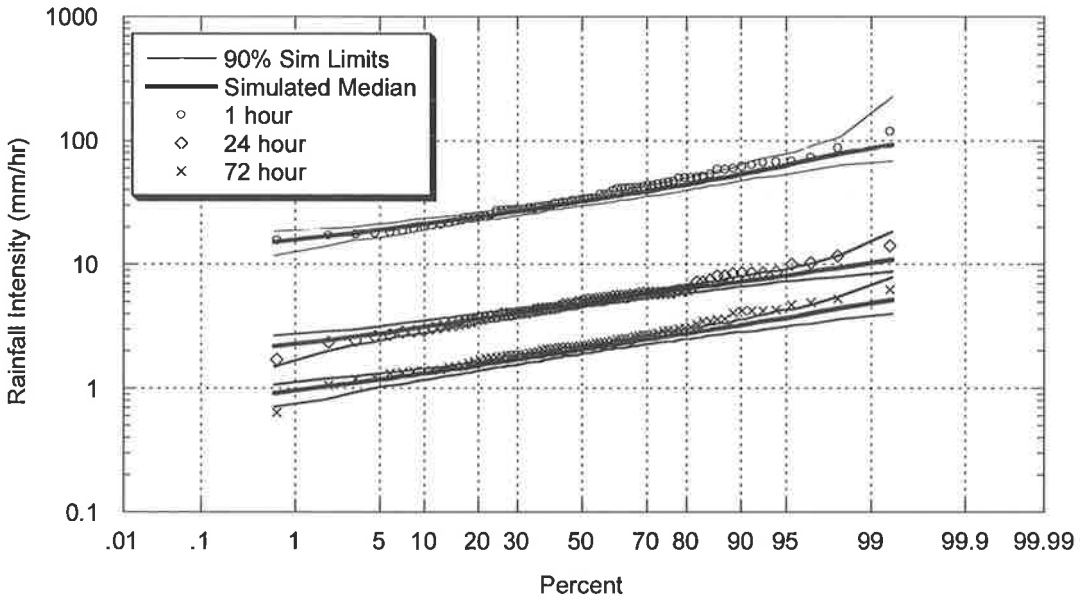

Figure B.5.22: Comparison between Observed and Simulated Annual Intensity Frequency Duration Relationship (Sydney)

APPENDIX C

Regionalisation with a Short Pluviograph Record - Results

C.1 Master – Adelaide, South Australia (BOM# 23034)

C.1.1 Target – Williamstown, South Australia (BOM# 23763)

C.1.1.1 Simulated and Observed Storm Event Characteristics

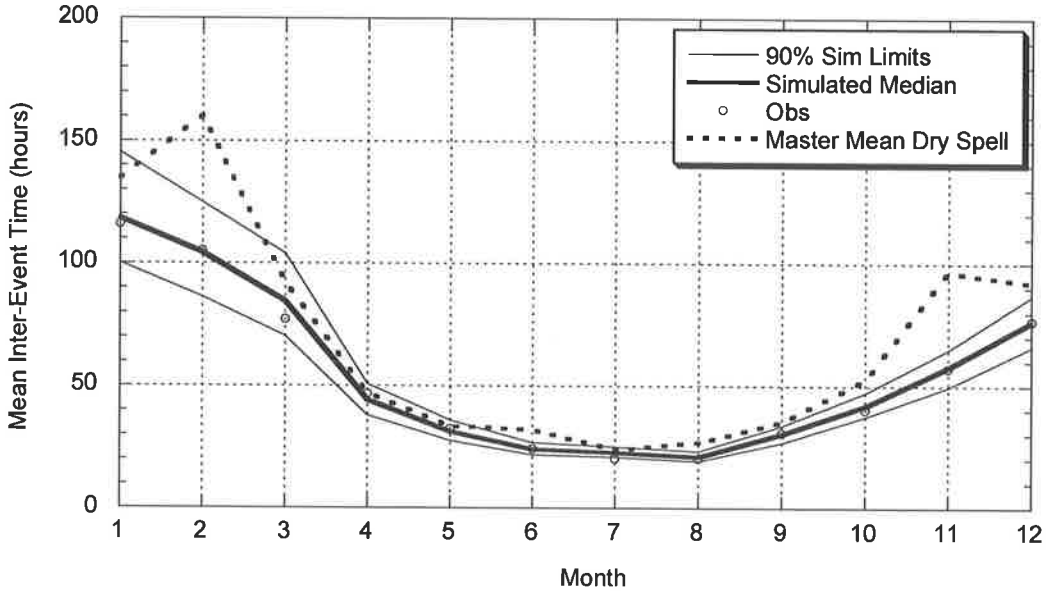

Figure C.1.1: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Adelaide Airport; Target – Williamstown)

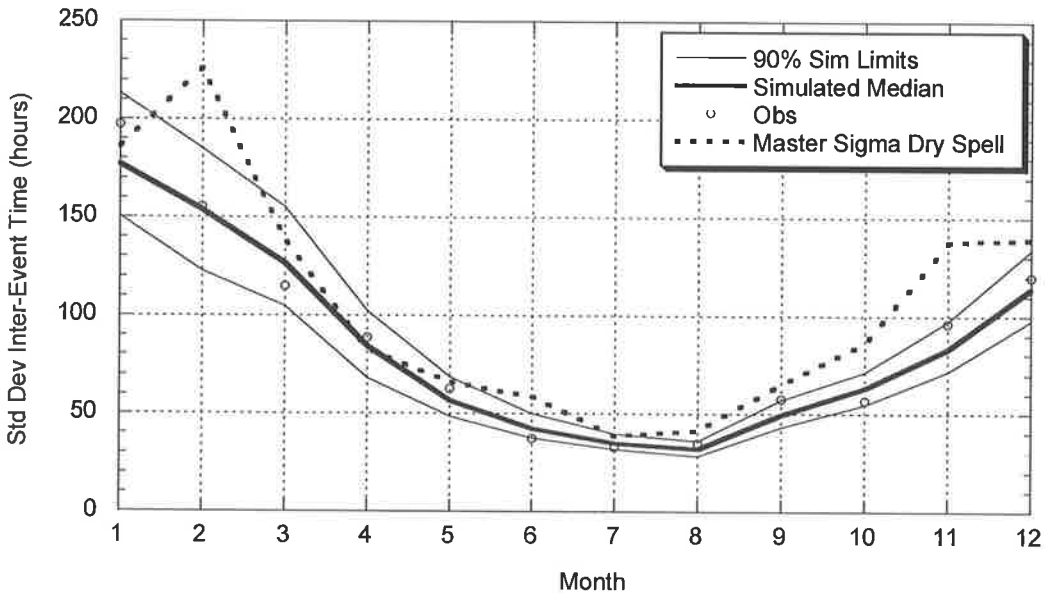

Figure C.1.2: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Adelaide Airport; Target – Williamstown)

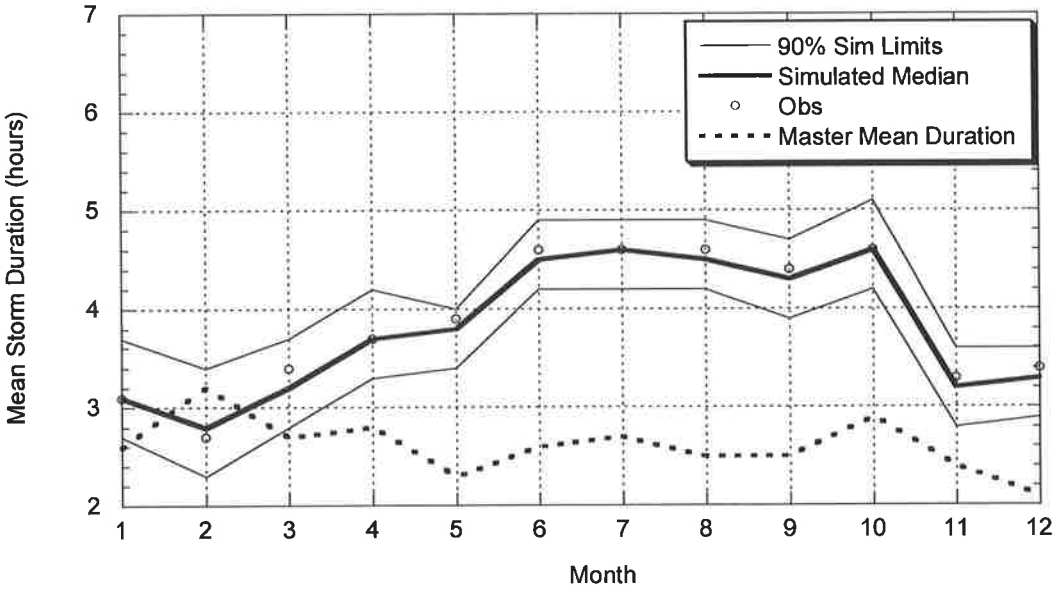

Figure C.1.3: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Adelaide Airport; Target – Williamstown)

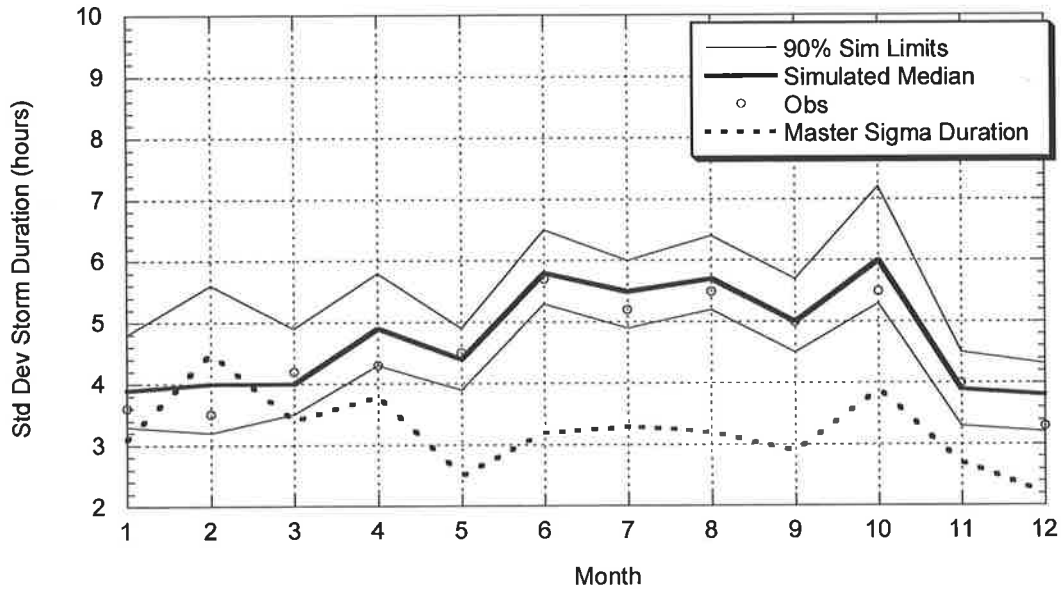

Figure C.1.4: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations (Master – Adelaide Airport; Target – Williamstown)

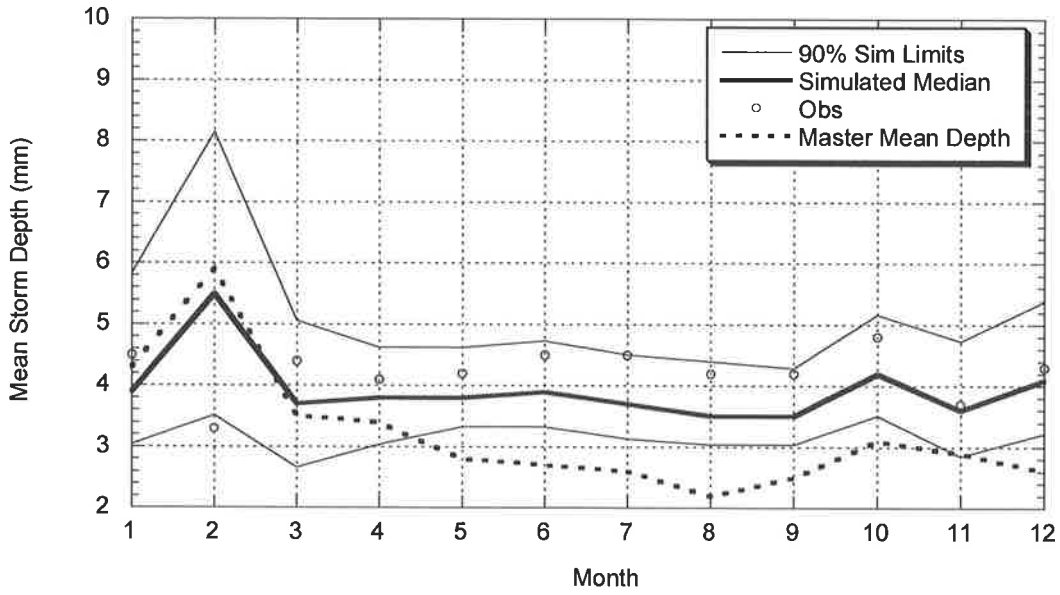

Figure C.1.5: Comparison between Observed and Target Simulated Average of Event Depths (Master – Adelaide Airport; Target – Williamstown)

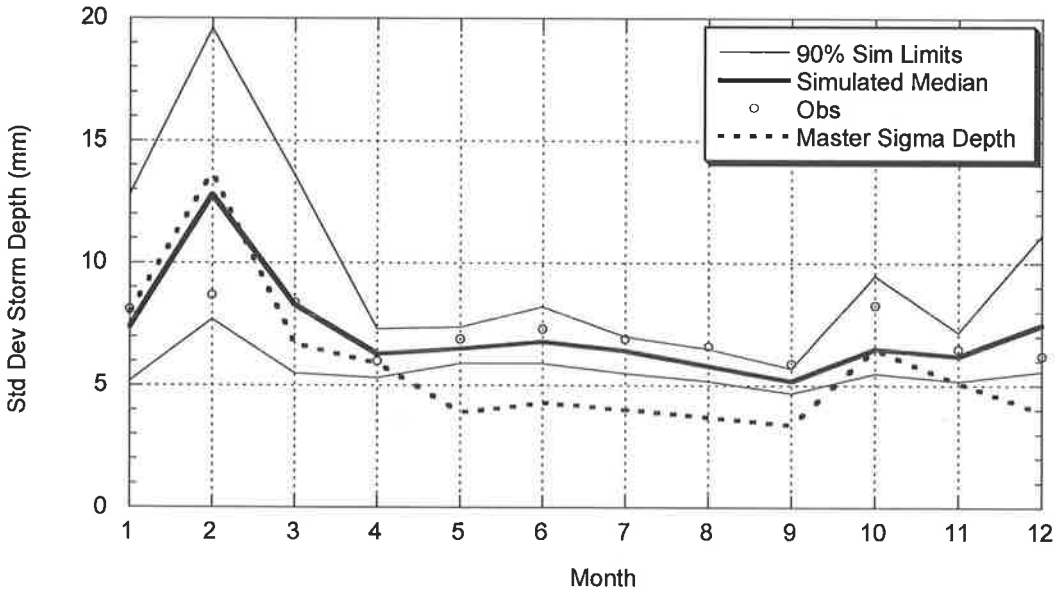

Figure C.1.6: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Adelaide Airport; Target – Williamstown)

C.1.1.2 Simulated and Observed Daily Statistics

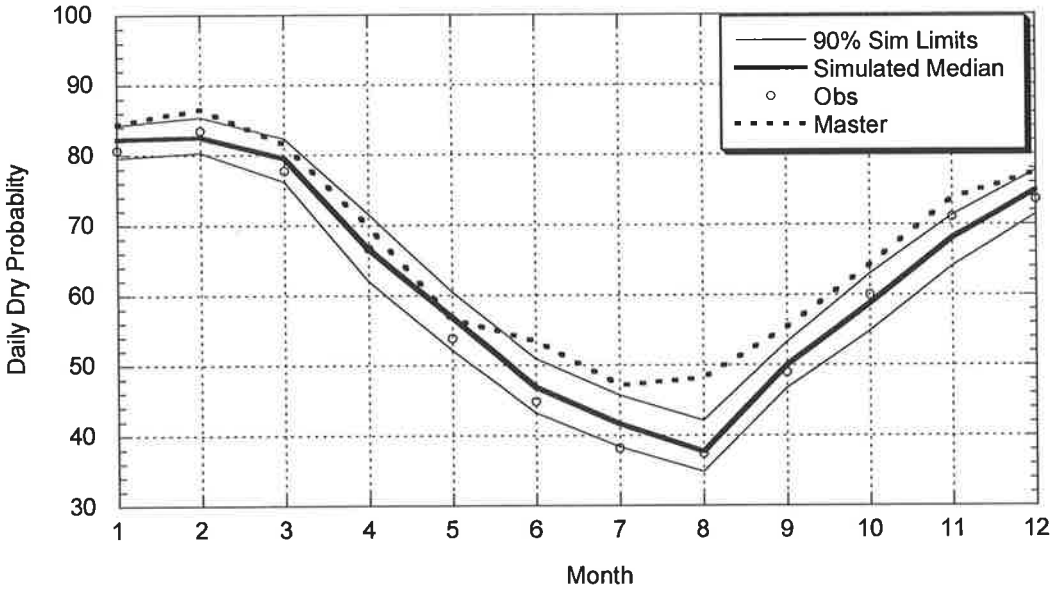

Figure C.1.7: Comparison between Observed and Target Simulated Daily Dry Probabilities (Master – Adelaide Airport; Target – Williamstown)

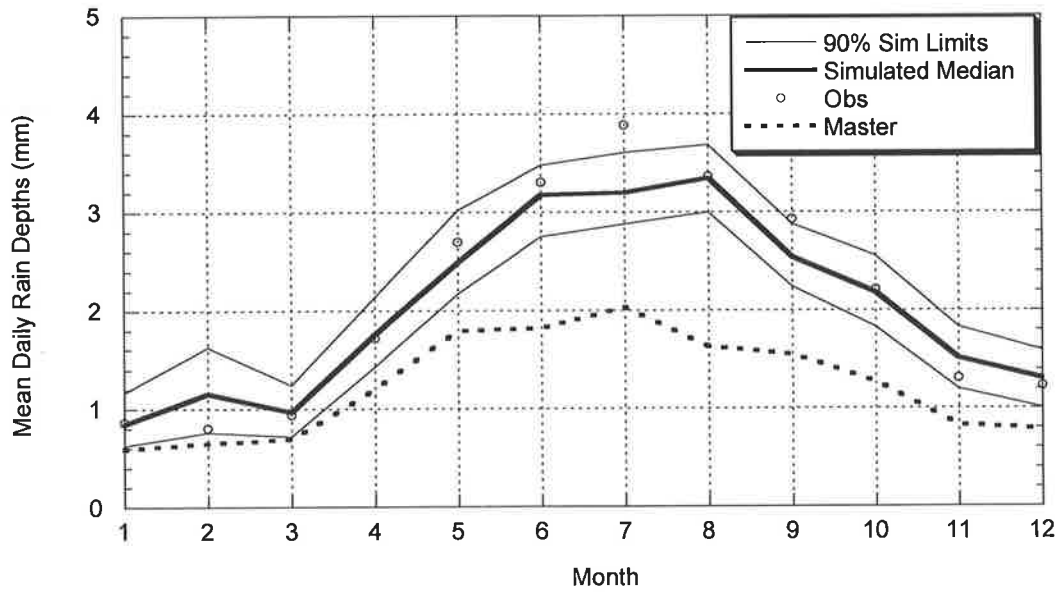

Figure C.1.8: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Adelaide Airport; Target – Williamstown)

C.1.1.3 Simulated and Observed Annual and Monthly Rainfall

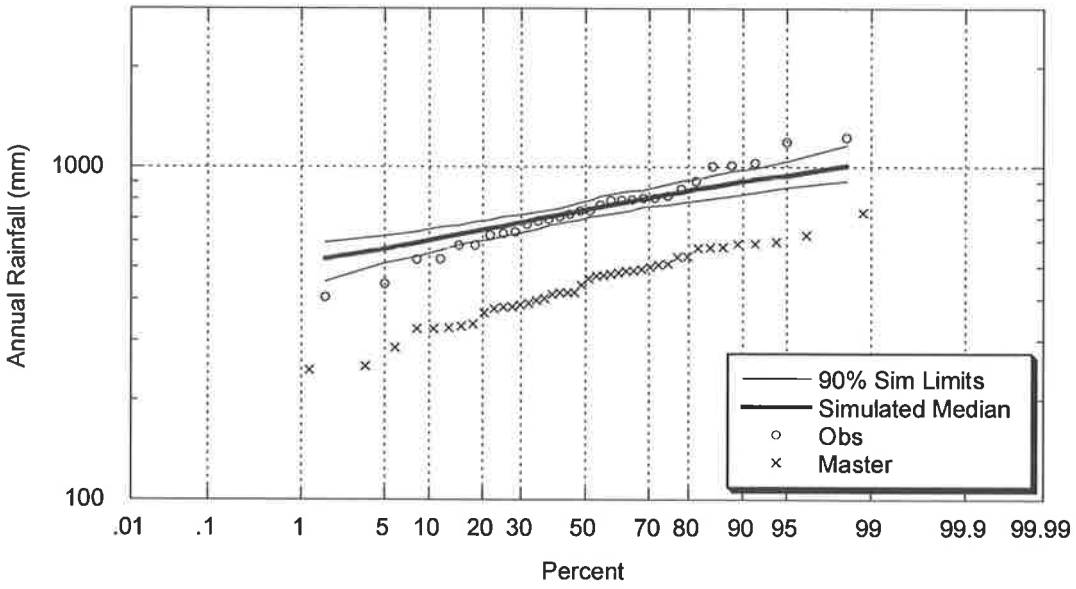

Figure C.1.9: Comparison between Observed and Target Simulated Annual Rainfall (Master – Adelaide Airport; Target – Williamstown)

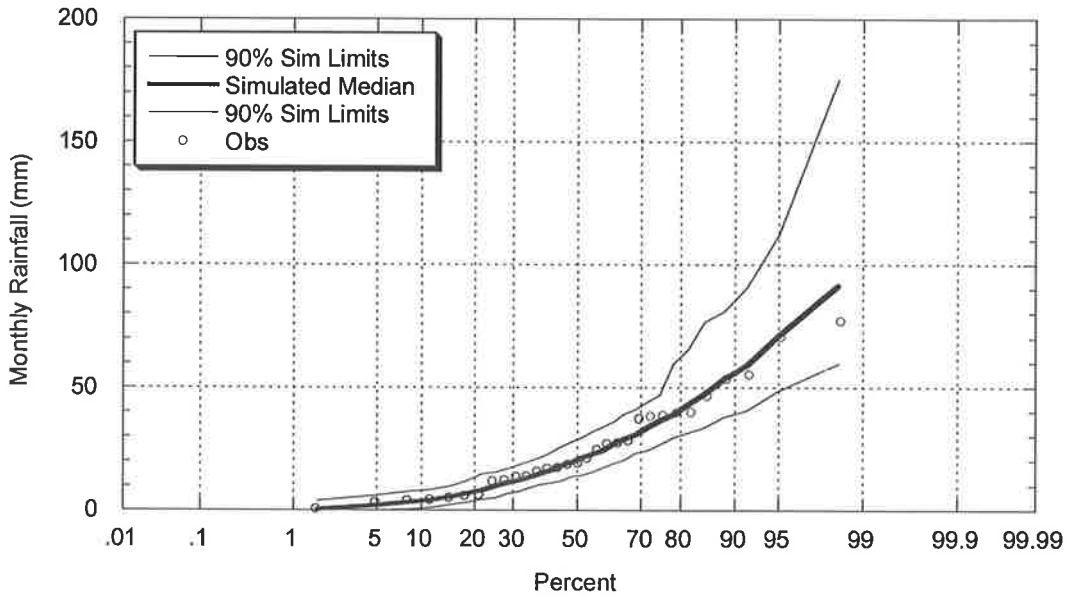

Figure C.1.10: Comparison between Observed and Target Simulated January Rainfall (Master – Adelaide Airport; Target – Williamstown)

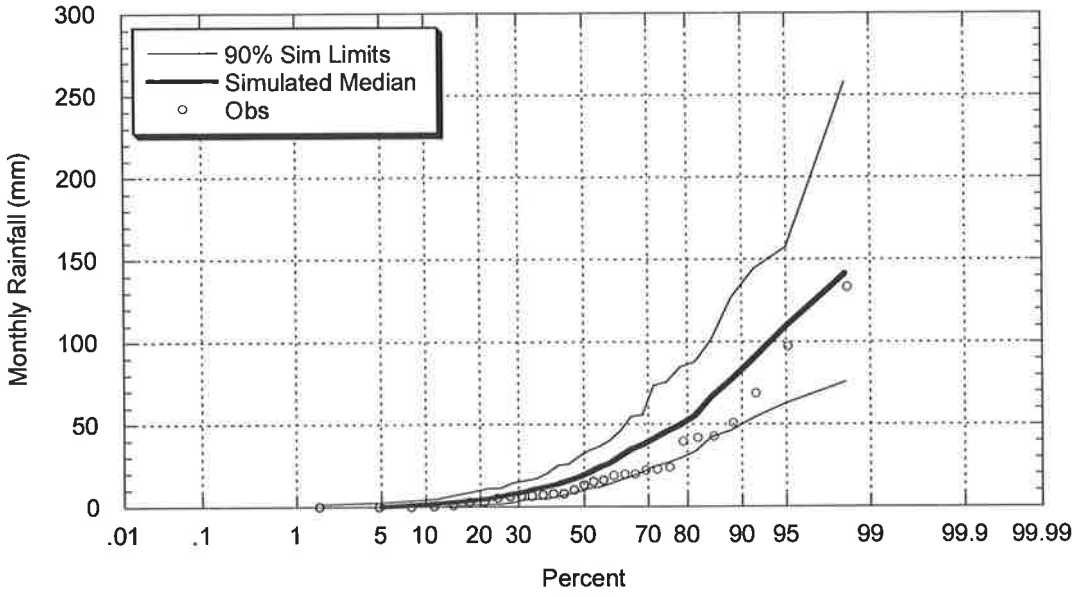

Figure C.1.11: Comparison between Observed and Target Simulated February Rainfall
(Master – Adelaide Airport; Target – Williamstown)

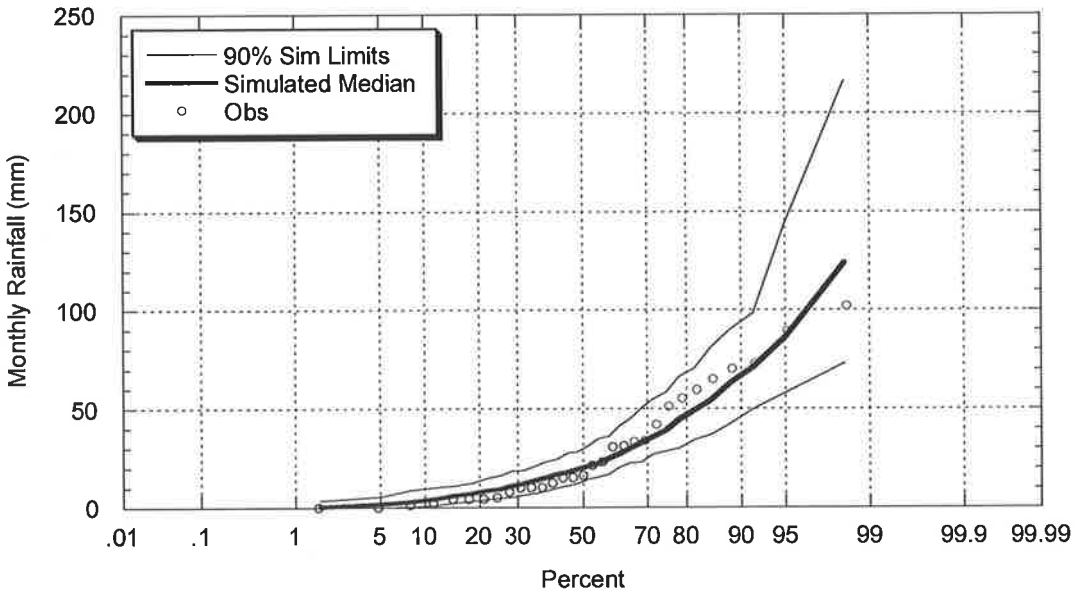

Figure C.1.12: Comparison between Observed and Target Simulated March Rainfall
(Master – Adelaide Airport; Target – Williamstown)

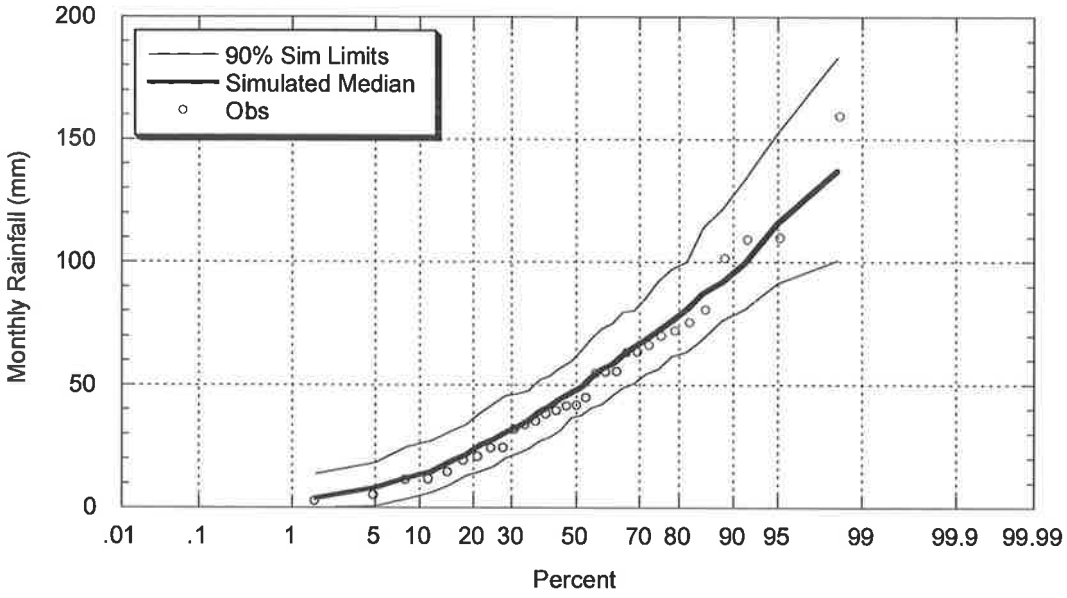

Figure C.1.13: Comparison between Observed and Target Simulated April Rainfall (Master – Adelaide Airport; Target – Williamstown)

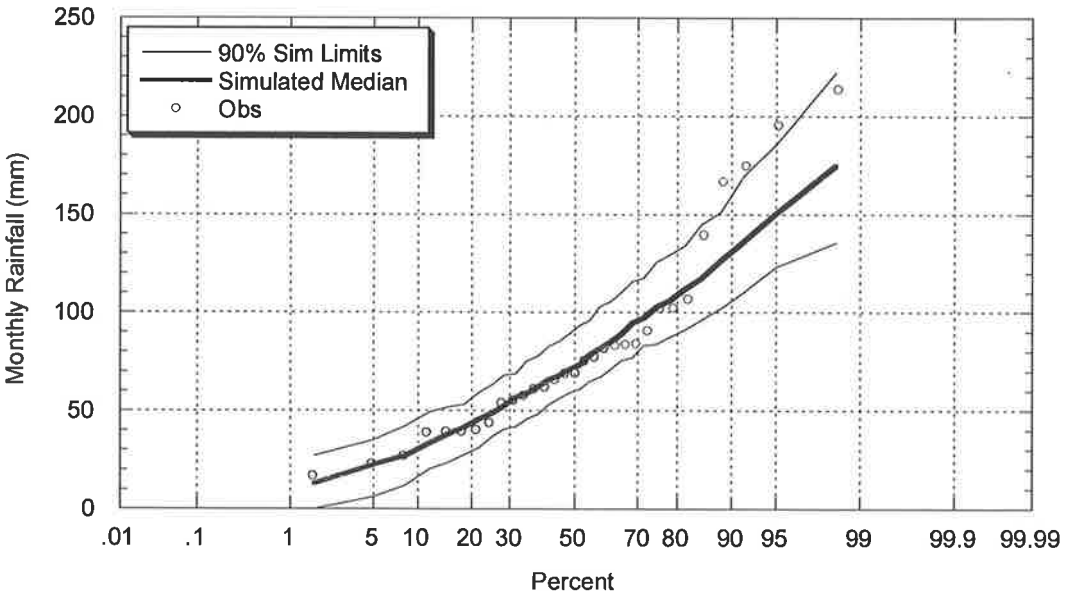

Figure C.1.14: Comparison between Observed and Target Simulated May Rainfall (Master – Adelaide Airport; Target – Williamstown)

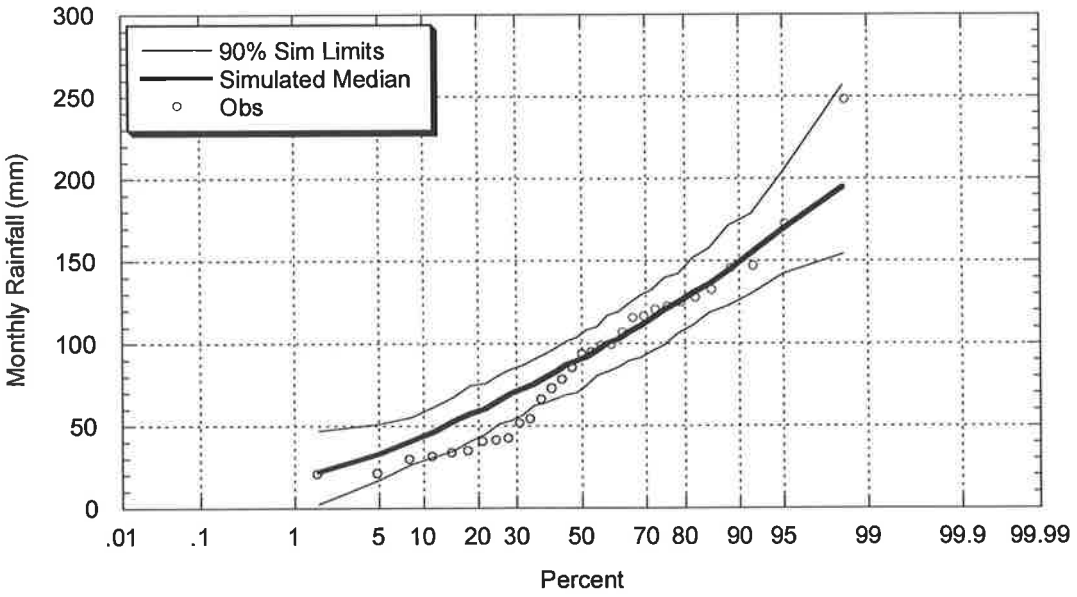

Figure C.1.15: Comparison between Observed and Target Simulated June Rainfall
(Master – Adelaide Airport; Target – Williamstown)

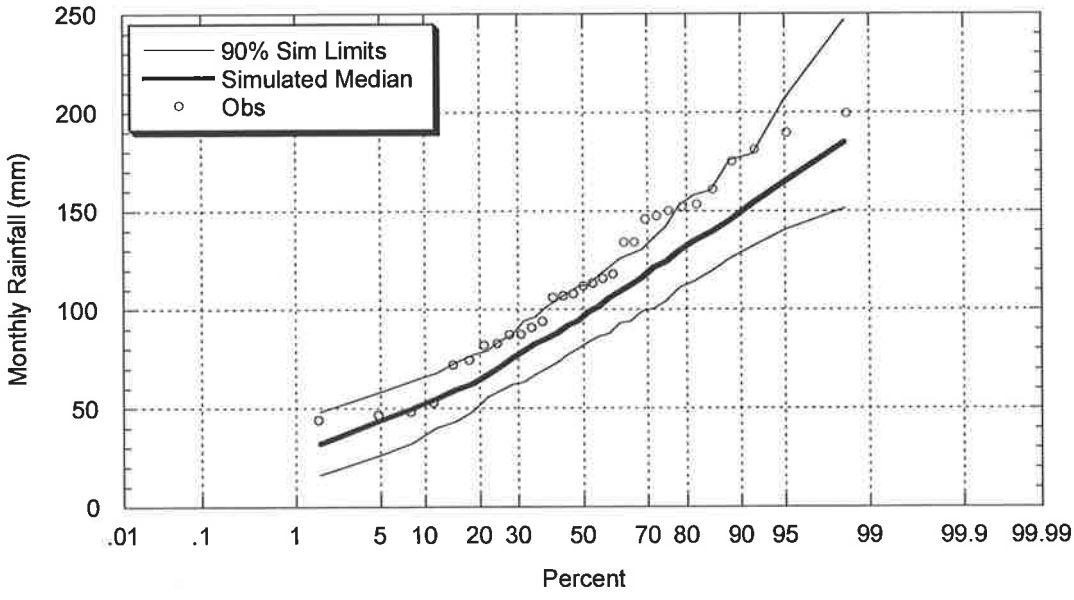

Figure C.1.16: Comparison between Observed and Target Simulated July Rainfall
(Master – Adelaide Airport; Target – Williamstown)

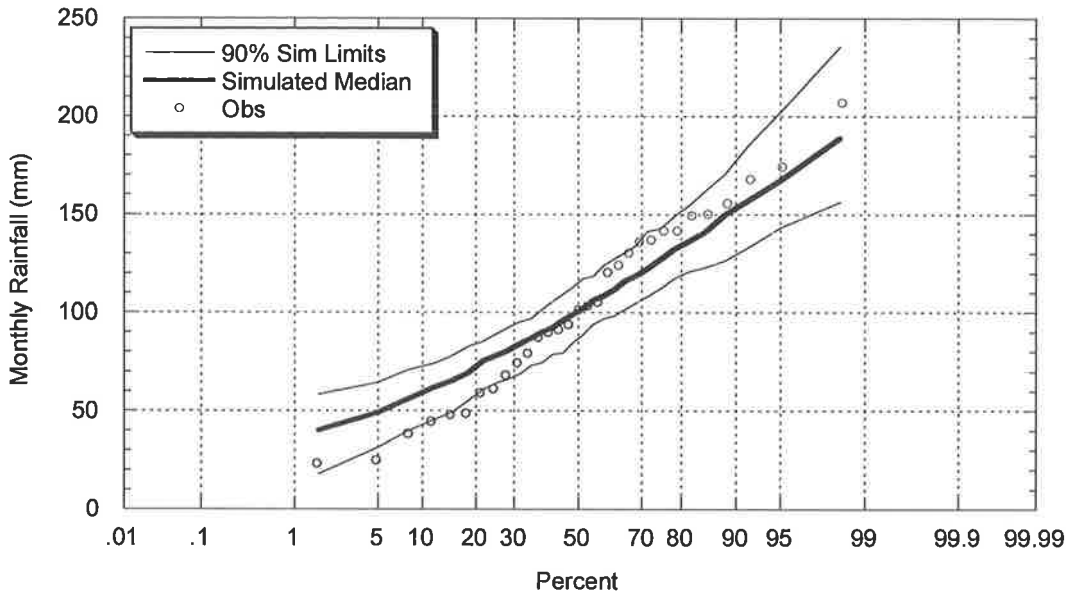

Figure C.1.17: Comparison between Observed and Target Simulated August Rainfall (Master – Adelaide Airport; Target – Williamstown)

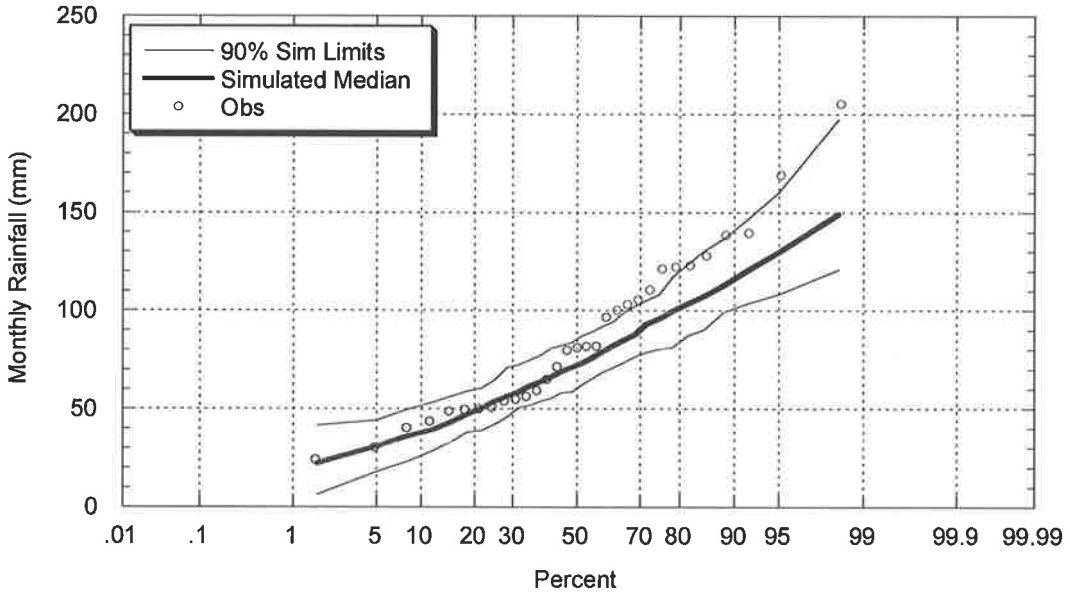

Figure C.1.18: Comparison between Observed and Target Simulated September Rainfall (Master – Adelaide Airport; Target – Williamstown)

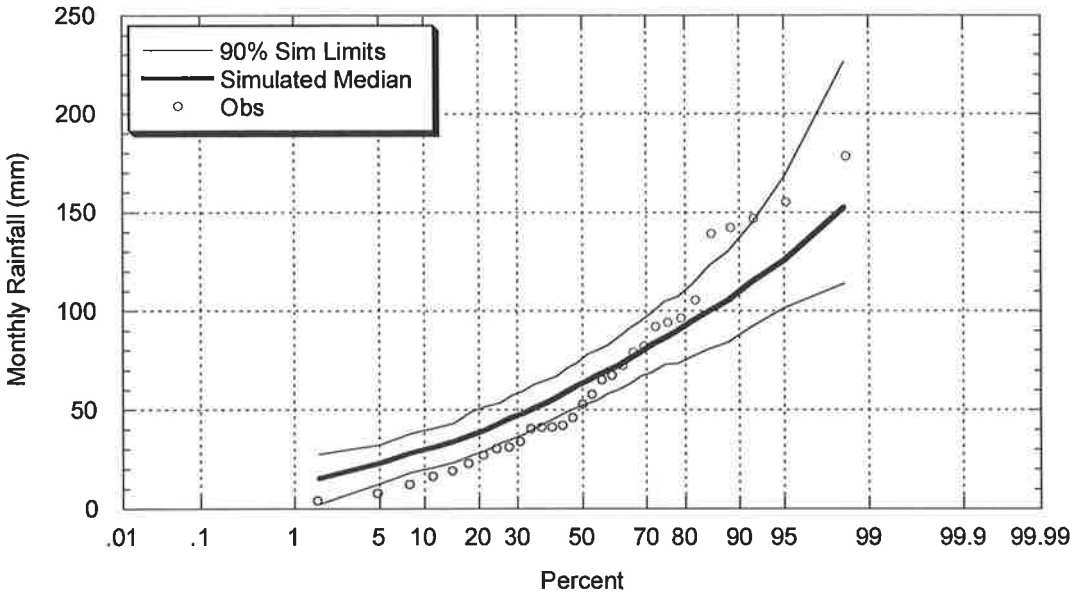

Figure C.1.19: Comparison between Observed and Target Simulated October Rainfall (Master – Adelaide Airport; Target – Williamstown)

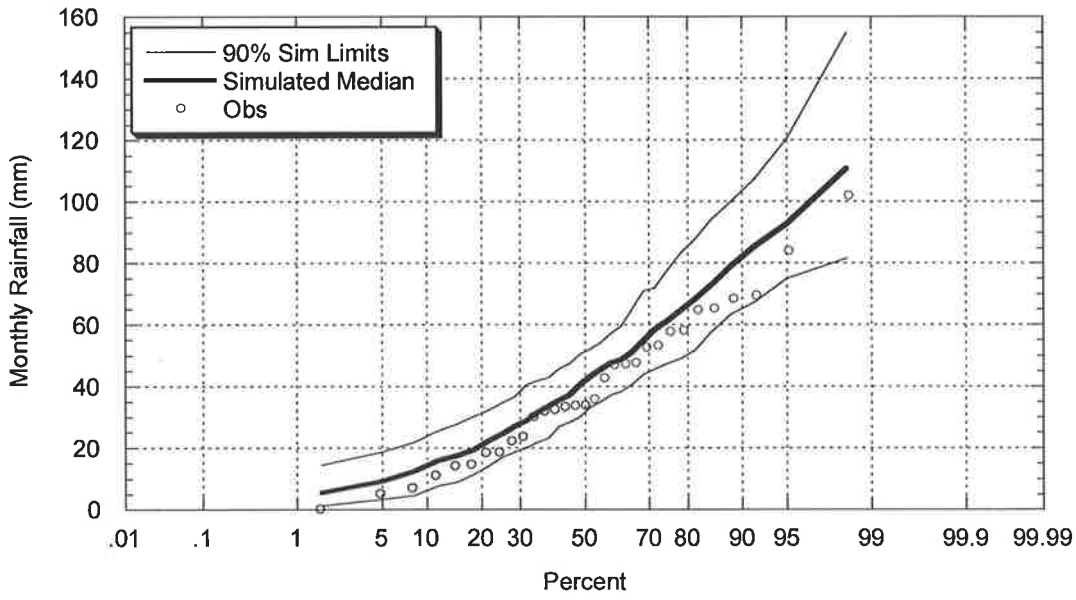

Figure C.1.20: Comparison between Observed and Target Simulated November Rainfall (Master – Adelaide Airport; Target – Williamstown)

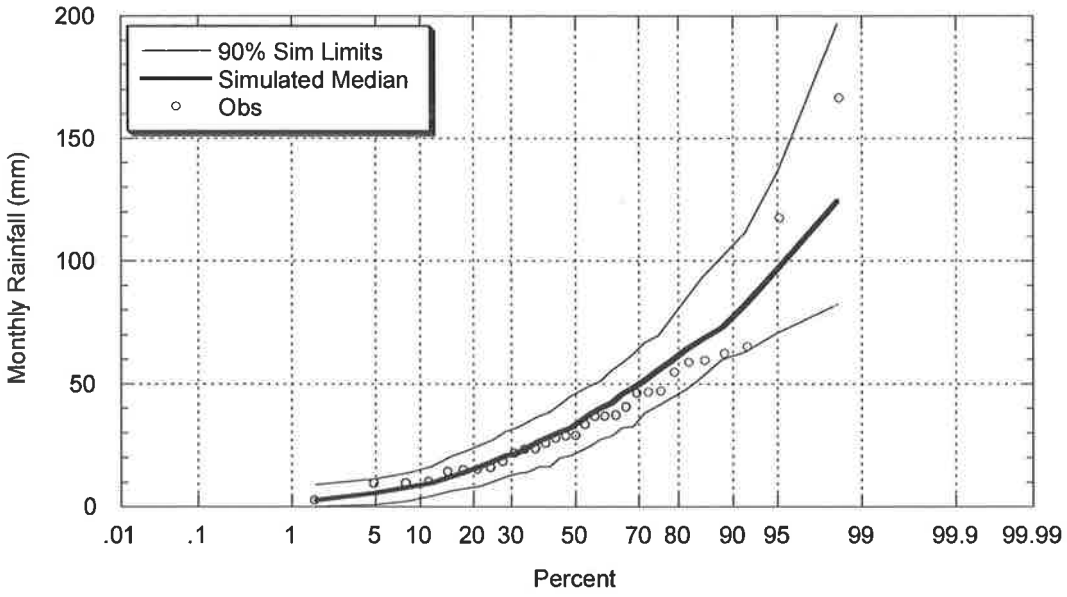

Figure C.1.21: Comparison between Observed and Target Simulated December Rainfall (Master – Adelaide Airport; Target – Williamstown)

C.1.1.4 Simulated and Observed Annual IFD Curves

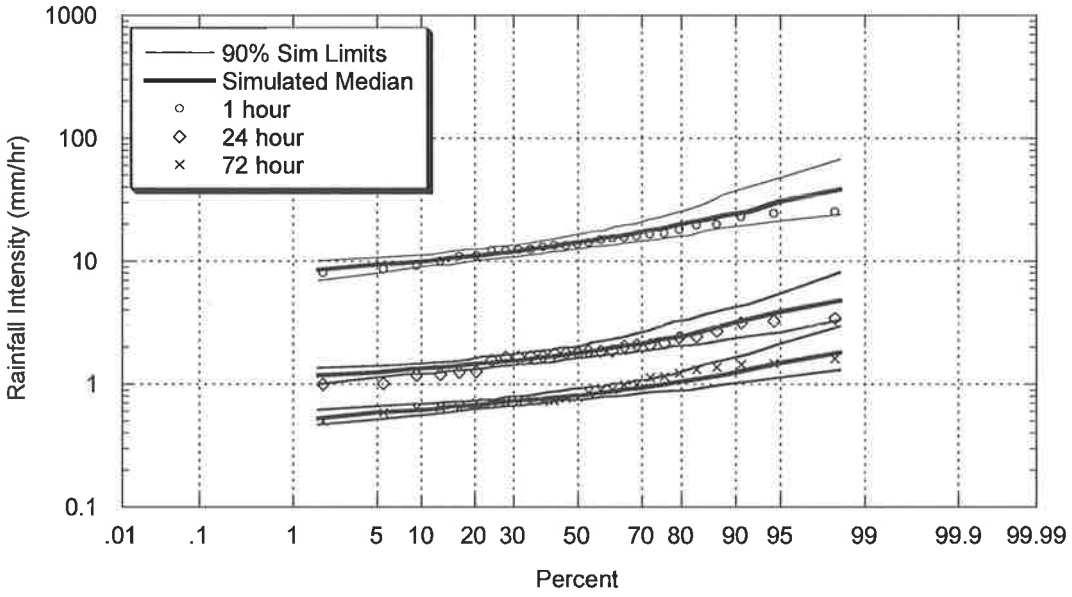

Figure C.1.22: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Adelaide Airport; Target – Williamstown)

C.1.2 Target – Stirling, South Australia (BOM# 23785)

C.1.2.1 Simulated and Observed Storm Event Characteristics

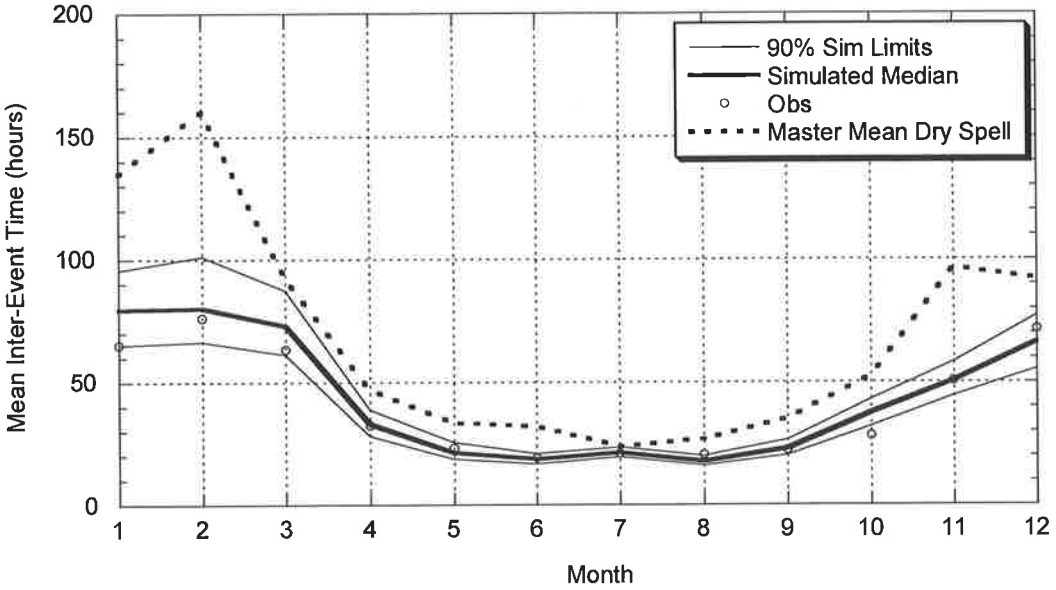

Figure C.1.23: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Adelaide Airport; Target – Stirling)

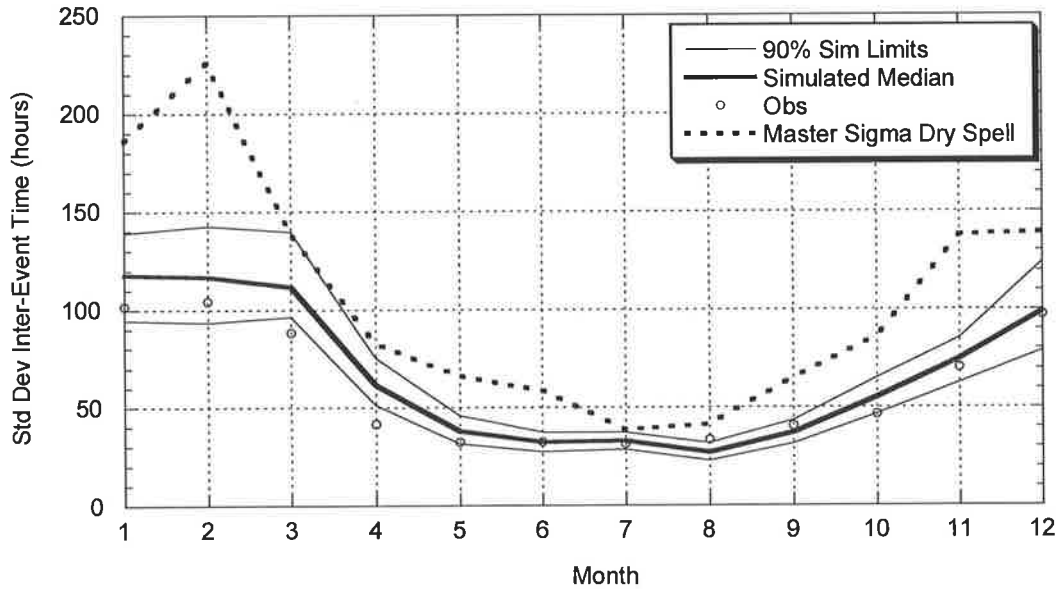

Figure C.1.24: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Adelaide Airport; Target – Stirling)

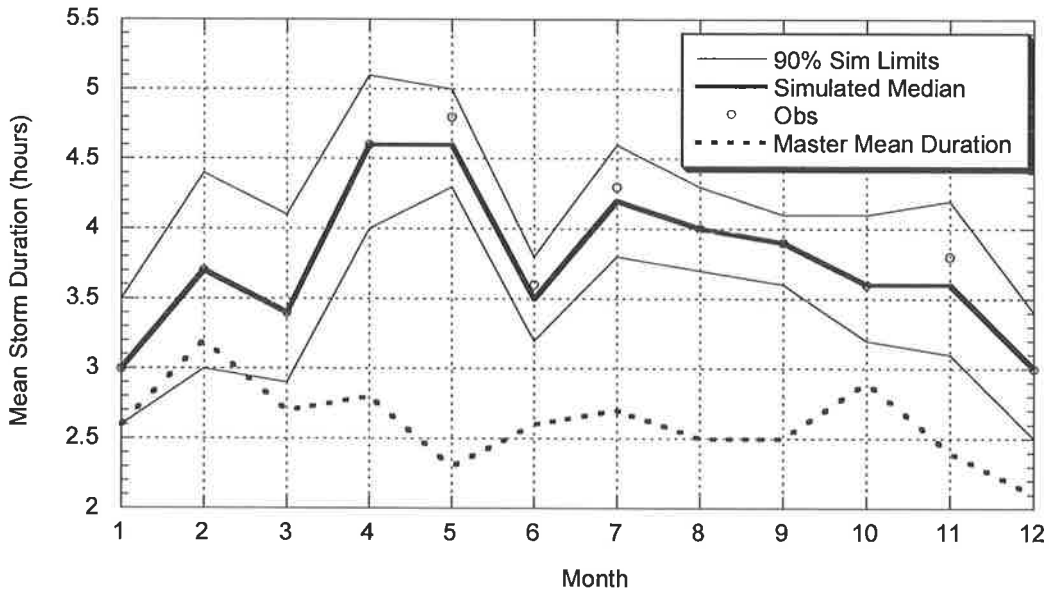

Figure C.1.25: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Adelaide Airport; Target – Stirling)

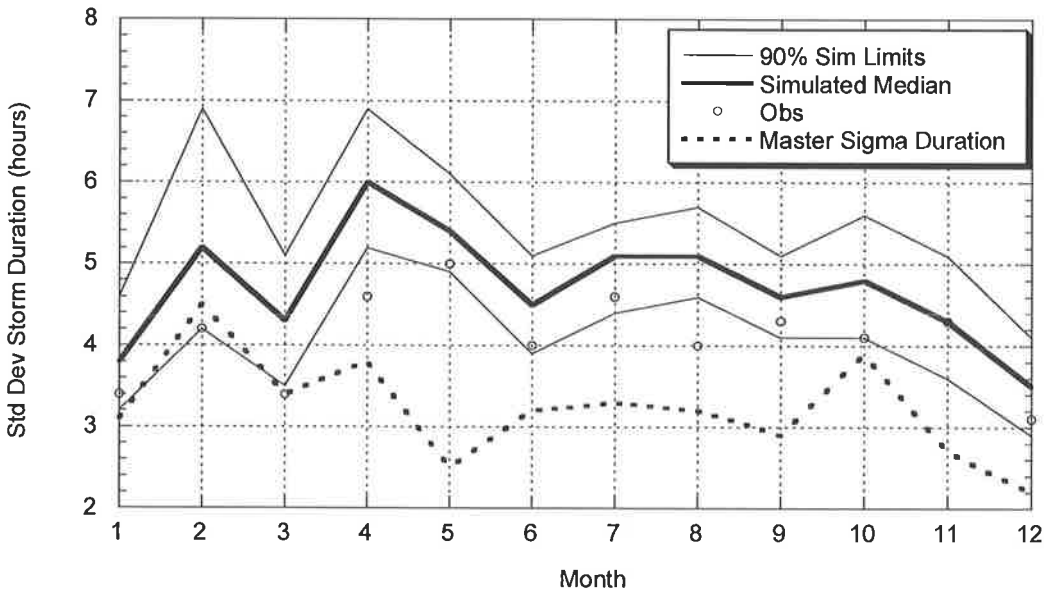

Figure C.1.26: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations (Master – Adelaide Airport; Target – Stirling)

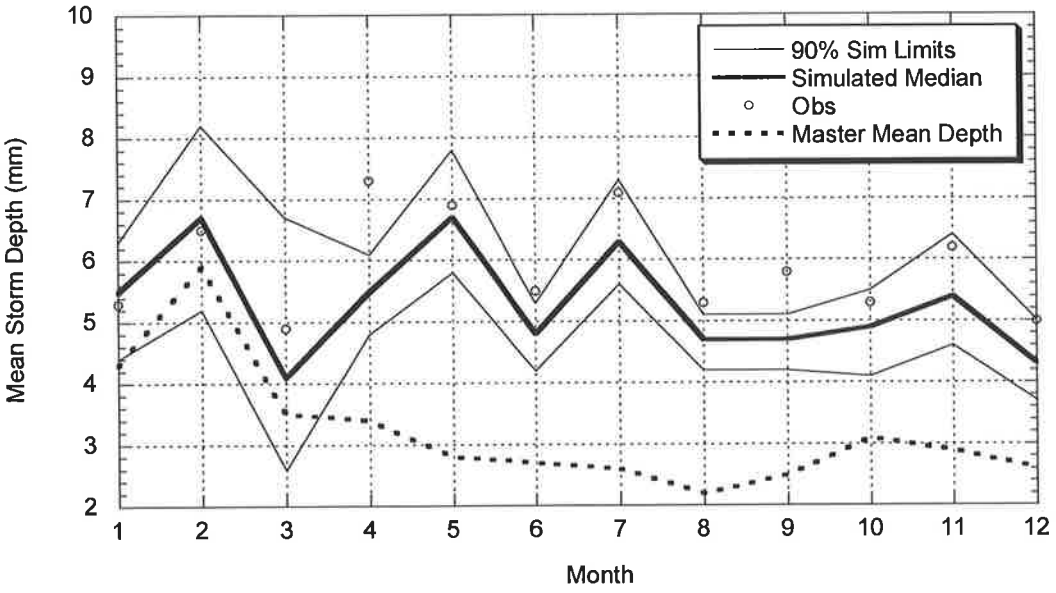

Figure C.1.27: Comparison between Observed and Target Simulated Average of Event Depths (Master – Adelaide Airport; Target – Stirling)

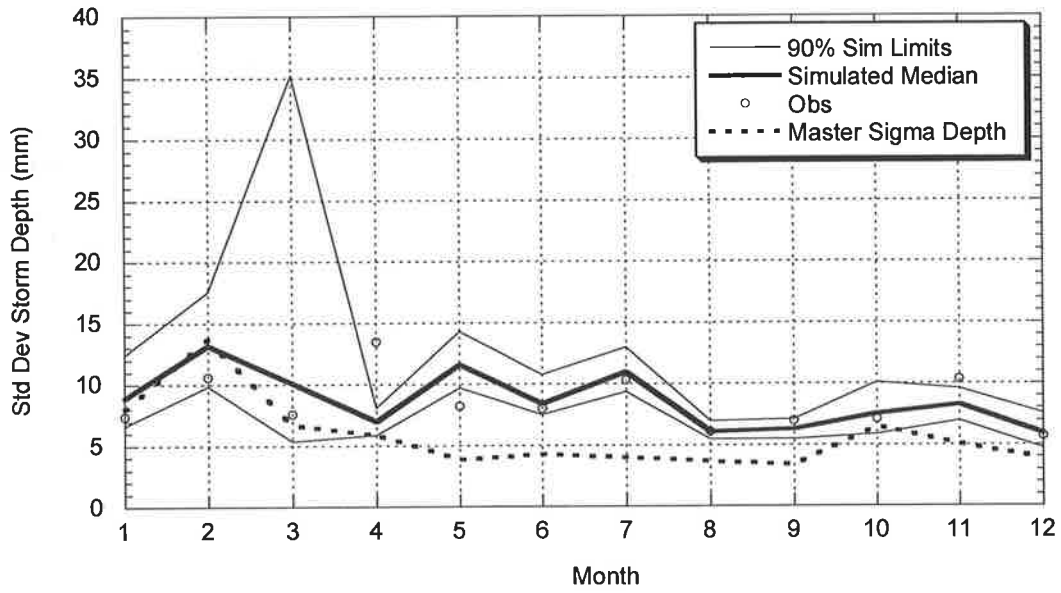

Figure C.1.28: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Adelaide Airport; Target – Stirling)

C.1.2.2 Simulated and Observed Daily Statistics

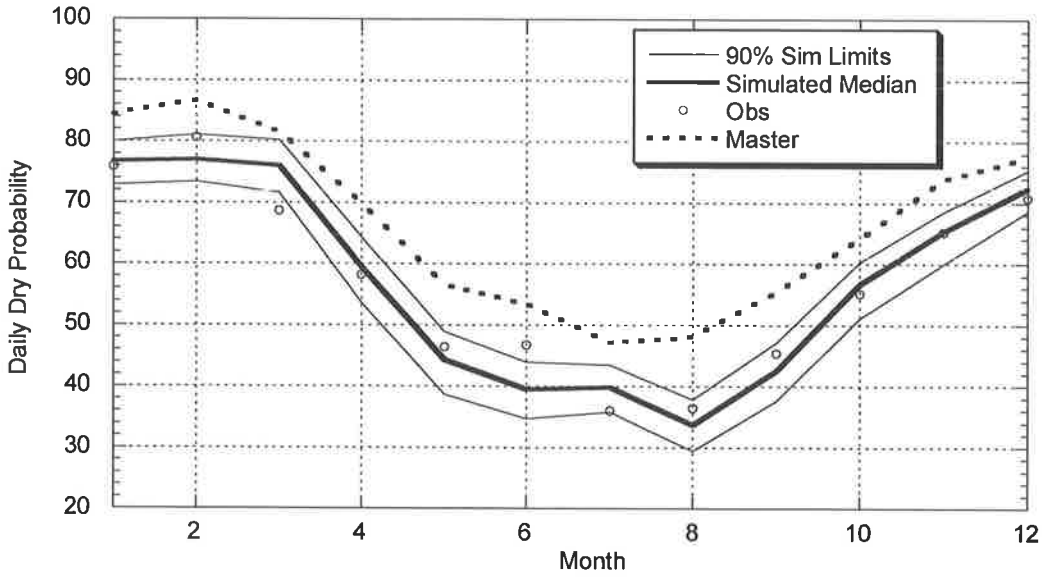

Figure C.1.29: Comparison between Observed and Target Simulated Daily Dry Probabilities (Master – Adelaide Airport; Target – Stirling)

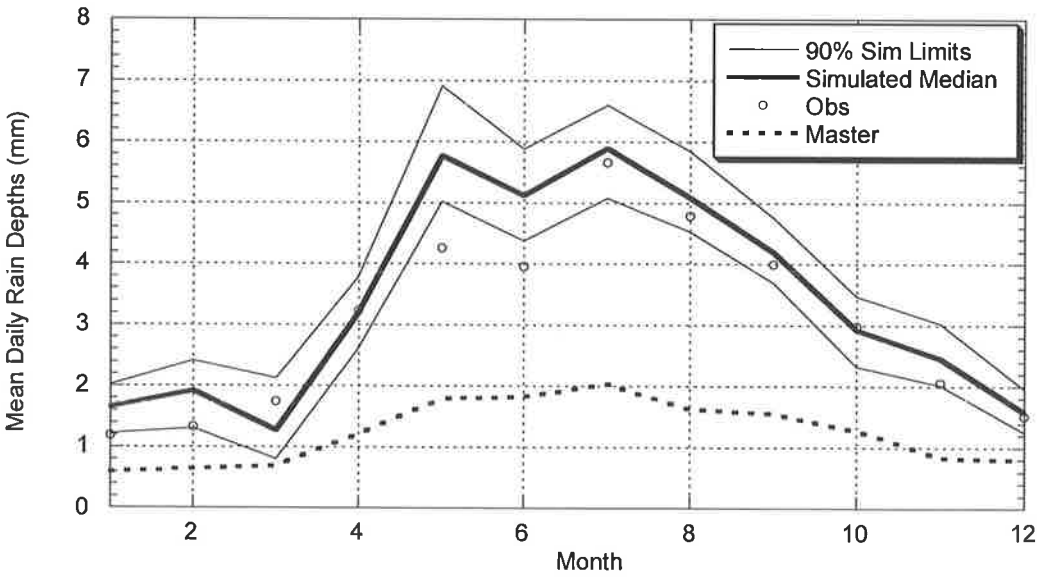

Figure C.1.30: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Adelaide Airport; Target – Stirling)

C.1.2.3 Simulated and Observed Annual and Monthly Rainfall

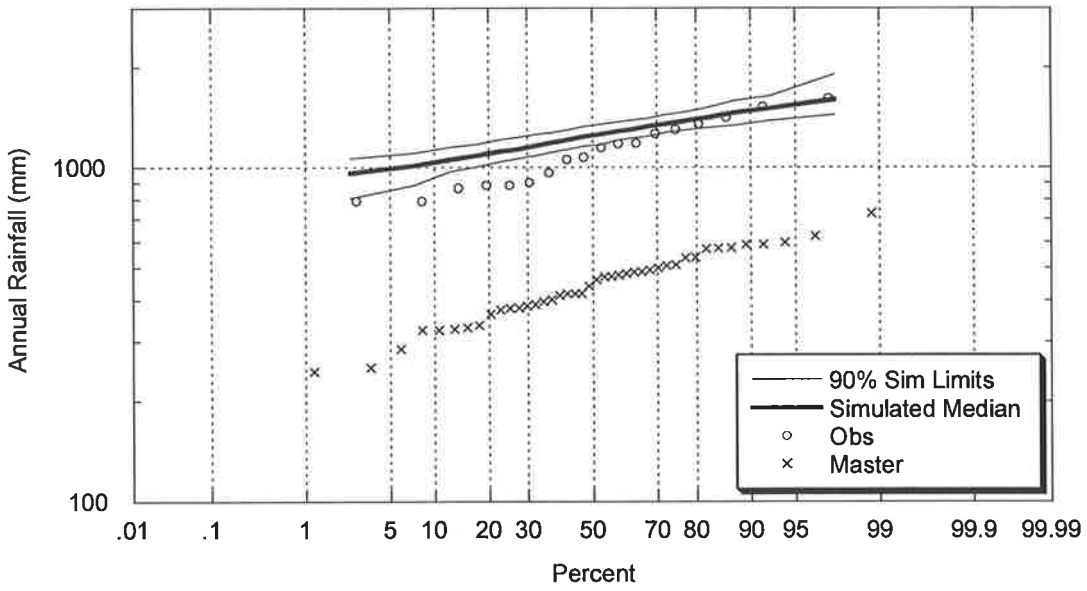

Figure C.1.31: Comparison between Observed and Target Simulated Annual Rainfall (Master – Adelaide Airport; Target – Stirling)

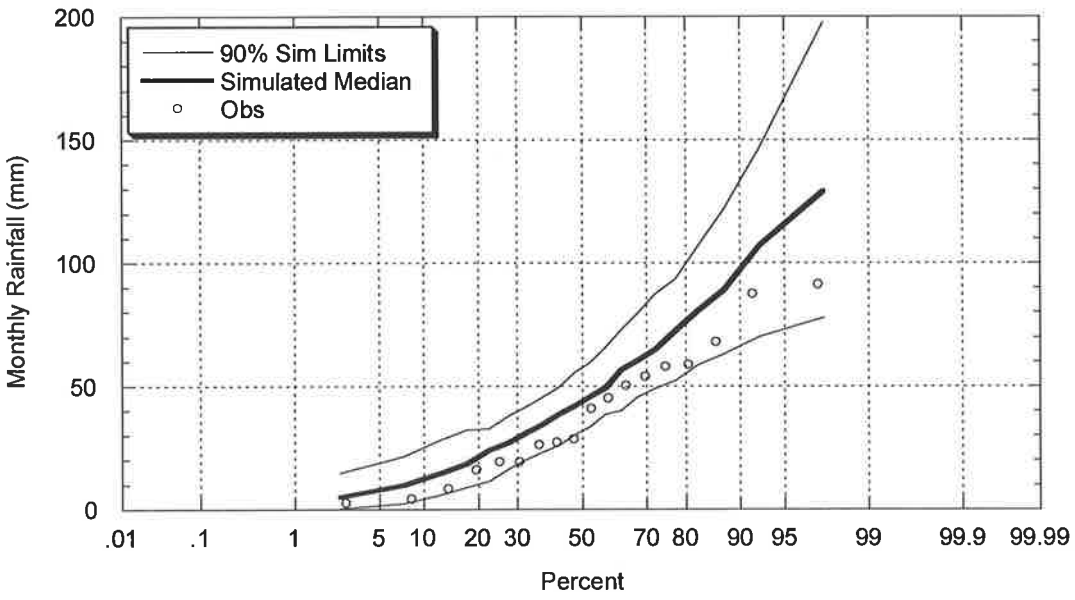

Figure C.1.32: Comparison between Observed and Target Simulated January Rainfall (Master – Adelaide Airport; Target – Stirling)

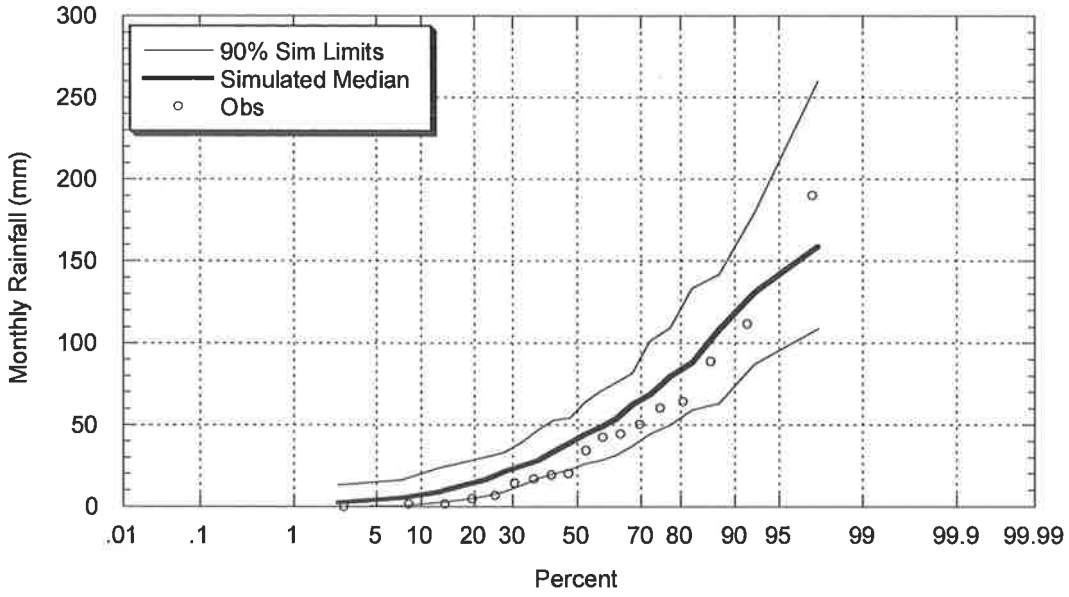

Figure C.1.33: Comparison between Observed and Target Simulated February Rainfall (Master – Adelaide Airport; Target – Stirling)

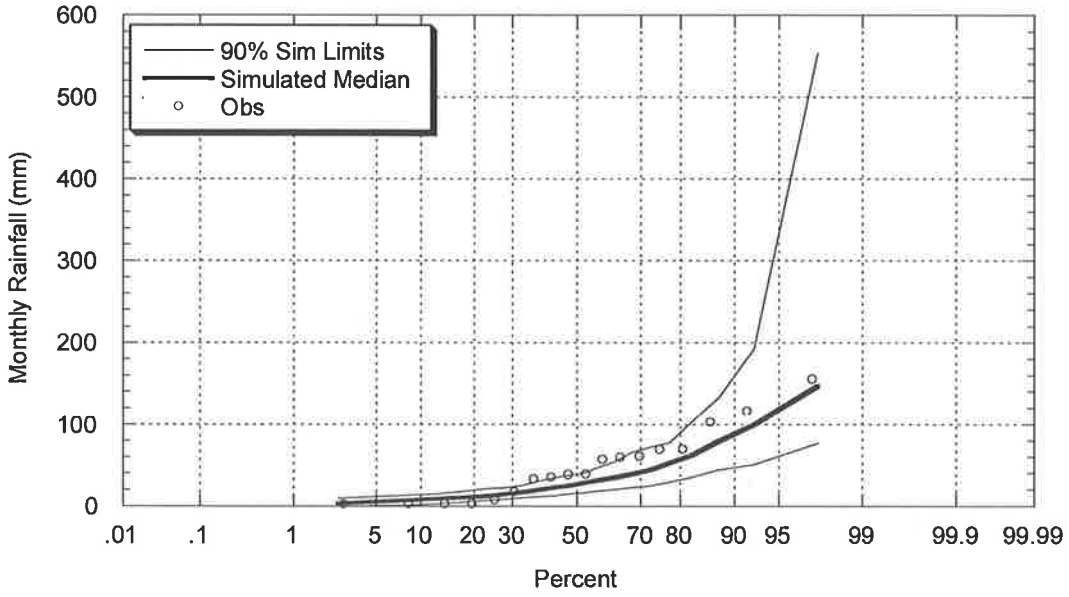

Figure C.1.34: Comparison between Observed and Target Simulated March Rainfall (Master – Adelaide Airport; Target – Stirling)

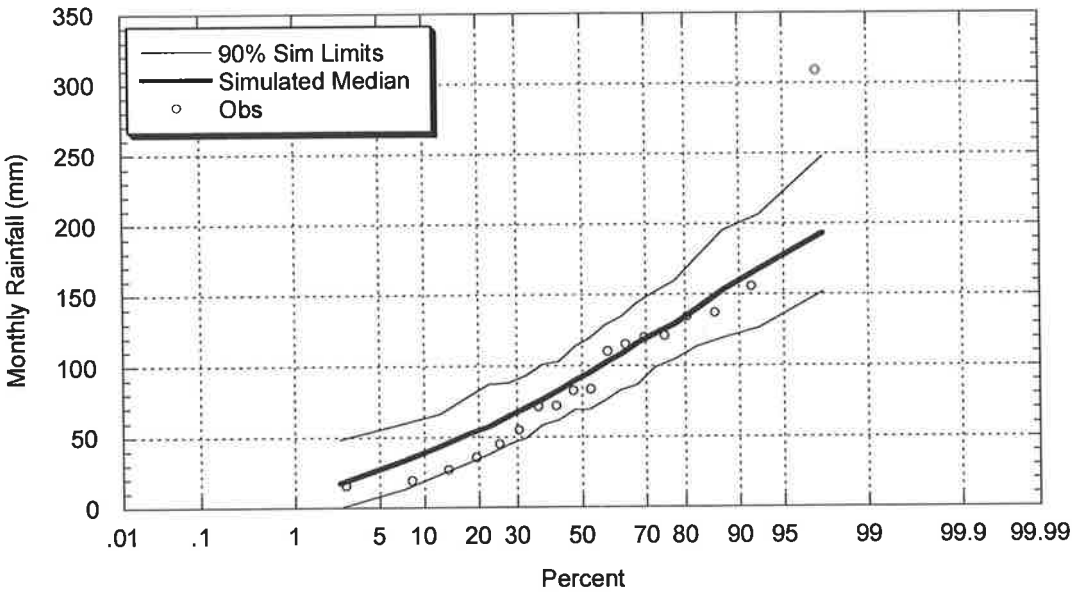

Figure C.1.35: Comparison between Observed and Target Simulated April Rainfall
(Master – Adelaide Airport; Target – Stirling)

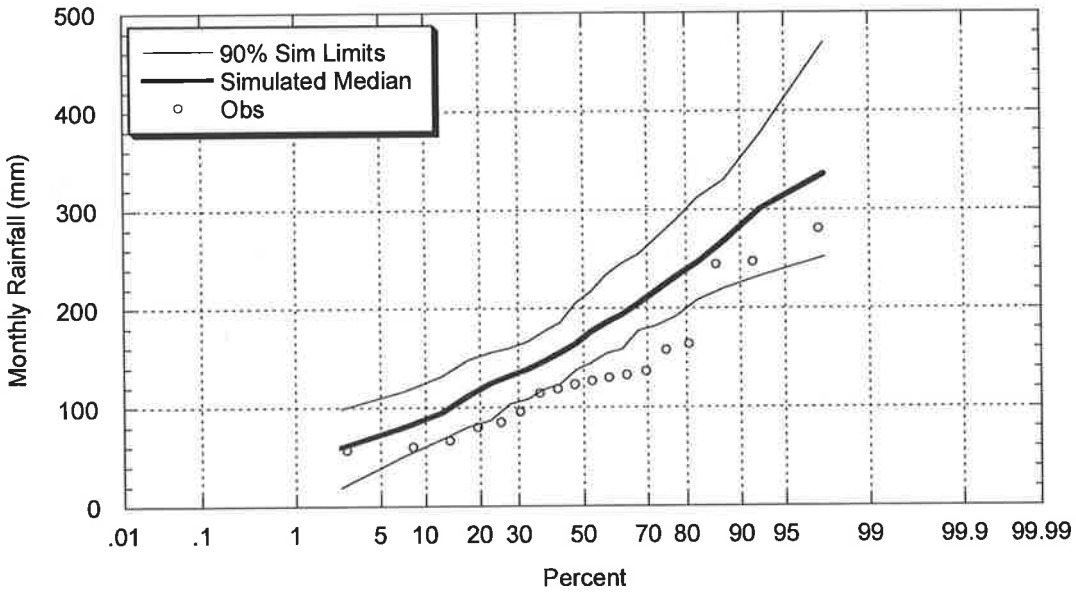

Figure C.1.36: Comparison between Observed and Target Simulated May Rainfall
(Master – Adelaide Airport; Target – Stirling)

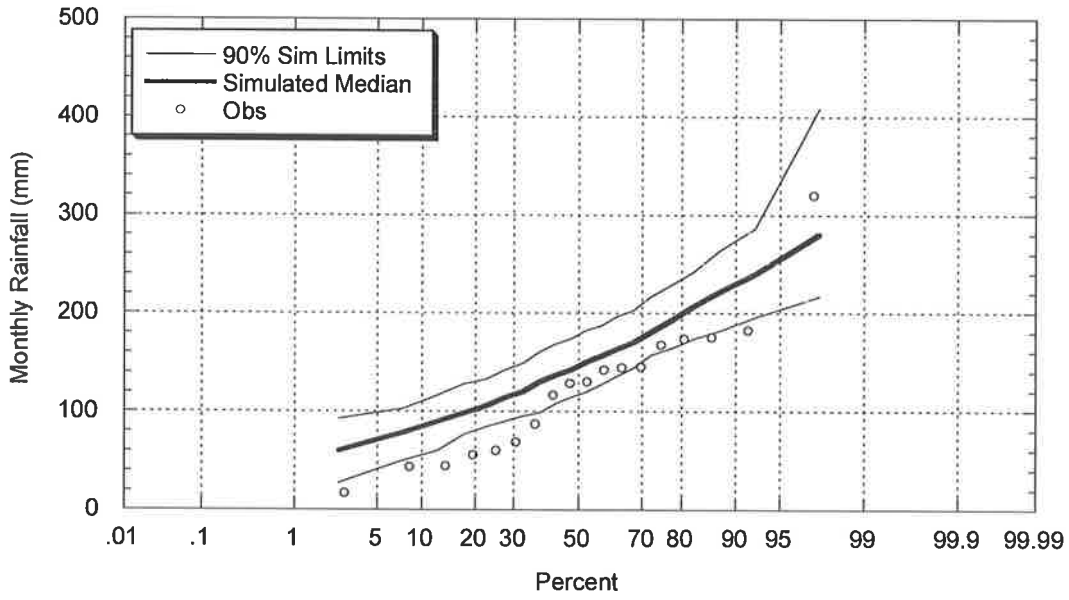

Figure C.1.37: Comparison between Observed and Target Simulated June Rainfall (Master – Adelaide Airport; Target – Stirling)

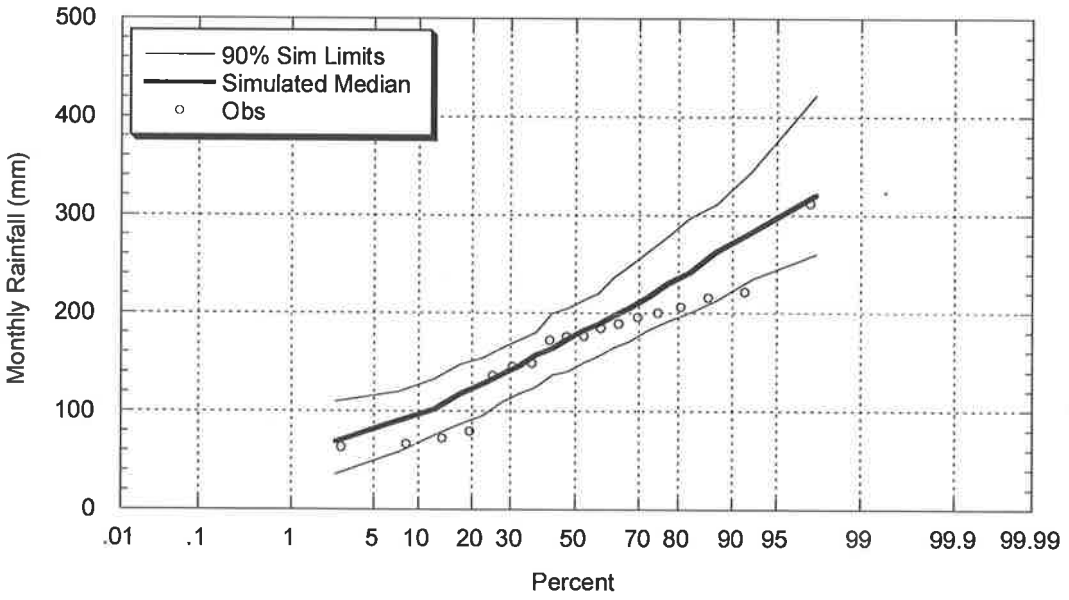

Figure C.1.38: Comparison between Observed and Target Simulated July Rainfall (Master – Adelaide Airport; Target – Stirling)

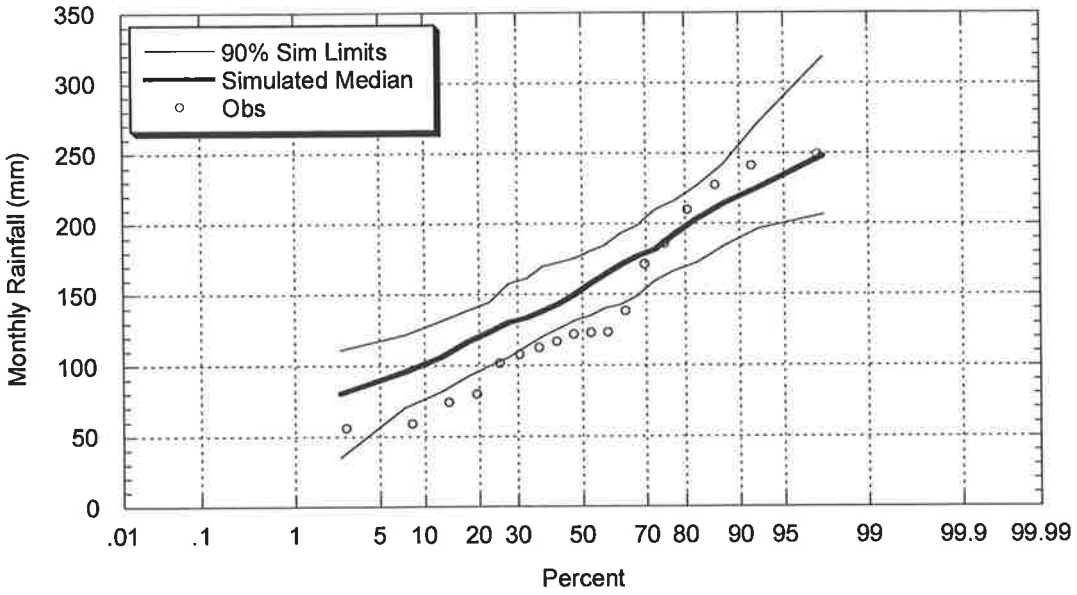

Figure C.1.39: Comparison between Observed and Target Simulated August Rainfall (Master – Adelaide Airport; Target – Stirling)

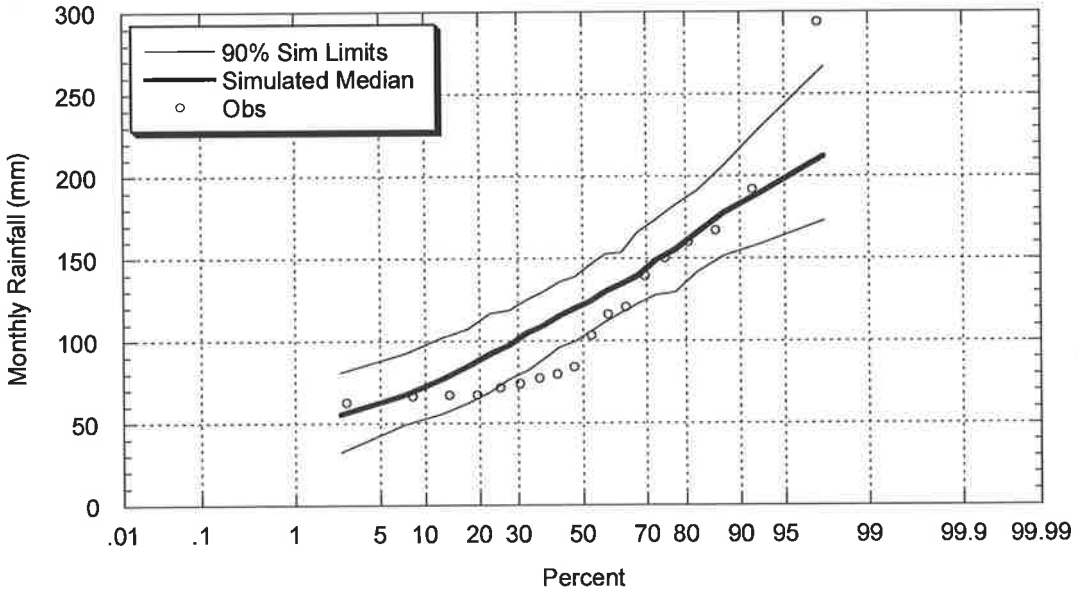

Figure C.1.40: Comparison between Observed and Target Simulated September Rainfall (Master – Adelaide Airport; Target – Stirling)

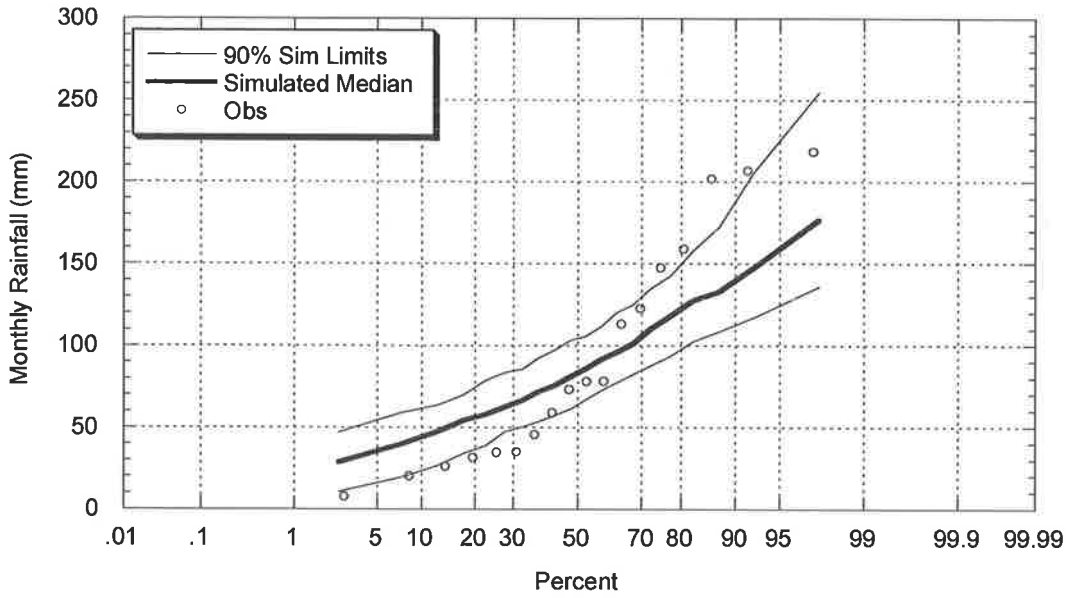

Figure C.1.41: Comparison between Observed and Target Simulated October Rainfall (Master – Adelaide Airport; Target – Stirling)

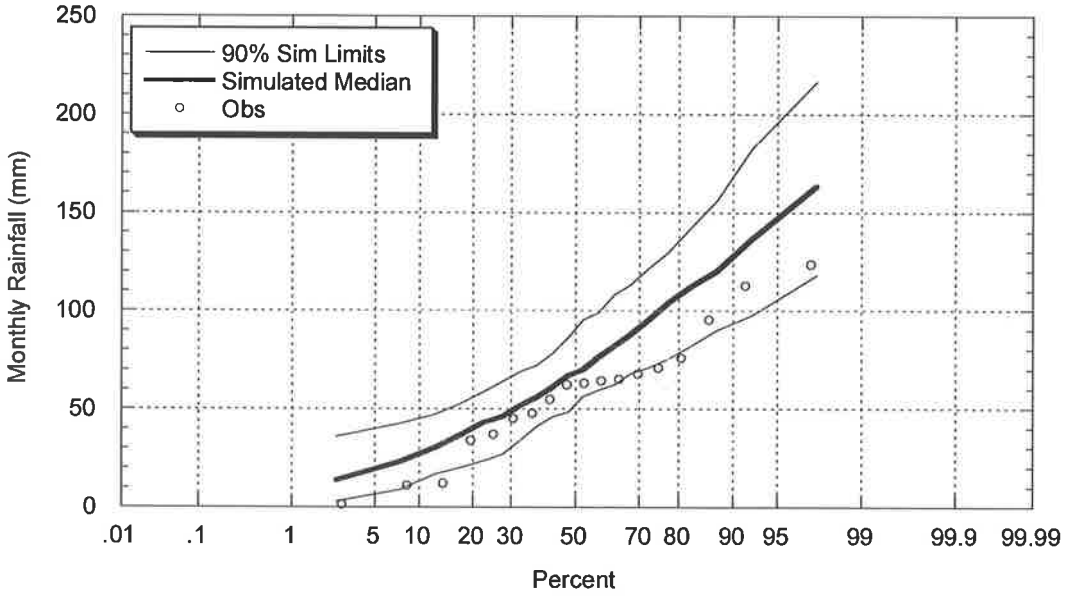

Figure C.1.42: Comparison between Observed and Target Simulated November Rainfall (Master – Adelaide Airport; Target – Stirling)

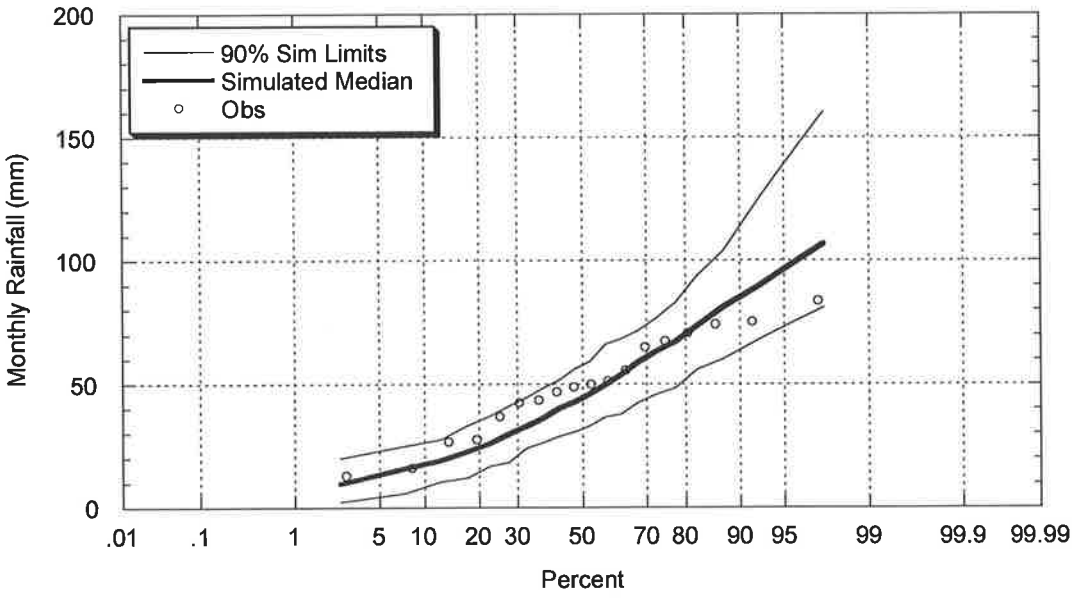

Figure C.1.43: Comparison between Observed and Target Simulated December Rainfall (Master – Adelaide Airport; Target – Stirling)

C.1.2.4 Simulated and Observed Annual Intensity – Frequency – Duration

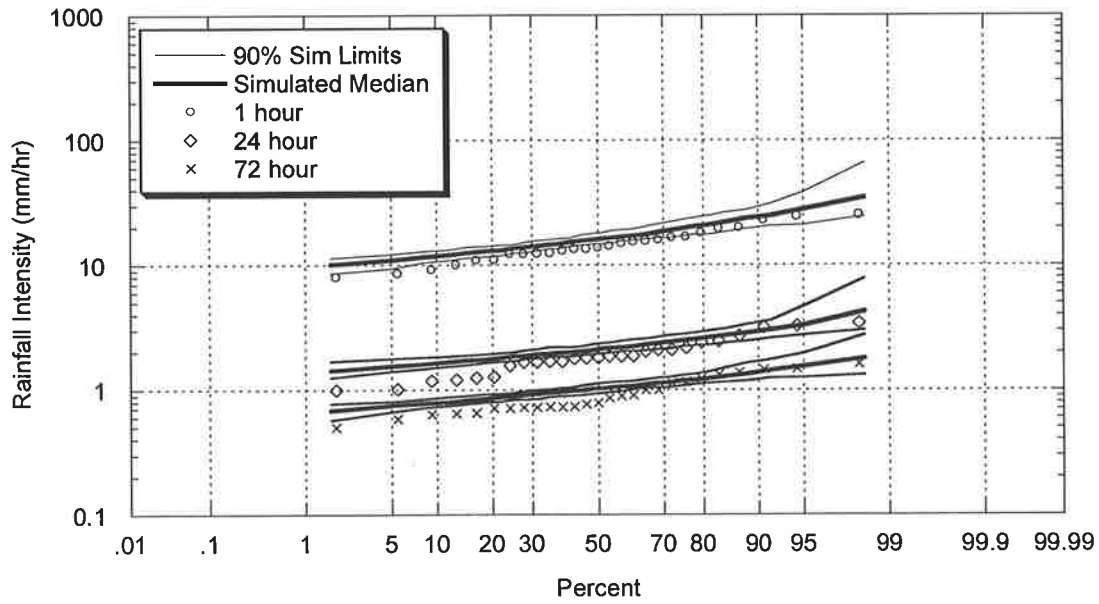

Figure C.1.44: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Adelaide Airport; Target – Stirling)

C.2 Master – Brisbane (RO), Queensland (BOM# 40214)

C.2.1 Target – Brisbane (AMO), Queensland (BOM# 40223)

C.2.1.1 Simulated and Observed Storm Event Characteristics

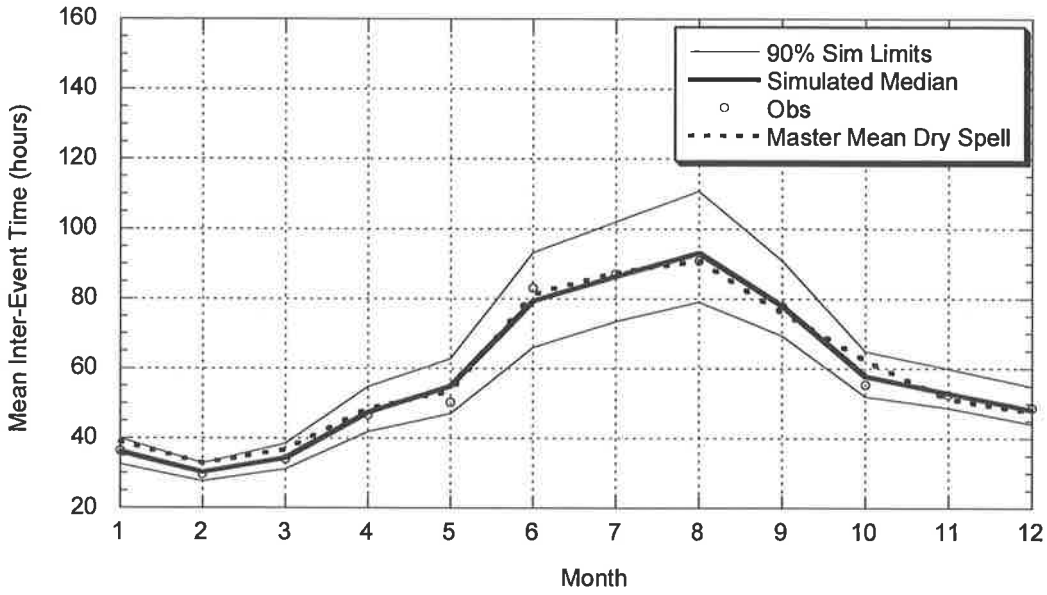

Figure C.2.1: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Brisbane Regional Office; Target – Brisbane AMO)

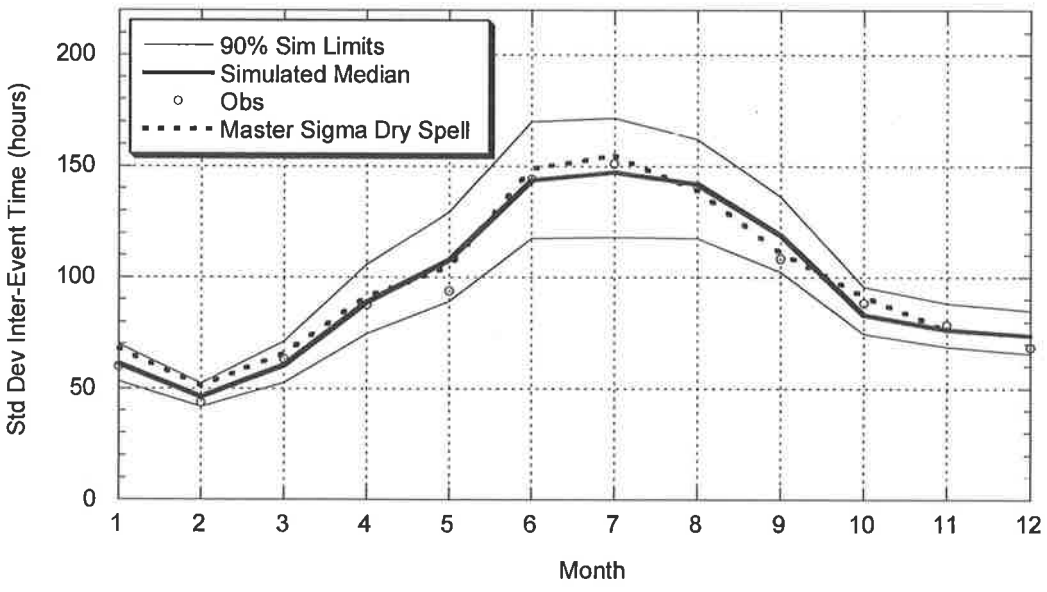

Figure C.2.2: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Brisbane Regional Office; Target – Brisbane AMO)

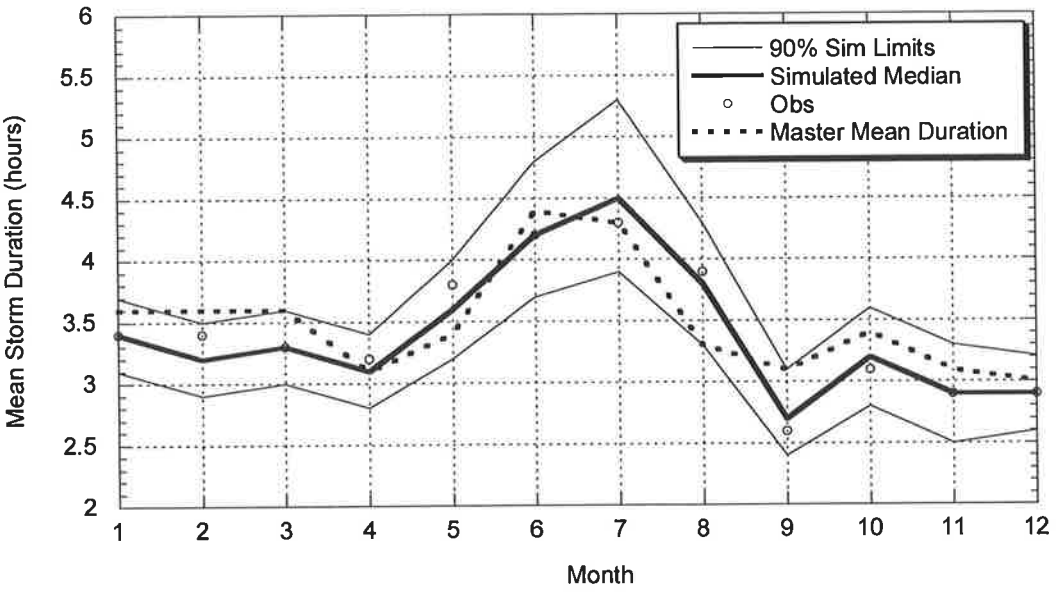

Figure C.2.3: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Brisbane Regional Office; Target – Brisbane AMO)

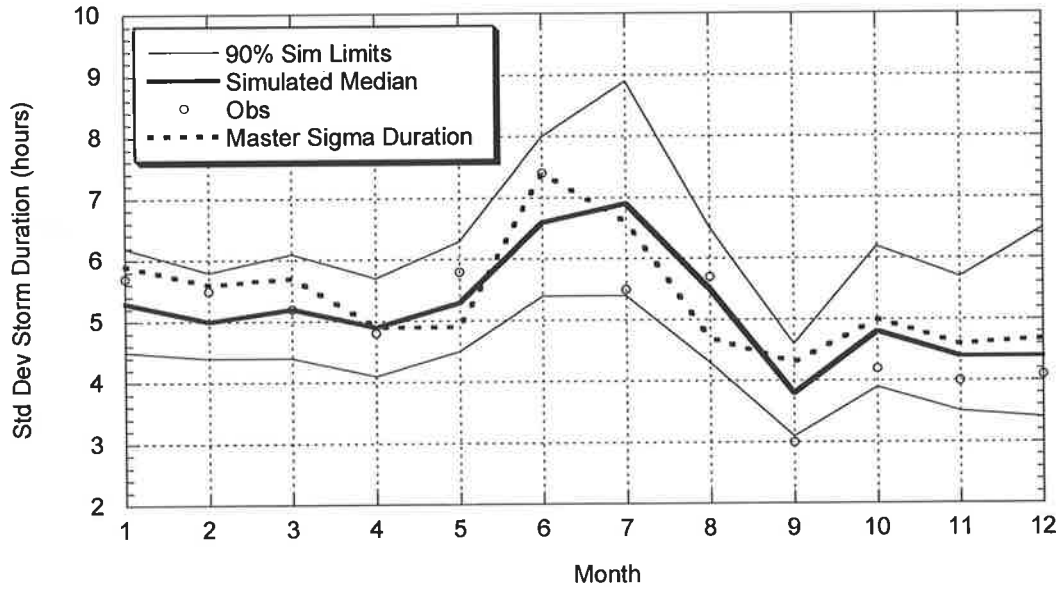

Figure C.2.4: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations (Master – Brisbane Regional Office; Target – Brisbane AMO)

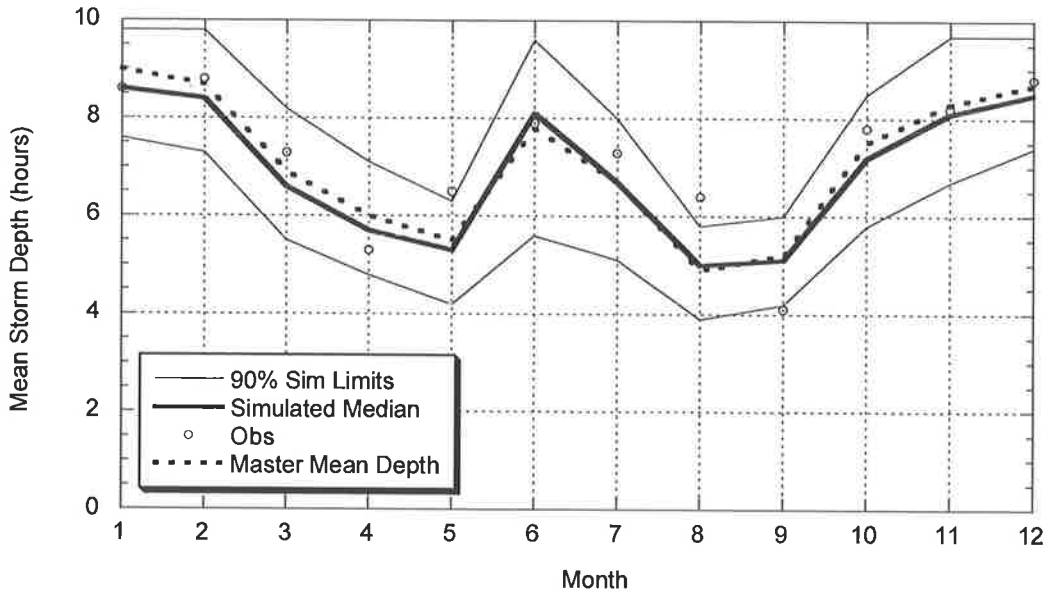

Figure C.2.5: Comparison between Observed and Target Simulated Average of Event Depths (Master – Brisbane Regional Office; Target – Brisbane AMO)

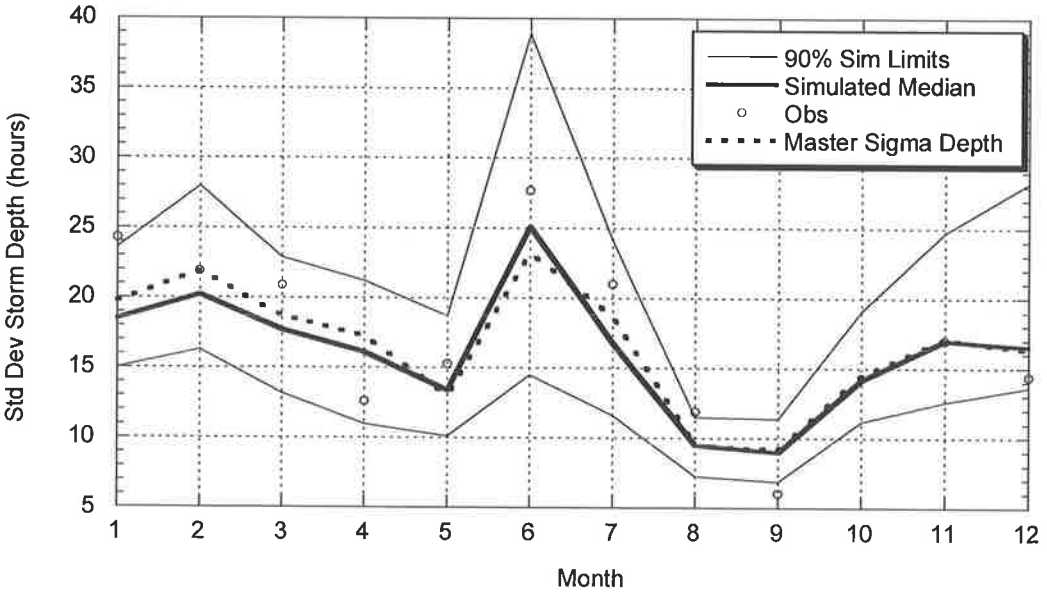

Figure C.2.6: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Brisbane Regional Office; Target – Brisbane AMO)

C.2.1.2 Simulated and Observed Daily Statistics

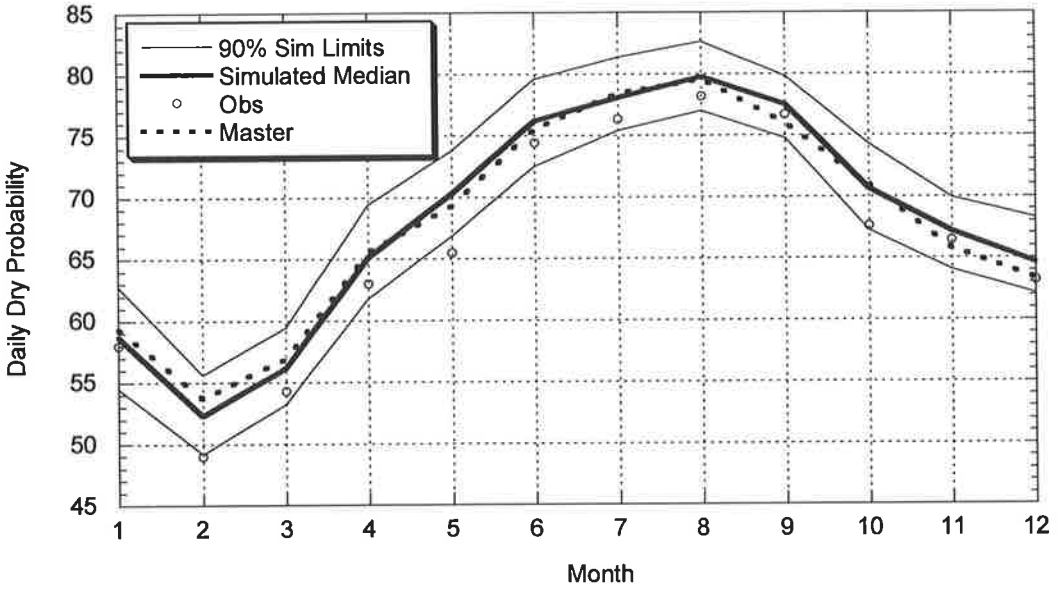

Figure C.2.7: Comparison between Observed and Target Simulated Daily Dry Probabilities (Master – Brisbane Regional Office; Target – Brisbane AMO)

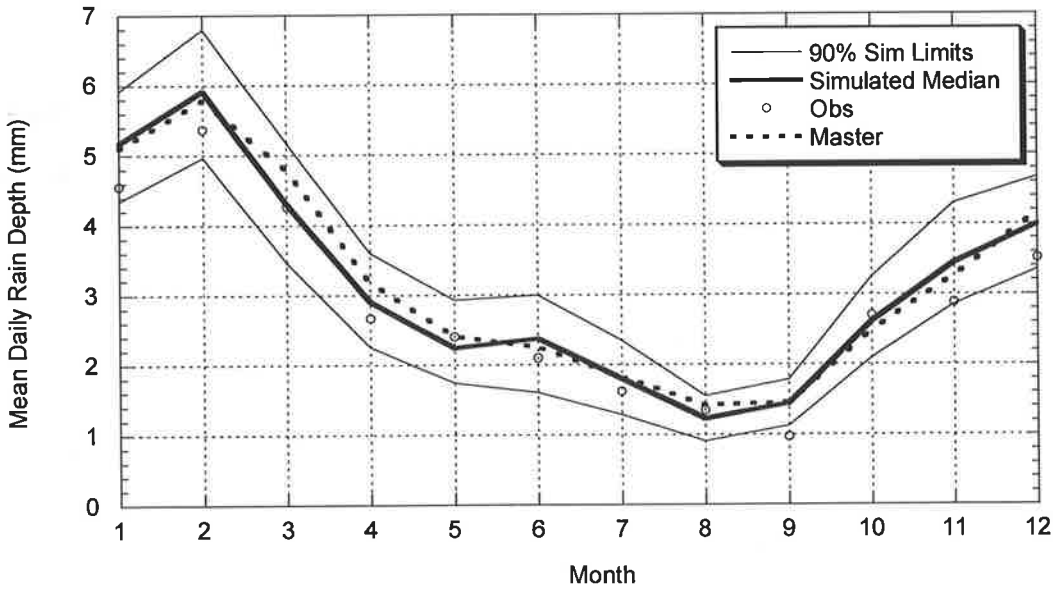

Figure C.2.8: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Brisbane Regional Office; Target – Brisbane AMO)

C.2.1.3 Simulated and Observed Annual and Monthly Rainfall

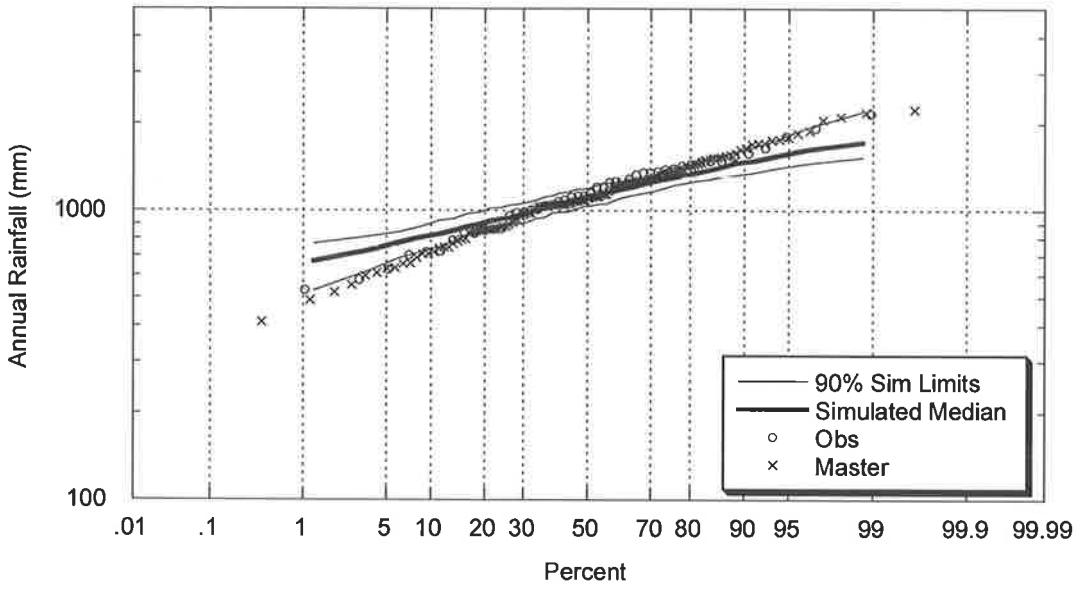

Figure C.2.9: Comparison between Observed and Target Simulated Annual Rainfall (Master – Brisbane Regional Office; Target – Brisbane AMO)

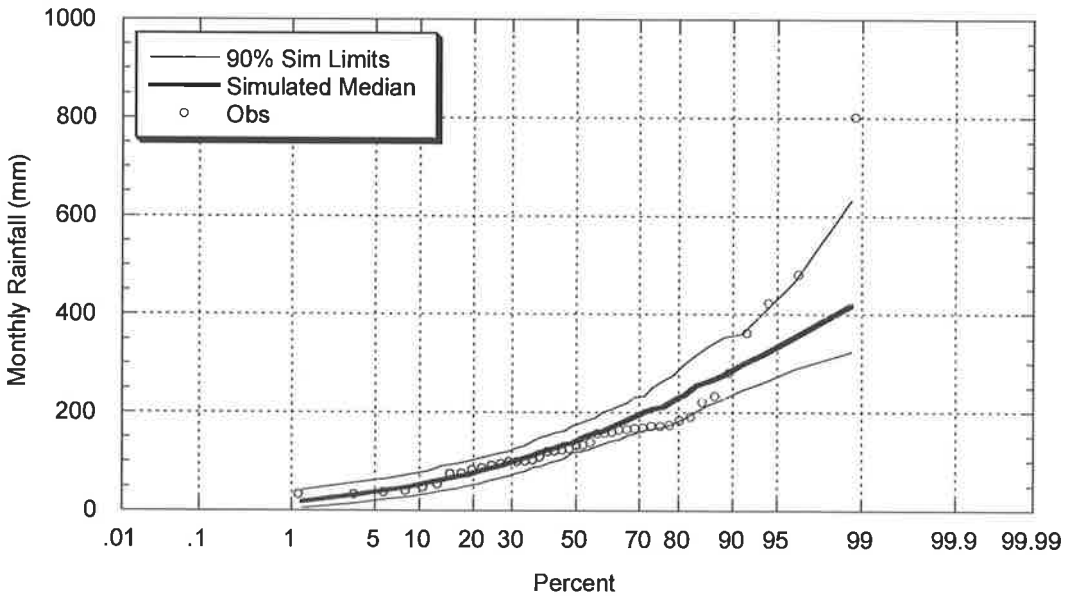

Figure C.2.10: Comparison between Observed and Target Simulated January Rainfall (Master – Brisbane Regional Office; Target – Brisbane AMO)

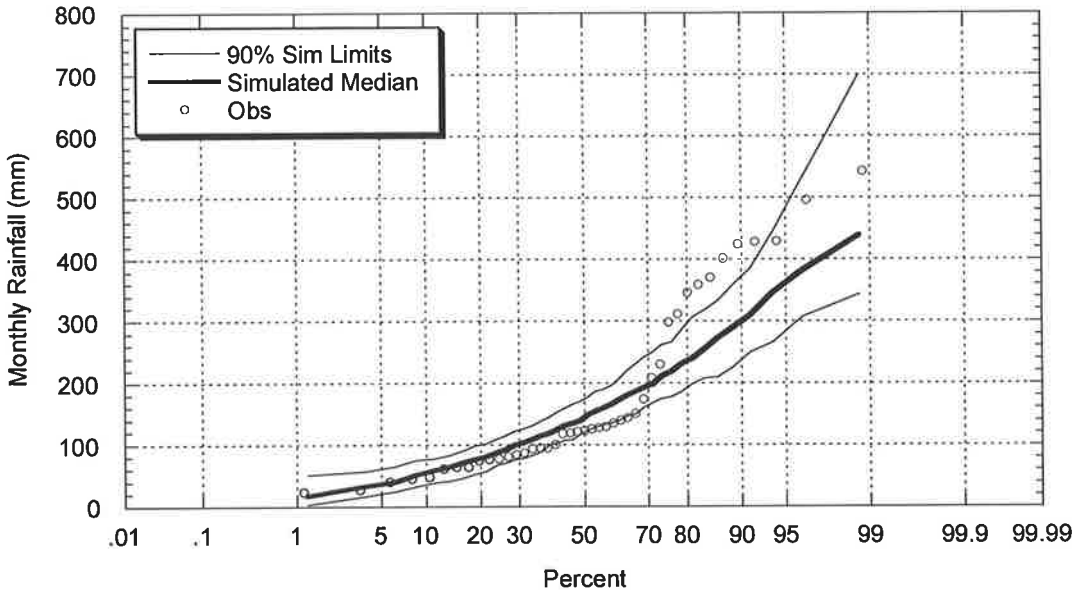

Figure C.2.11: Comparison between Observed and Target Simulated February Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

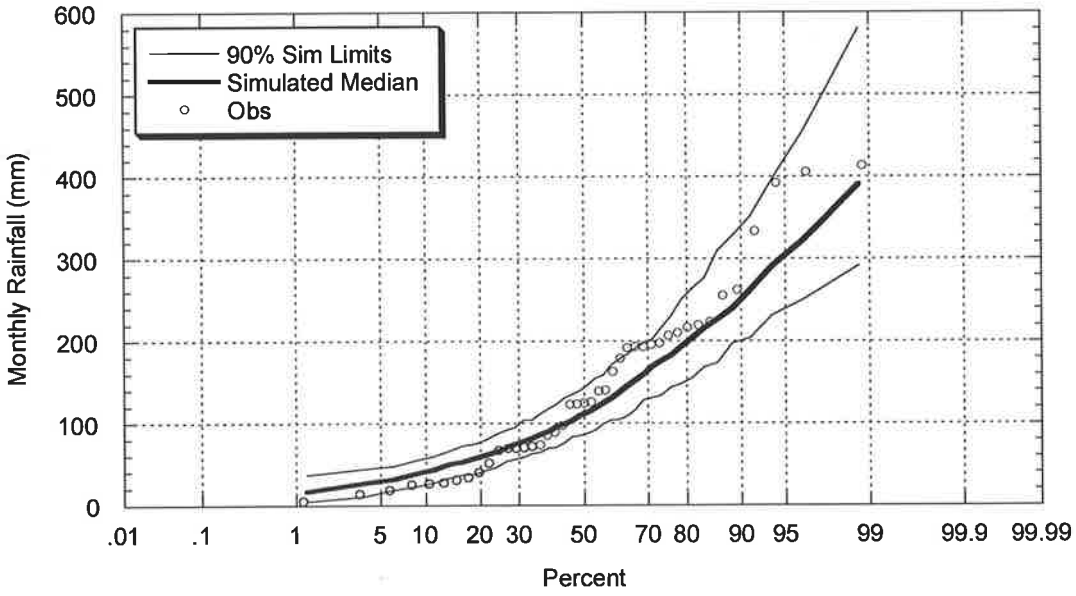

Figure C.2.12: Comparison between Observed and Target Simulated March Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

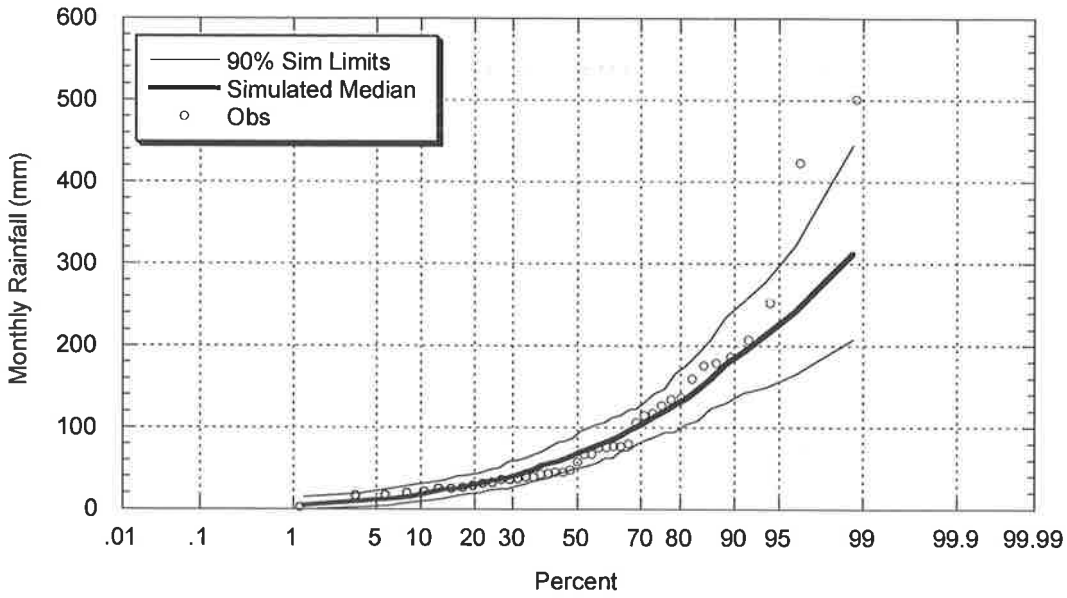

Figure C.2.13: Comparison between Observed and Target Simulated April Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

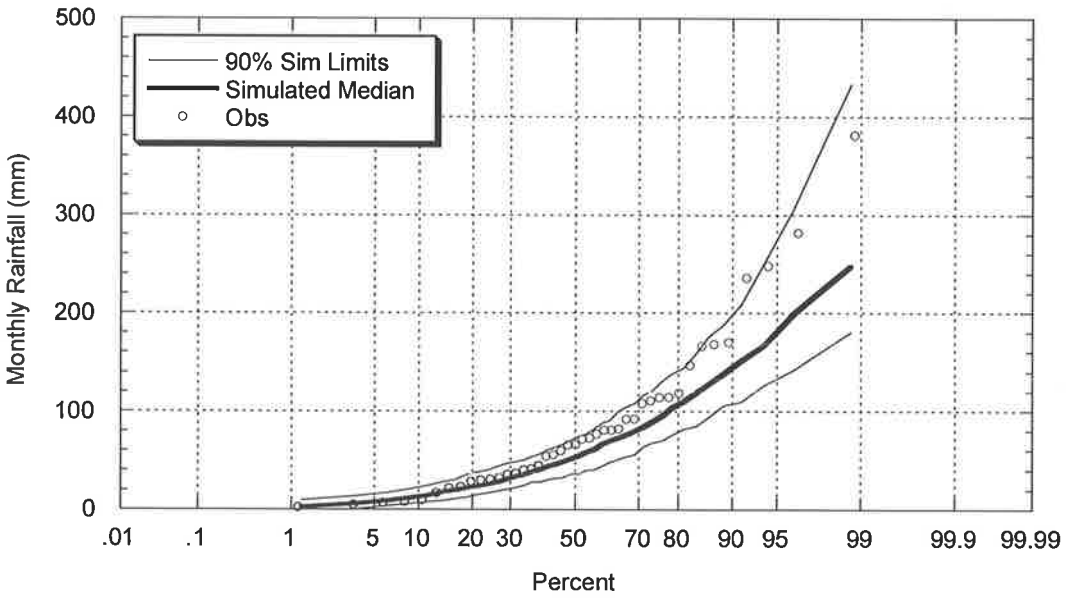

Figure C.2.14: Comparison between Observed and Target Simulated May Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

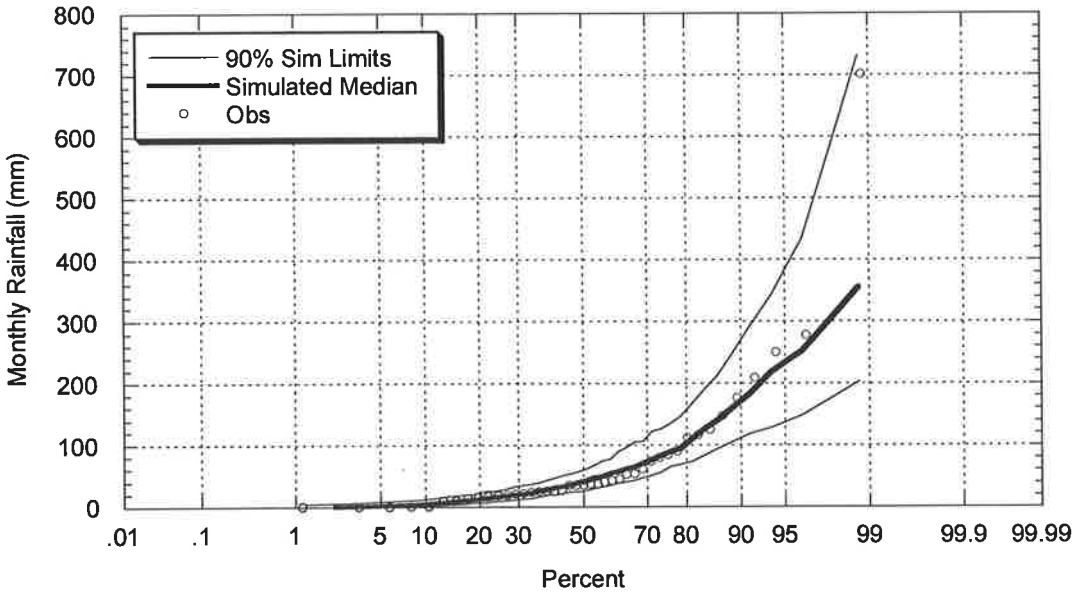

Figure C.2.15: Comparison between Observed and Target Simulated June Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

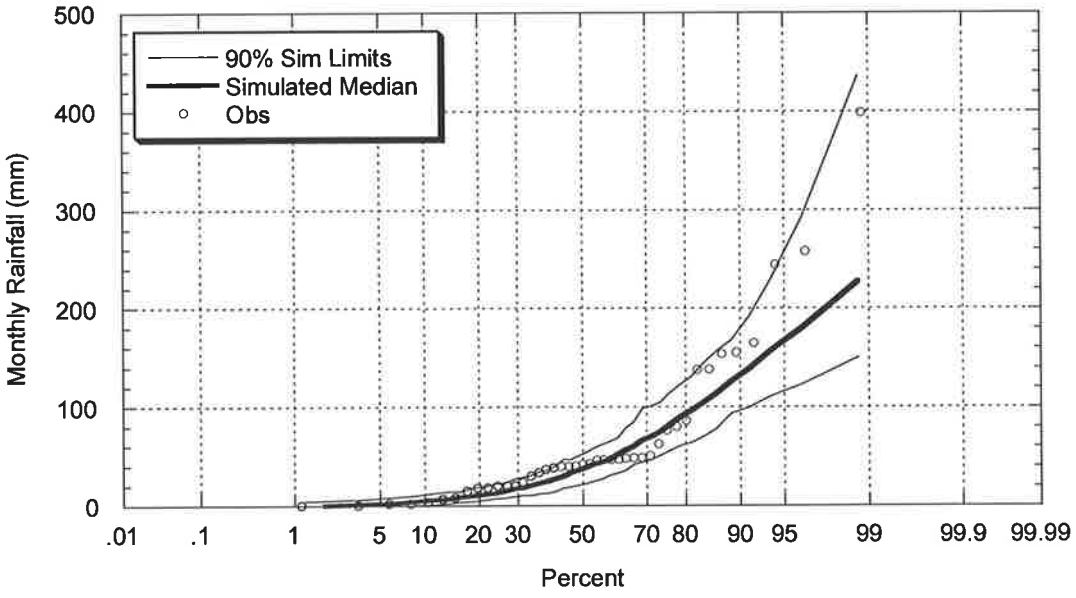

Figure C.2.16: Comparison between Observed and Target Simulated July Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

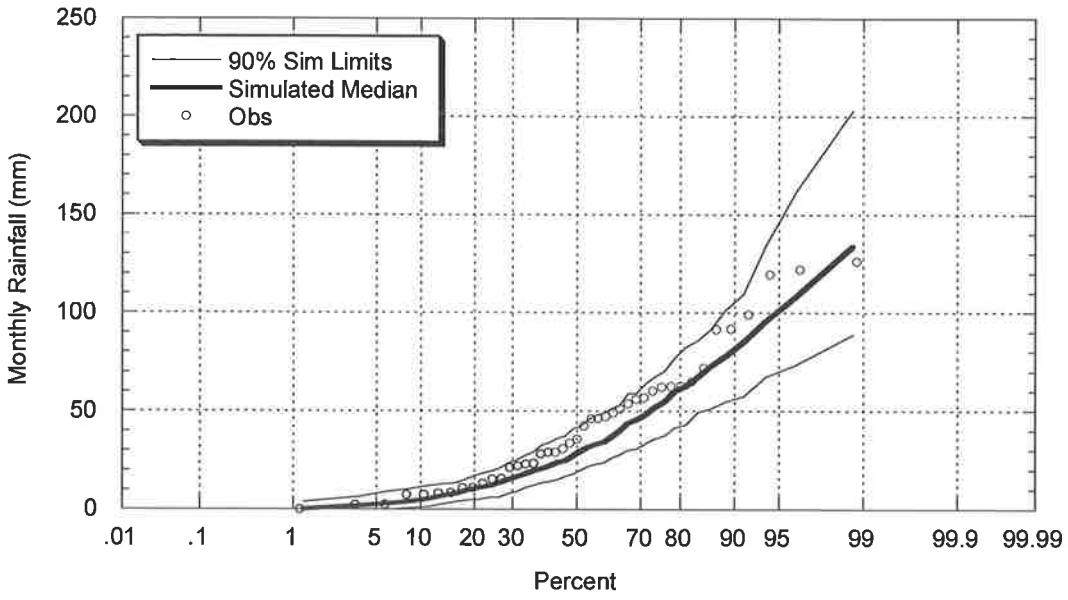

Figure C.2.17: Comparison between Observed and Target Simulated August Rainfall (Master – Brisbane Regional Office; Target – Brisbane AMO)

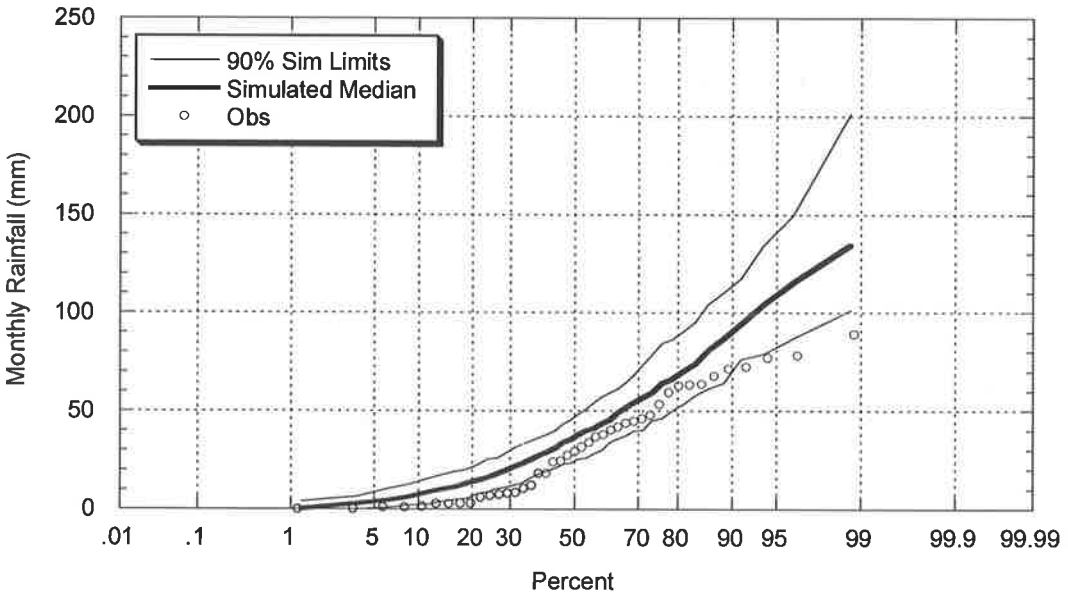

Figure C.2.18: Comparison between Observed and Target Simulated September Rainfall (Master – Brisbane Regional Office; Target – Brisbane AMO)

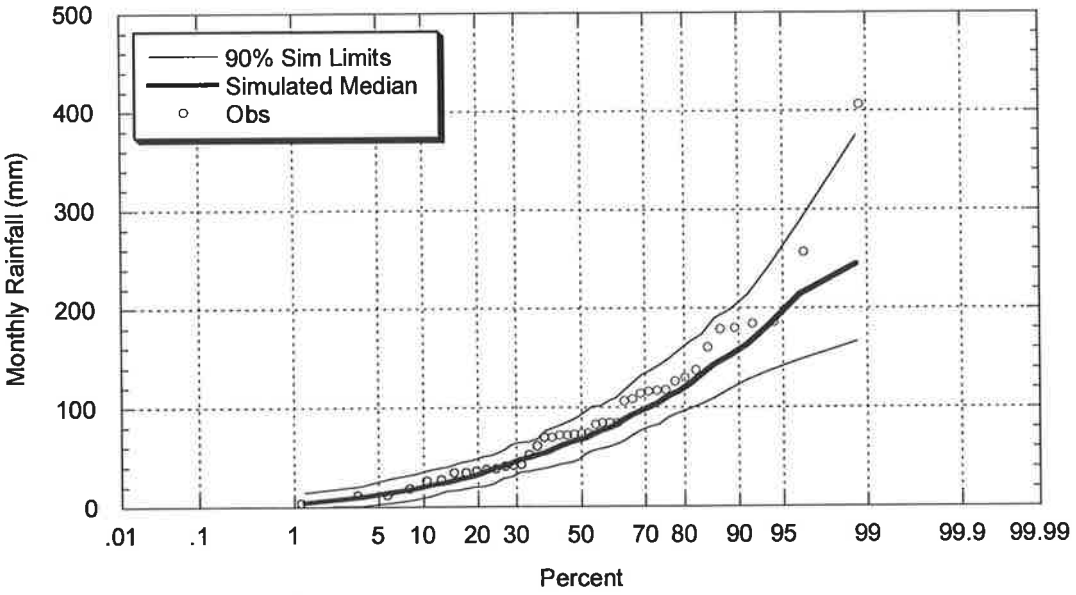

Figure C.2.19: Comparison between Observed and Target Simulated October Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

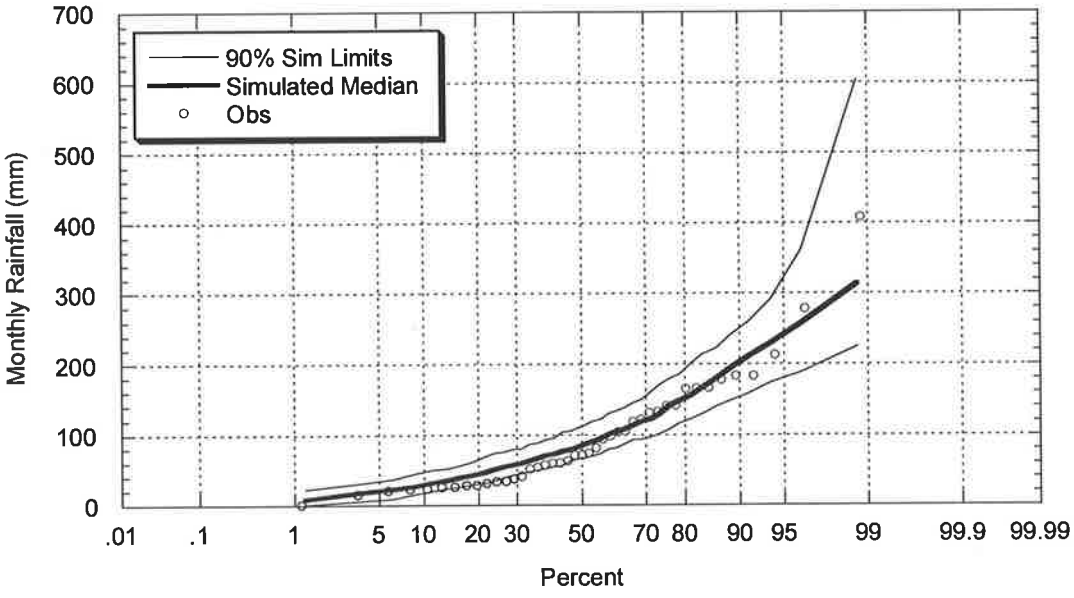

Figure C.2.20: Comparison between Observed and Target Simulated November Rainfall
(Master – Brisbane Regional Office; Target – Brisbane AMO)

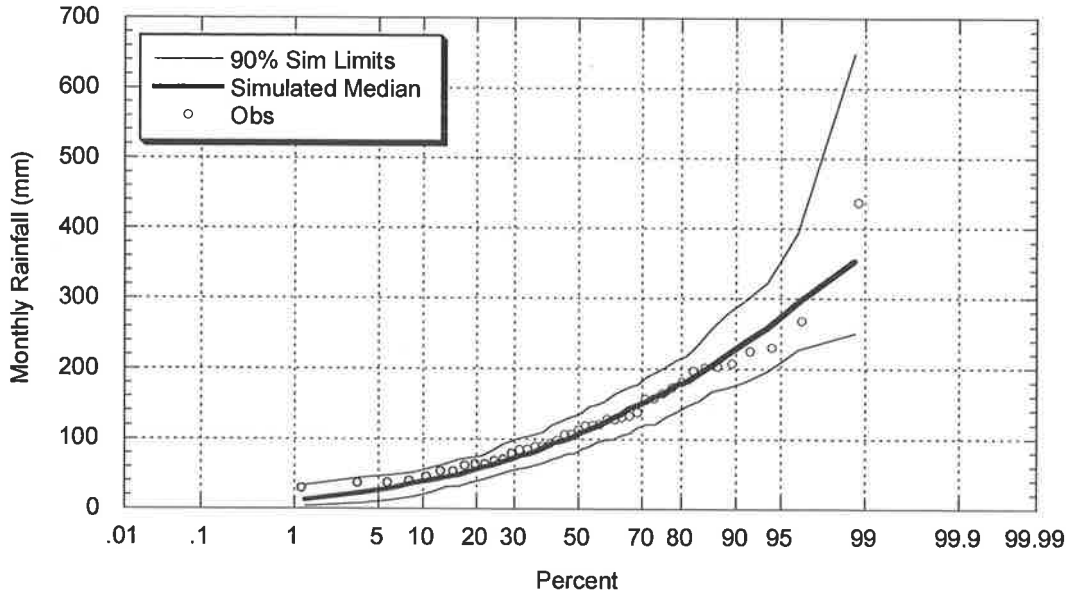

Figure C.2.21: Comparison between Observed and Target Simulated December Rainfall (Master – Brisbane Regional Office; Target – Brisbane AMO)

C.2.1.4 Simulated and Observed Annual Intensity – Frequency – Duration

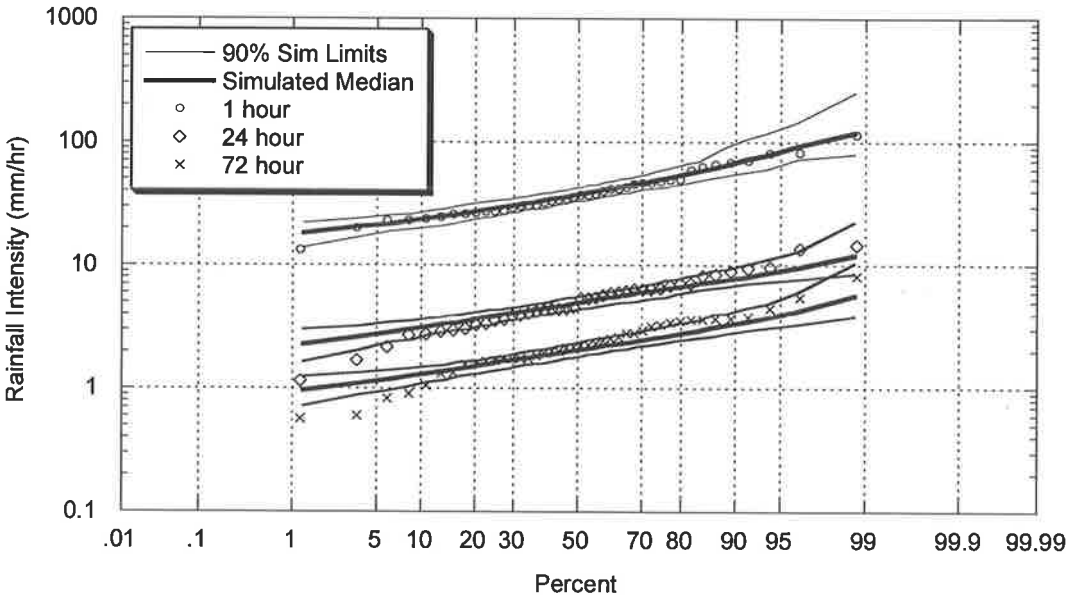

Figure C.2.22: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Brisbane Regional Office; Target – Brisbane AMO)

C.2.2 Target – Kirkleigh, Queensland (BOM# 40318)

C.2.2.1 Simulated and Observed Storm Event Characteristics

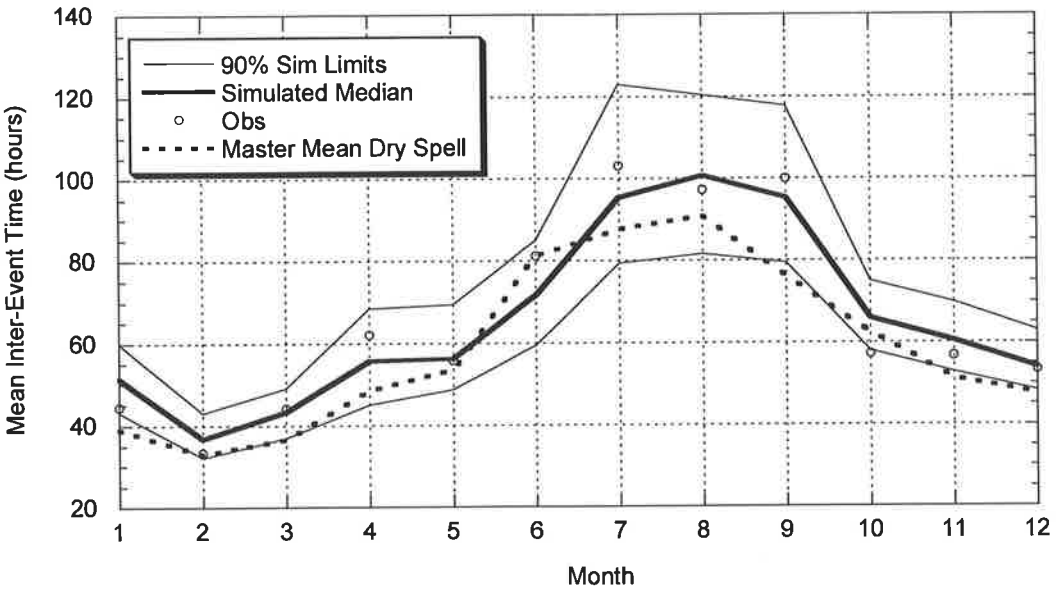

Figure C.2.23: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Brisbane Regional Office; Target – Kirkleagh)

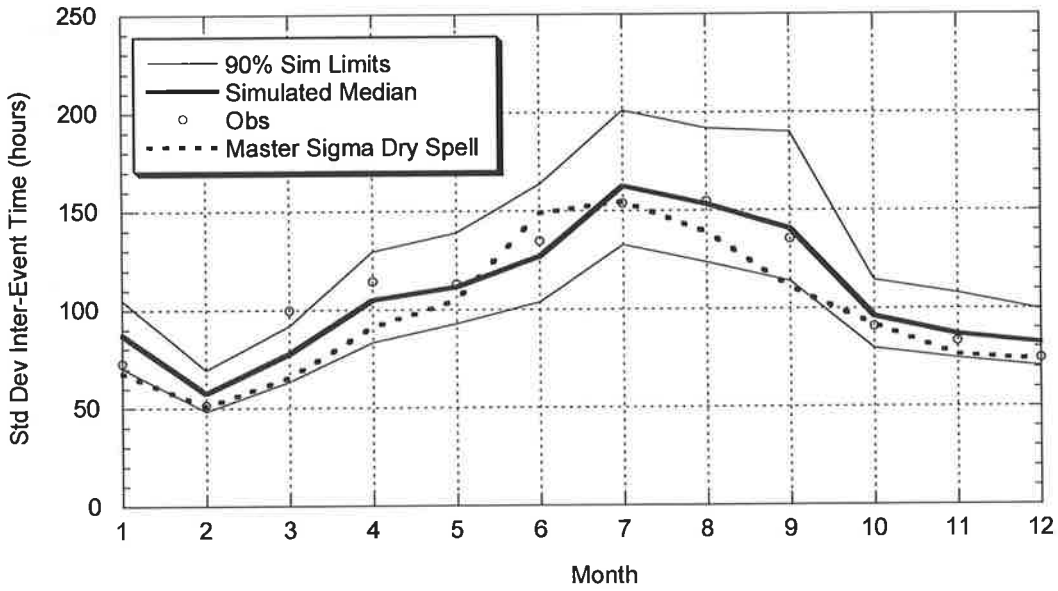

Figure C.2.24: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times Master – Brisbane Regional Office; Target – Kirkleagh)

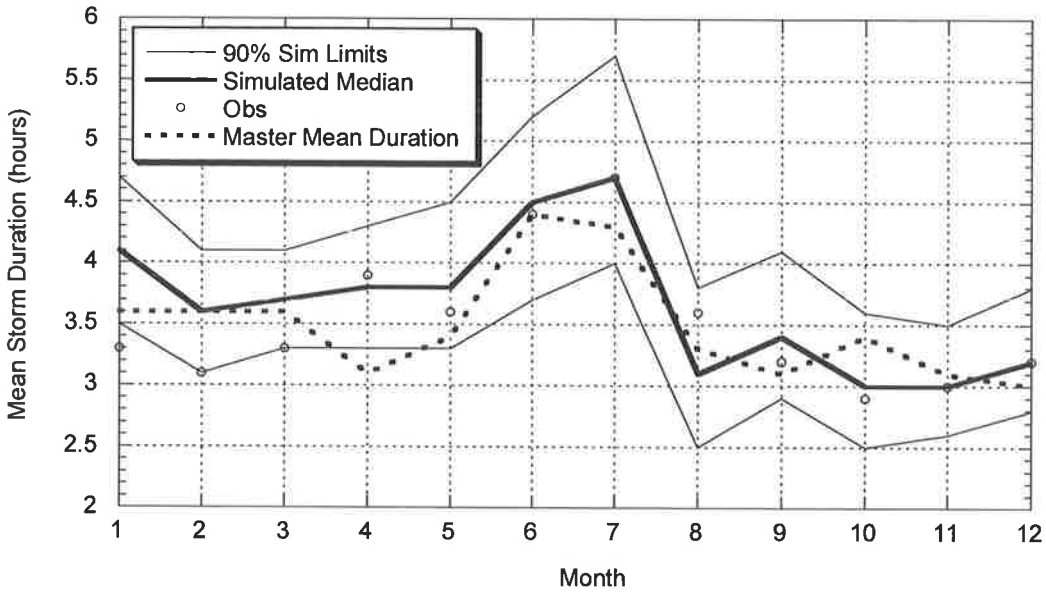

Figure C.2.25: Comparison between Observed and Target Simulated Mean of Event Storm Durations Master – Brisbane Regional Office; Target – Kirkleagh)

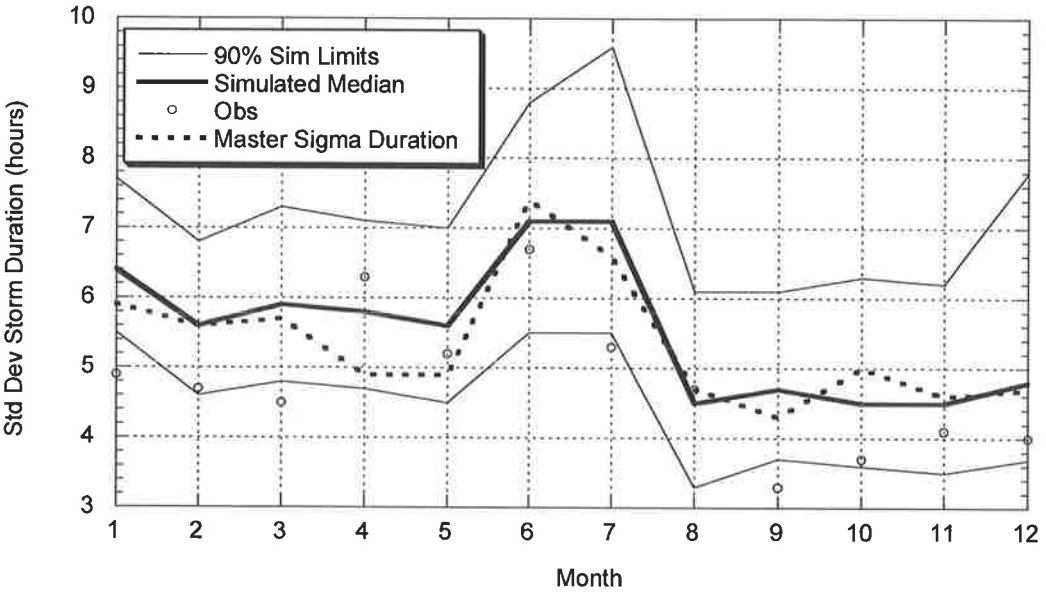

Figure C.2.26: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations Master – Brisbane Regional Office; Target – Kirkleagh)

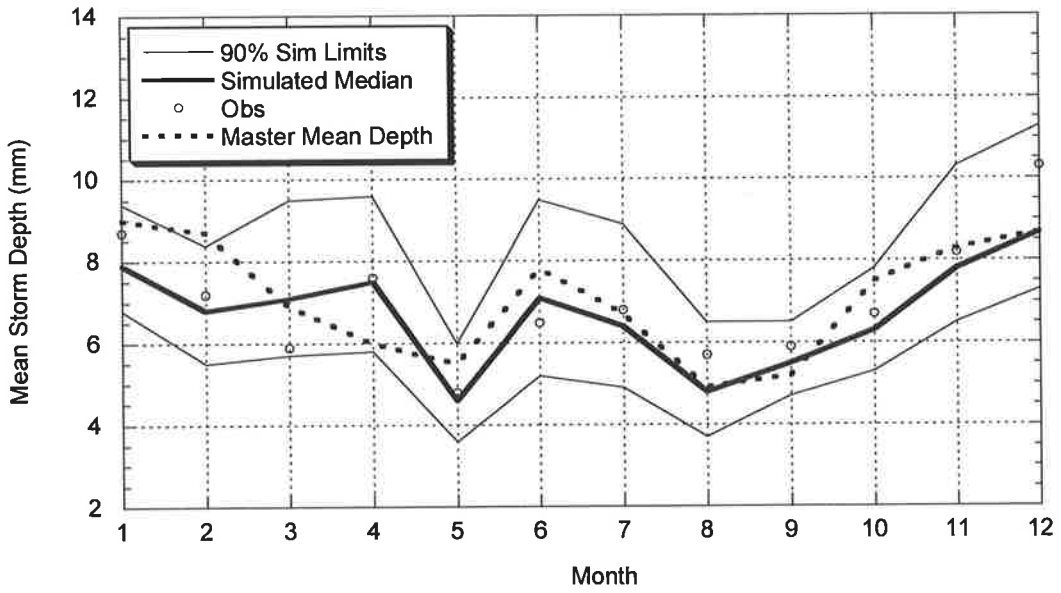

Figure C.2.27: Comparison between Observed and Target Simulated Average of Event Depths Master – Brisbane Regional Office; Target – Kirkleagh)

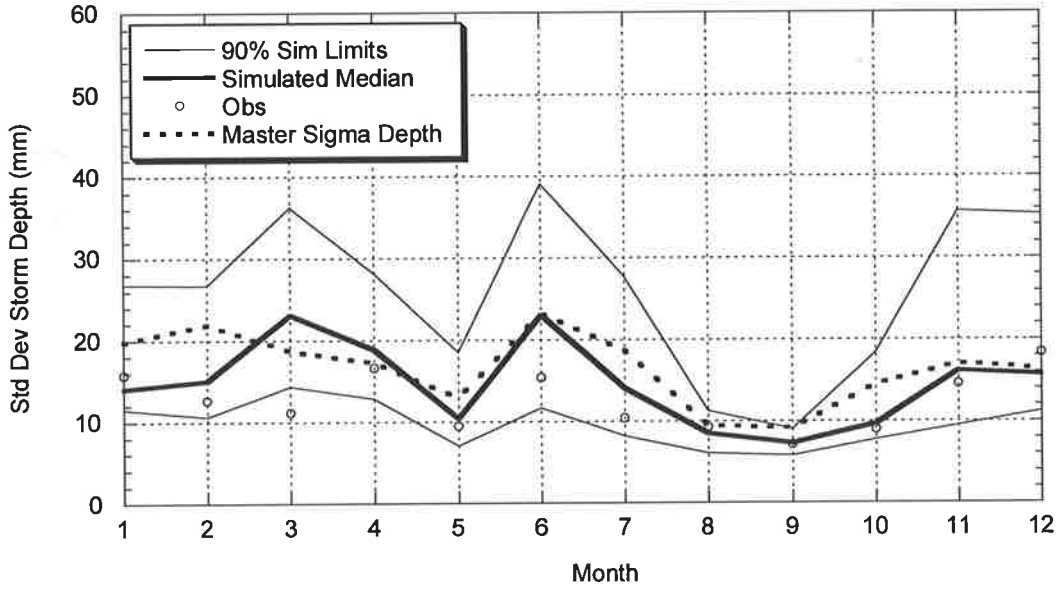

Figure C.2.28: Comparison between Observed and Target Simulated Standard Deviation of Event Depths Master – Brisbane Regional Office; Target – Kirkleagh)

C.2.2.2 Simulated and Observed Daily Statistics

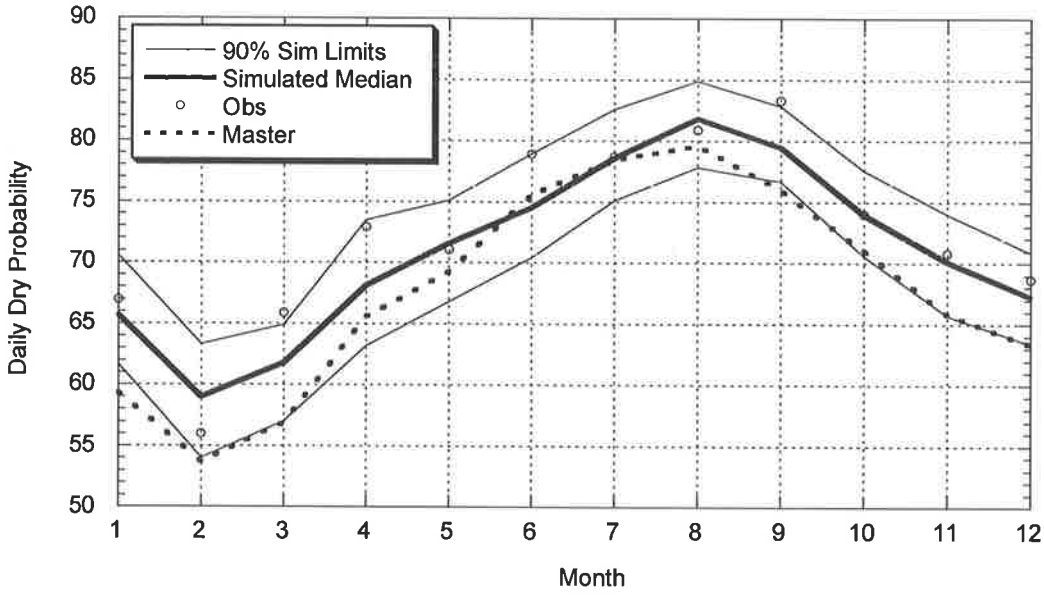

Figure C.2.29: Comparison between Observed and Target Simulated Daily Dry Probabilities Master – Brisbane Regional Office; Target – Kirkleagh)

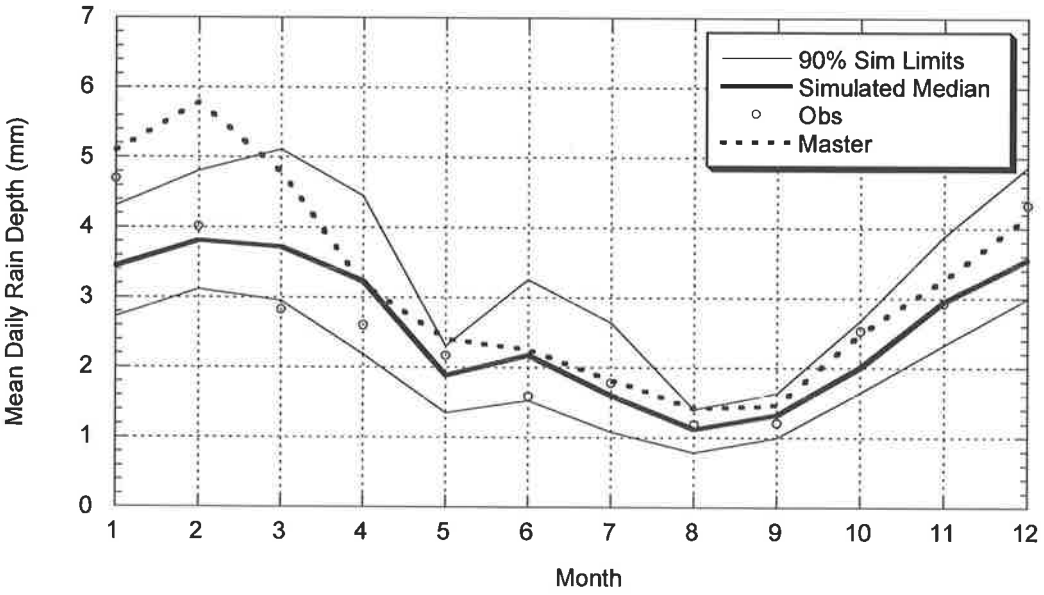

Figure C.2.30: Comparison between Observed and Target Simulated Daily Mean Depth Master – Brisbane Regional Office; Target – Kirkleagh)

C.2.2.3 Simulated and Observed Annual and Monthly Rainfall

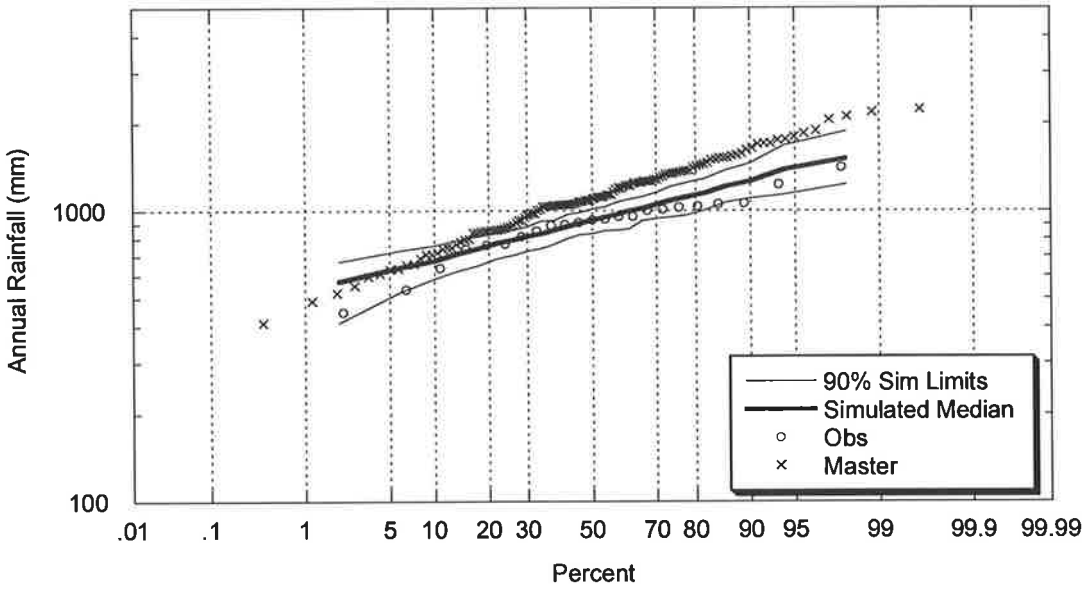

Figure C.2.31: Comparison between Observed and Target Simulated Annual Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

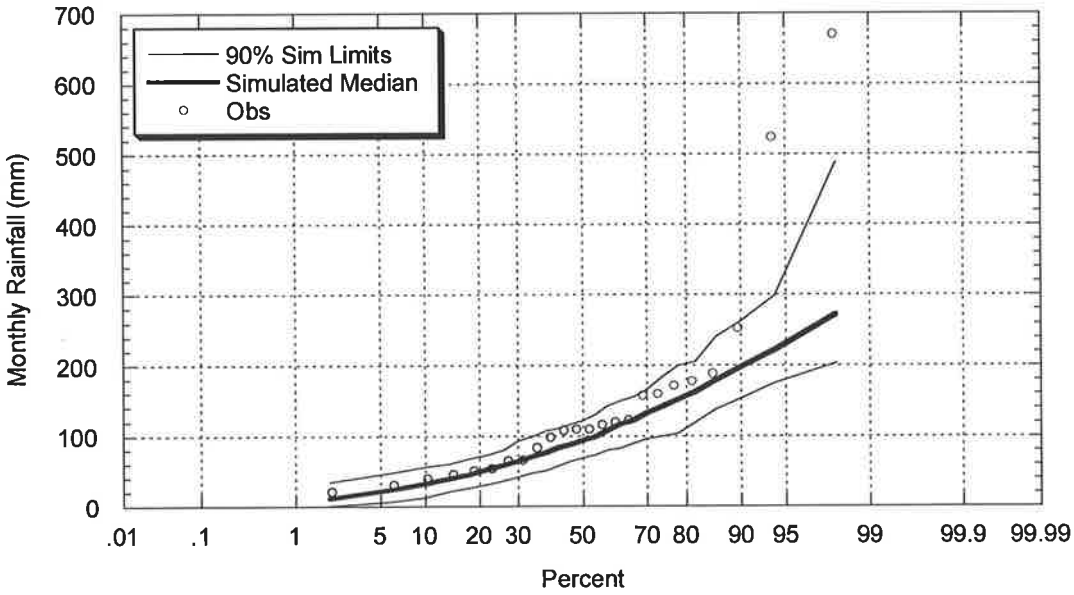

Figure C.2.32: Comparison between Observed and Target Simulated January Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

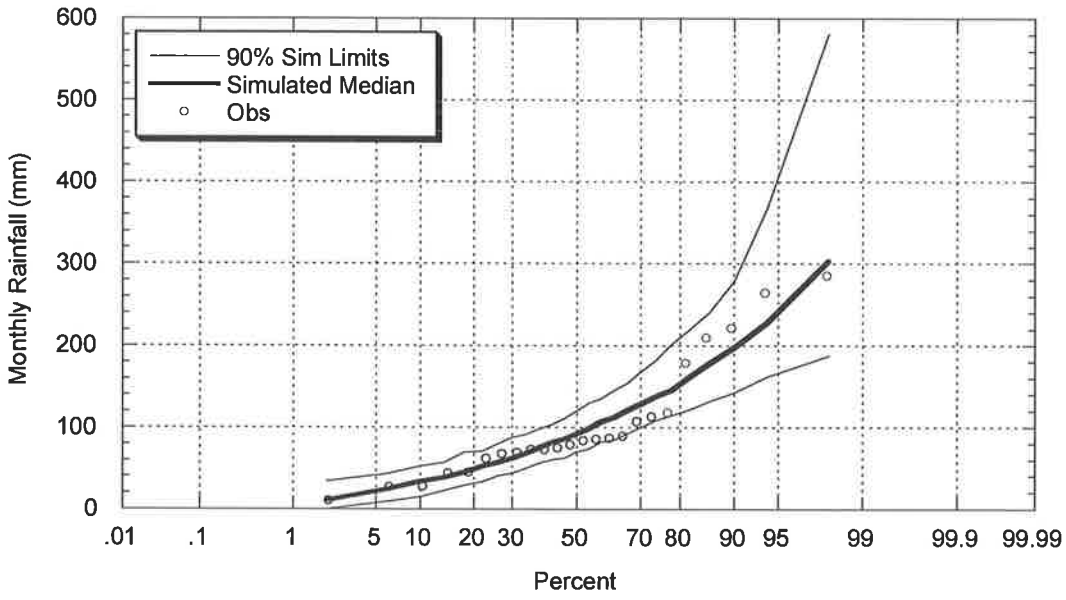

Figure C.2.33: Comparison between Observed and Target Simulated February Rainfall
Master – Brisbane Regional Office; Target – Kirkleagh)

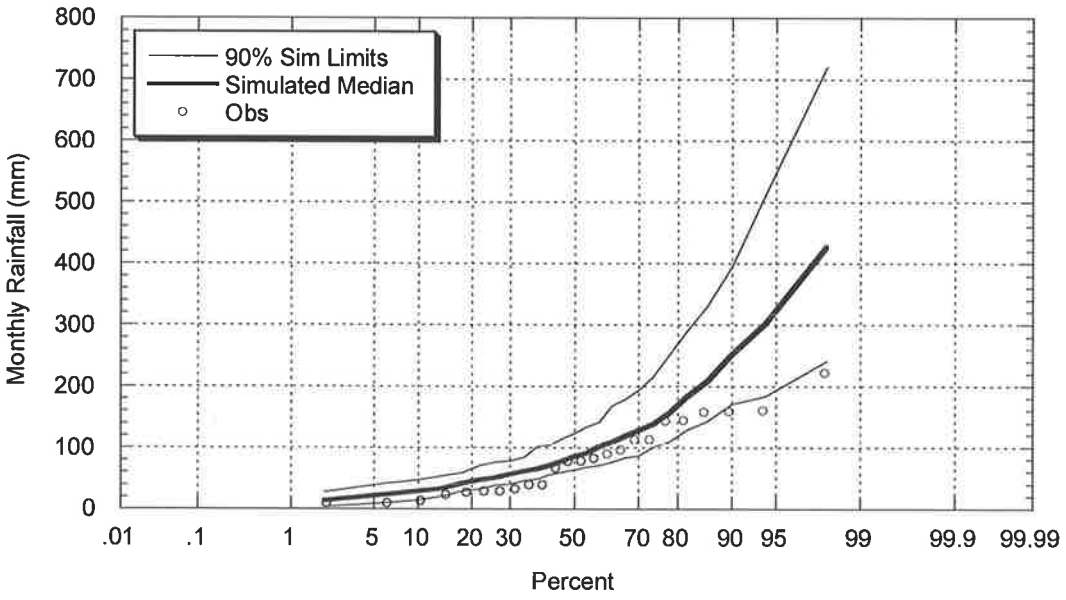

Figure C.2.34: Comparison between Observed and Target Simulated March Rainfall
Master – Brisbane Regional Office; Target – Kirkleagh)

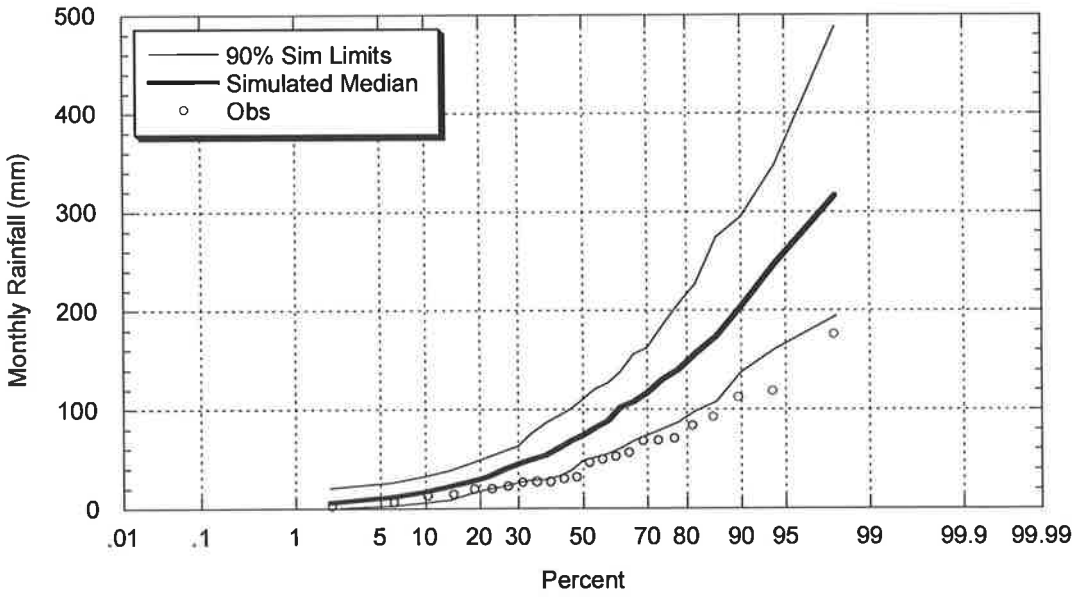

Figure C.2.35: Comparison between Observed and Target Simulated April Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

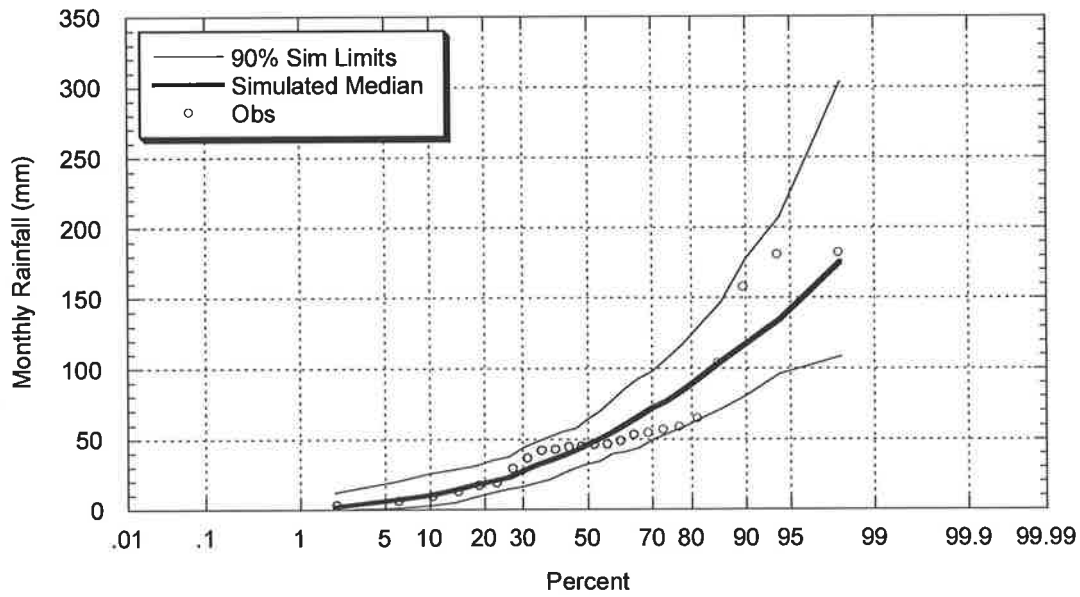

Figure C.2.36: Comparison between Observed and Target Simulated May Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

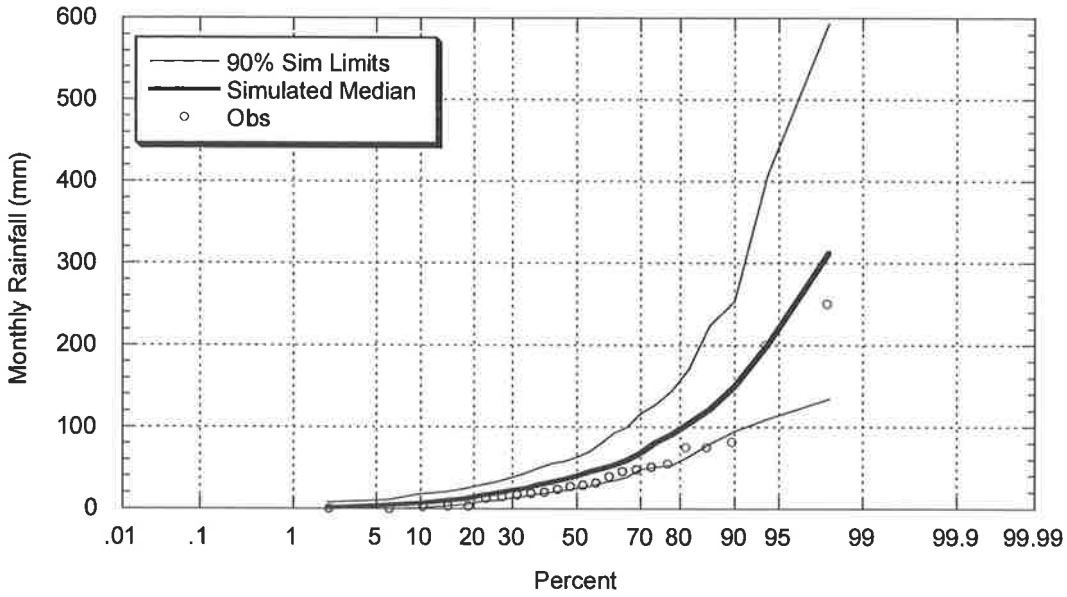

Figure C.2.37: Comparison between Observed and Target Simulated June Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

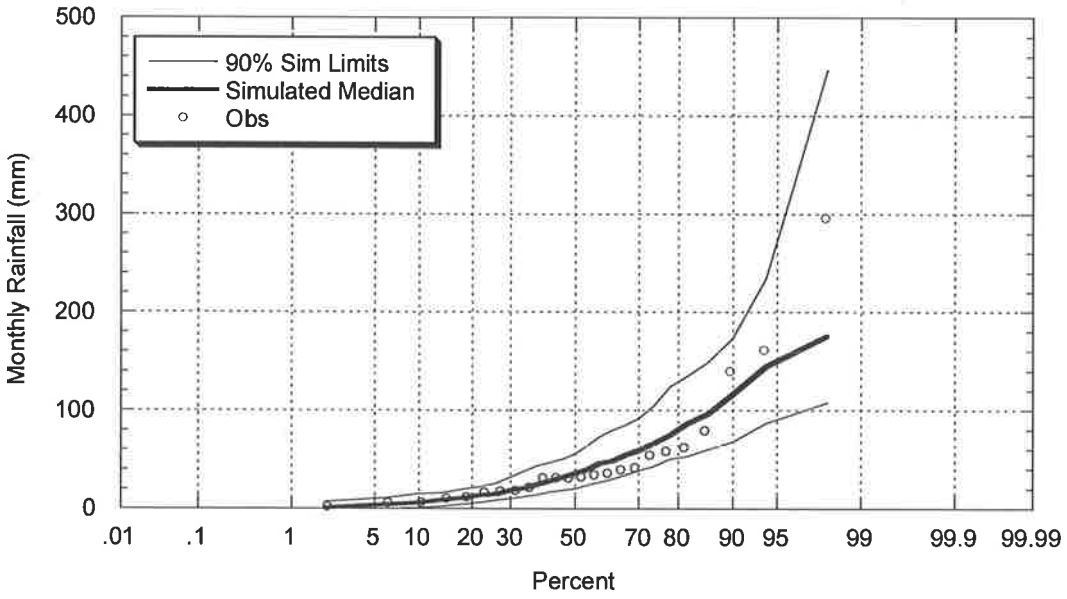

Figure C.2.38: Comparison between Observed and Target Simulated July Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

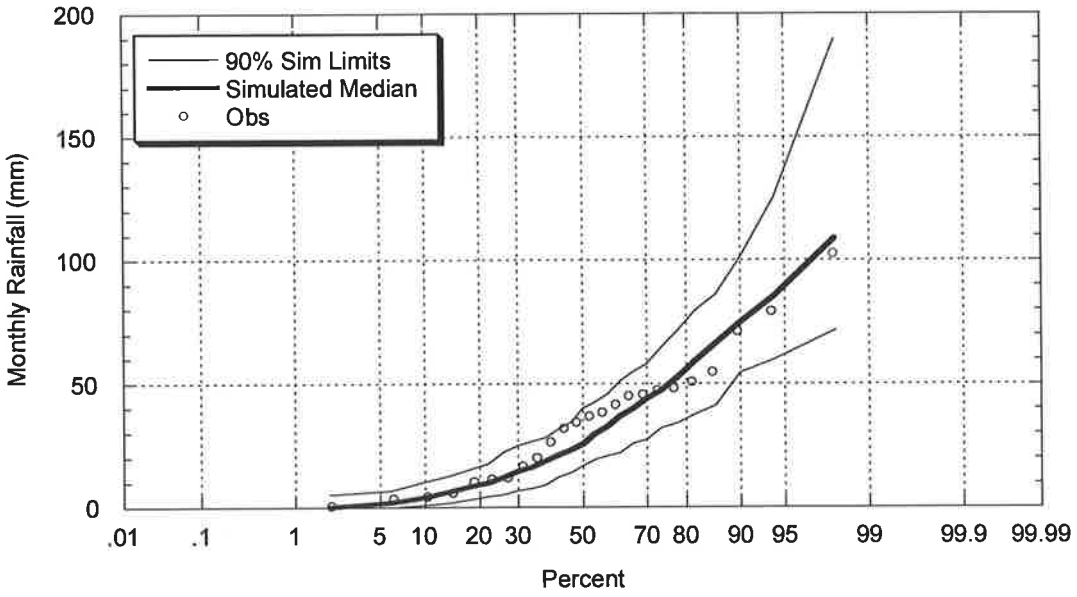

Figure C.2.39: Comparison between Observed and Target Simulated August Rainfall
Master – Brisbane Regional Office; Target – Kirkleagh)

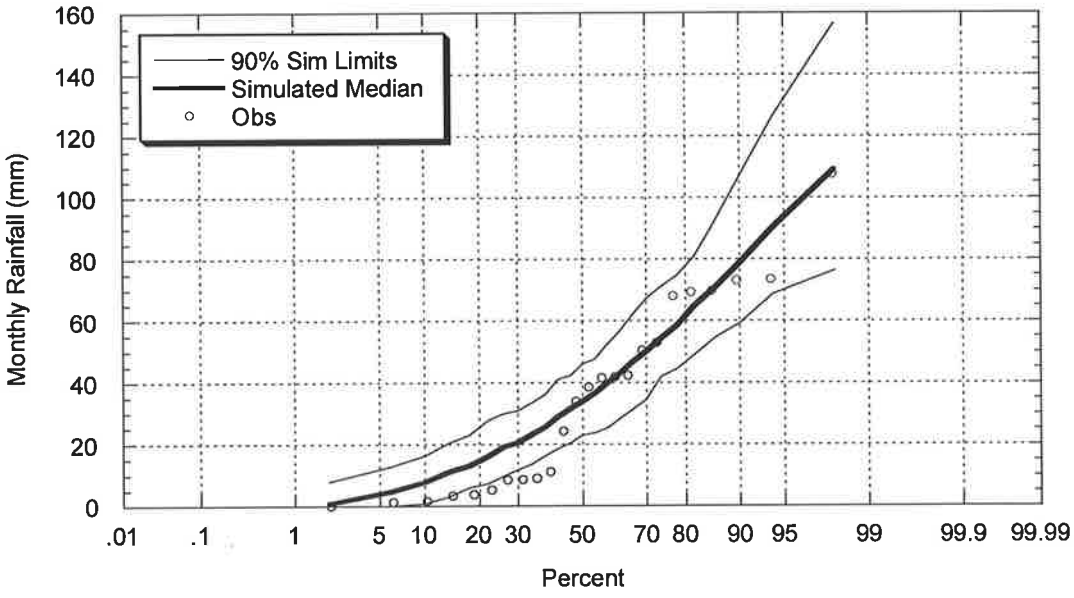

Figure C.2.40: Comparison between Observed and Target Simulated September Rainfall
Master – Brisbane Regional Office; Target – Kirkleagh)

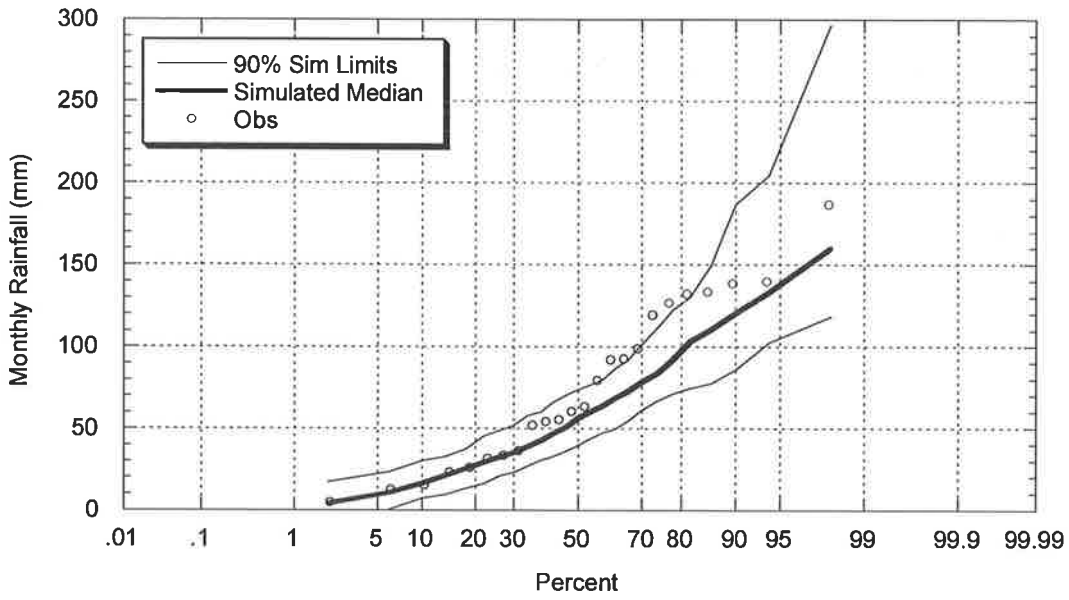

Figure C.2.41: Comparison between Observed and Target Simulated October Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

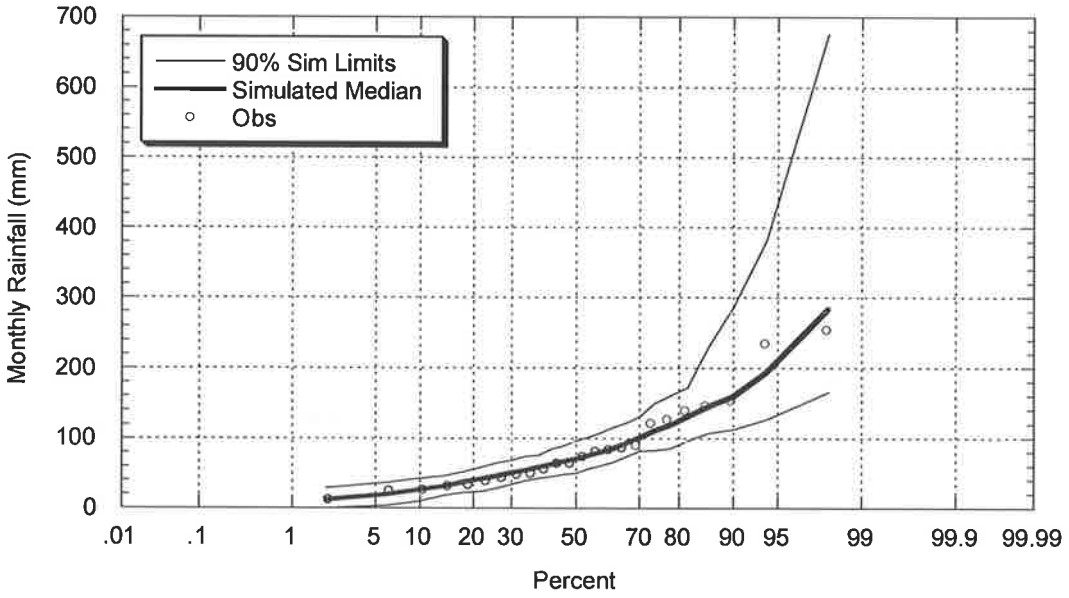

Figure C.2.42: Comparison between Observed and Target Simulated November Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

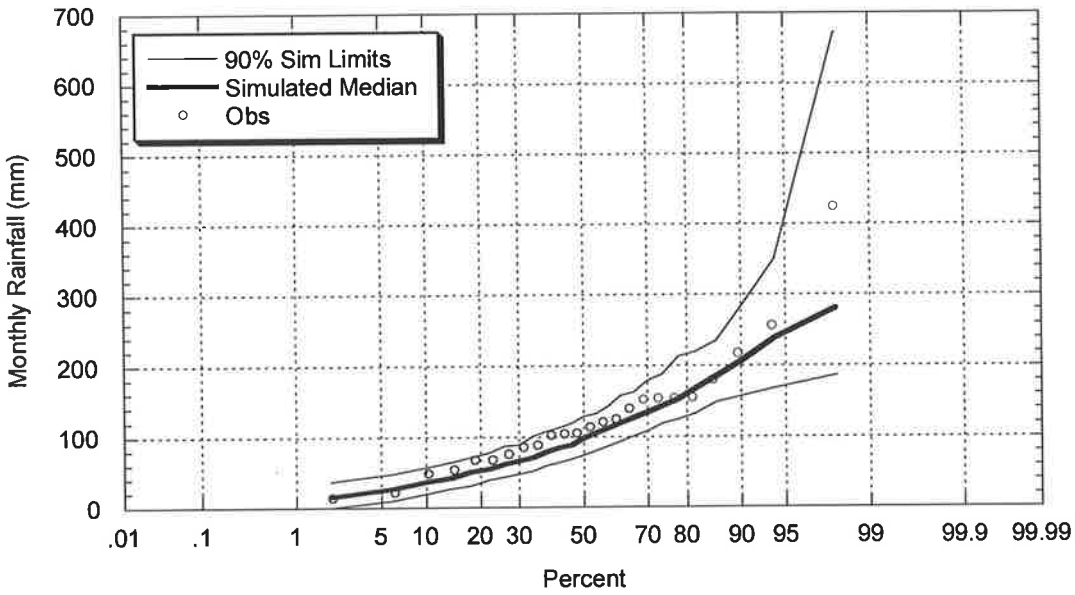

Figure C.2.43: Comparison between Observed and Target Simulated December Rainfall Master – Brisbane Regional Office; Target – Kirkleagh)

C.2.2.4 Simulated and Observed Annual Intensity – Frequency – Duration

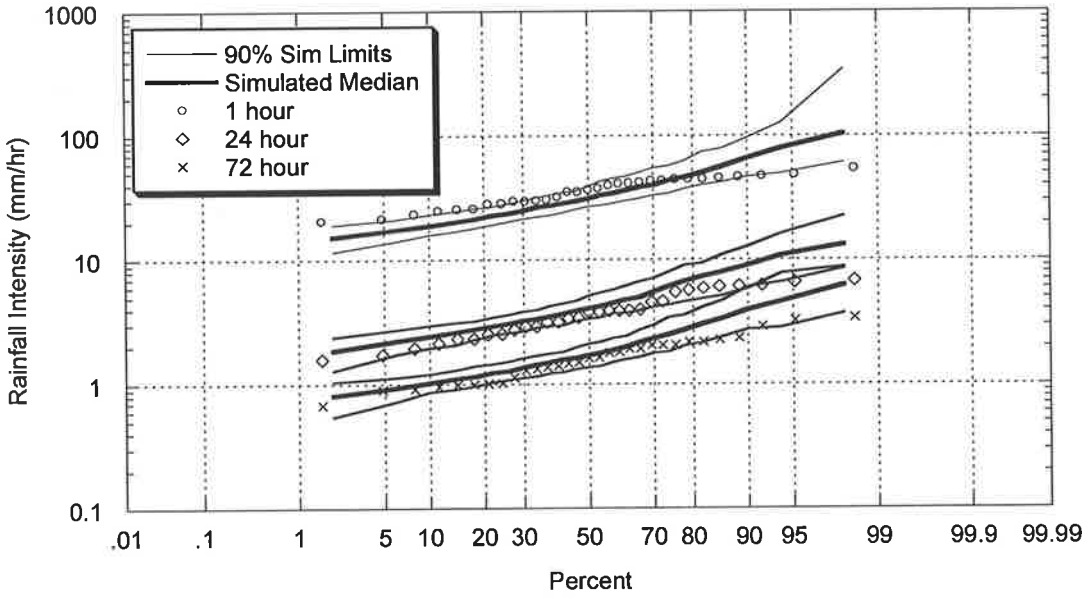

Figure C.2.44: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship Master – Brisbane Regional Office; Target – Kirkleagh)

C.3 Master – Melbourne, Victoria (BOM# 86071)

C.3.1 Target – East Sale, Victoria (BOM# 85072)

C.3.1.1 Simulated and Observed Storm Event Characteristics

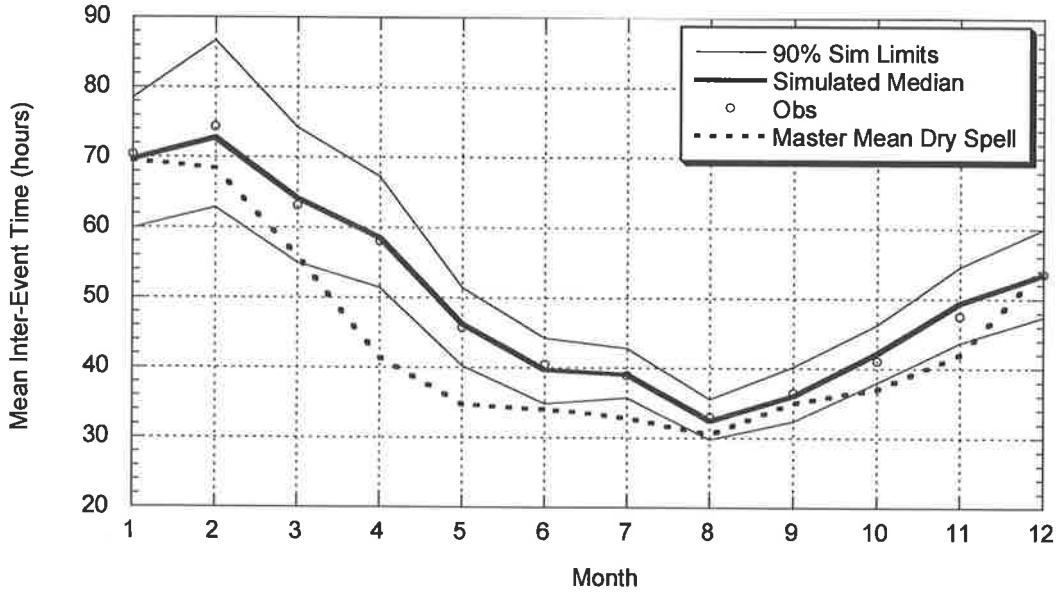

Figure C.3.1: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Melbourne; Target – East Sale)

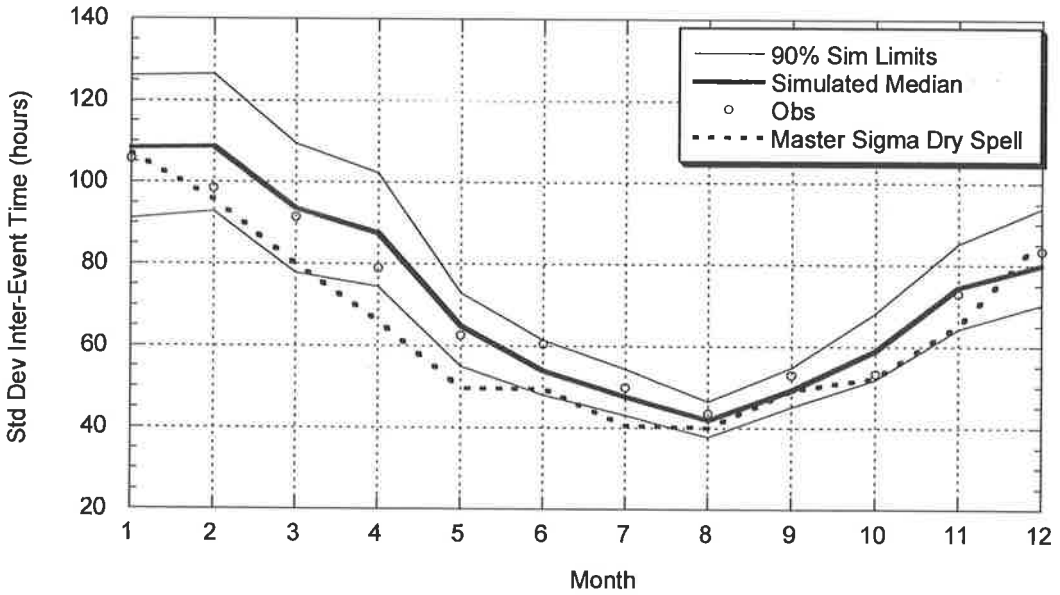

Figure C.3.2: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Melbourne; Target – East Sale)

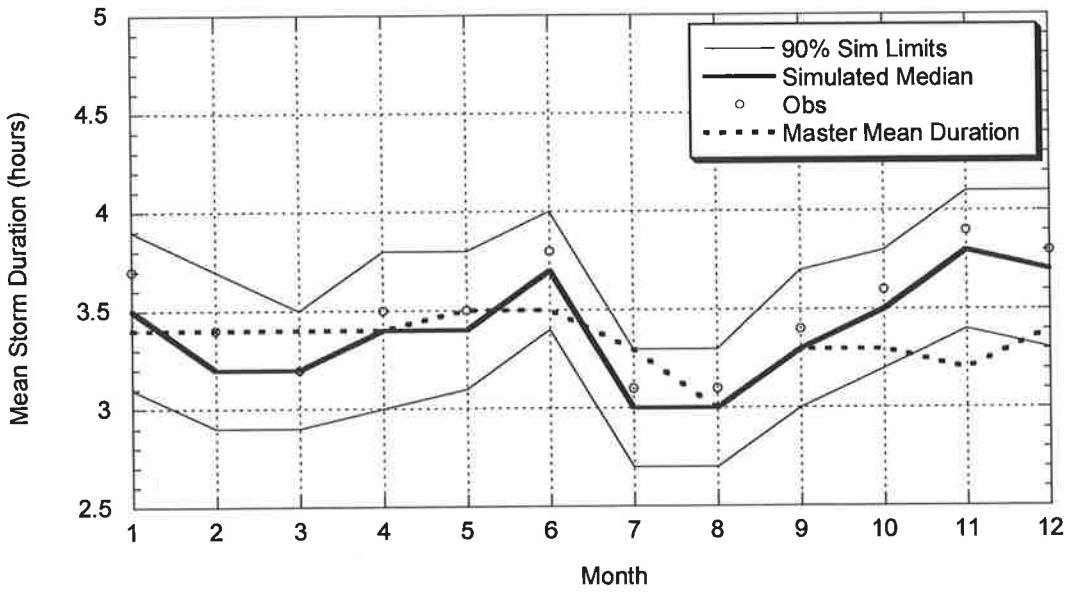

Figure C.3.3: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Melbourne; Target – East Sale)

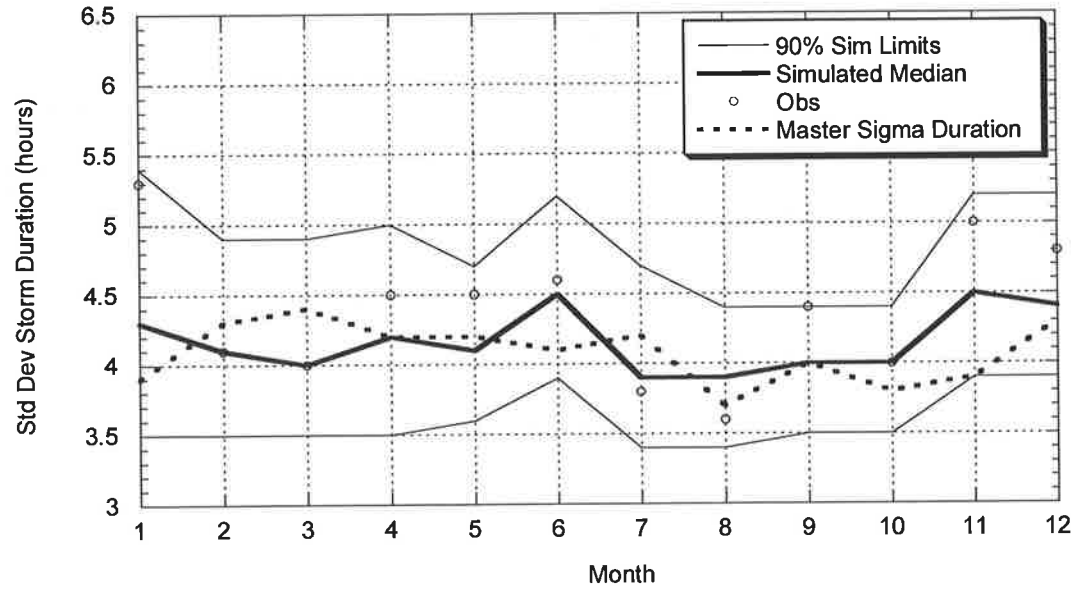

Figure C.3.4: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations (Master – Melbourne; Target – East Sale)

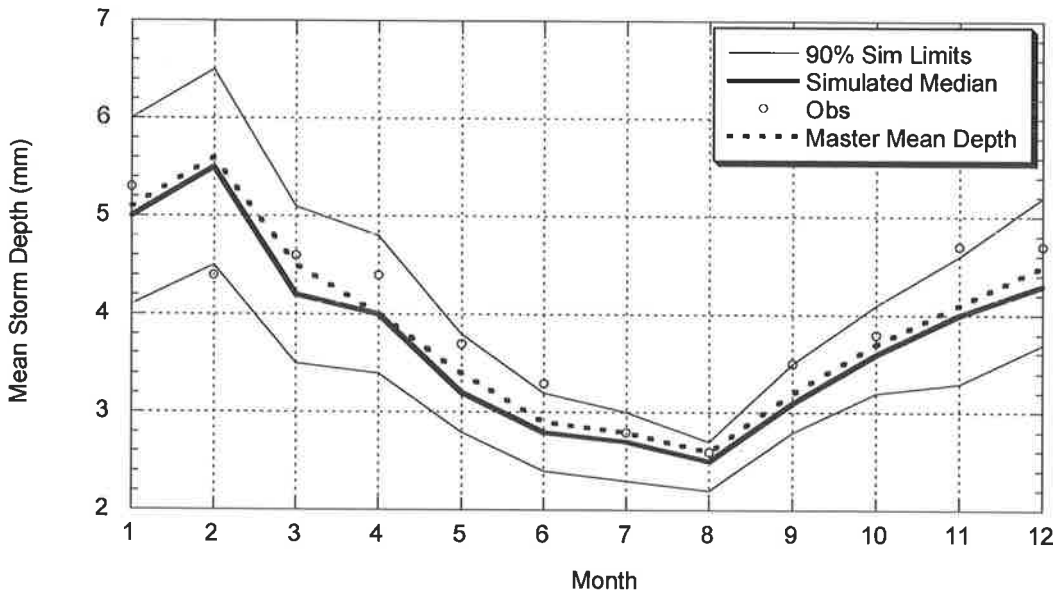

Figure C.3.5: Comparison between Observed and Target Simulated Average of Event Depths (Master – Melbourne; Target – East Sale)

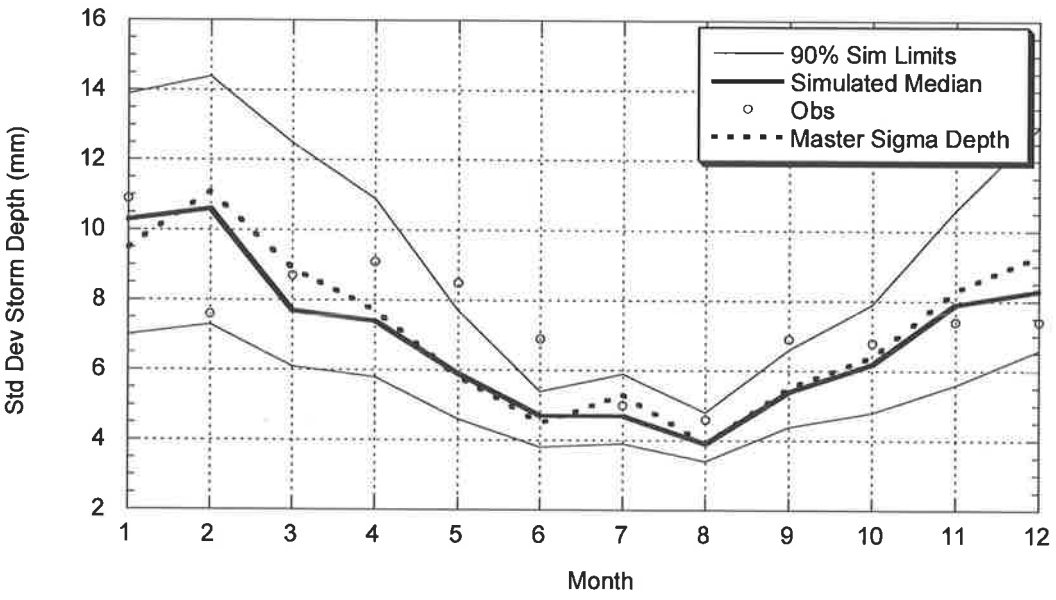

Figure C.3.6: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Melbourne; Target – East Sale)

C.3.1.2 Simulated and Observed Daily Statistics

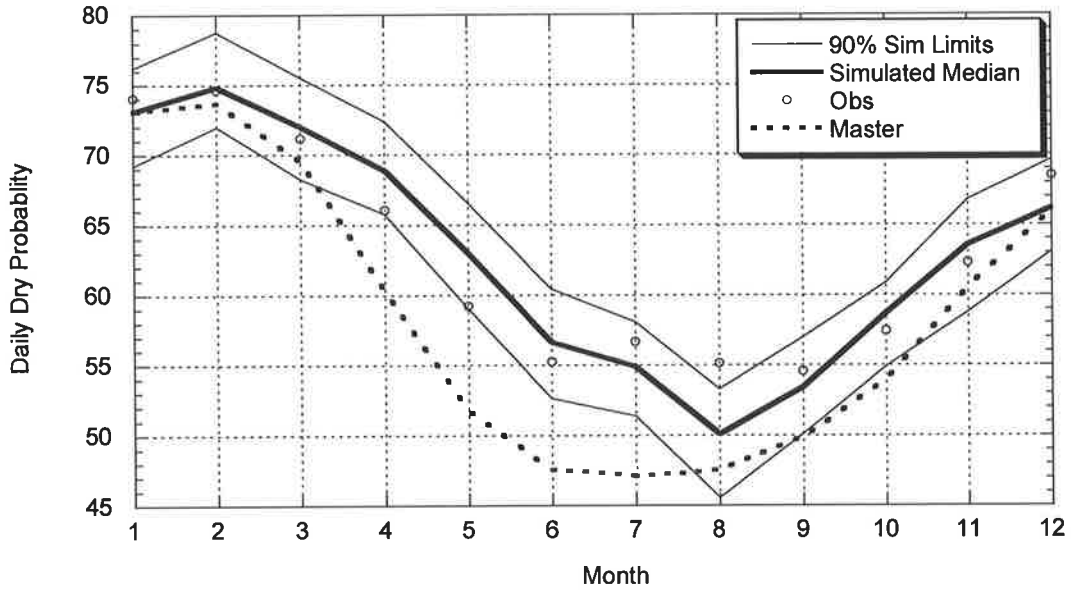

Figure C.3.7: Comparison between Observed and Target Simulated Daily Dry Probabilities (Master – Melbourne; Target – East Sale)

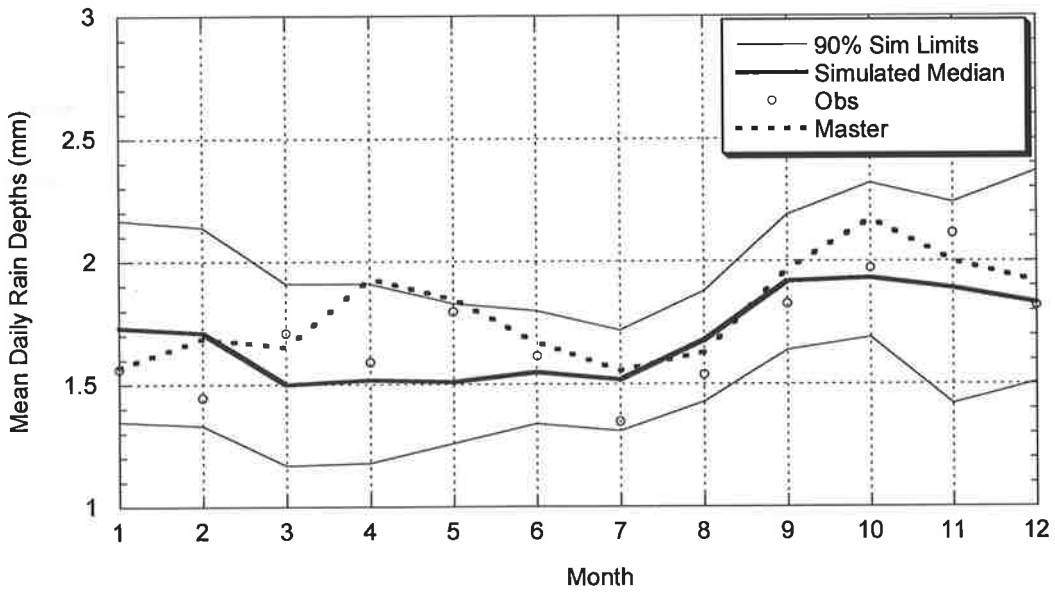

Figure C.3.8: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Melbourne; Target – East Sale)

C.3.1.3 Simulated and Observed Annual and Monthly Rainfall

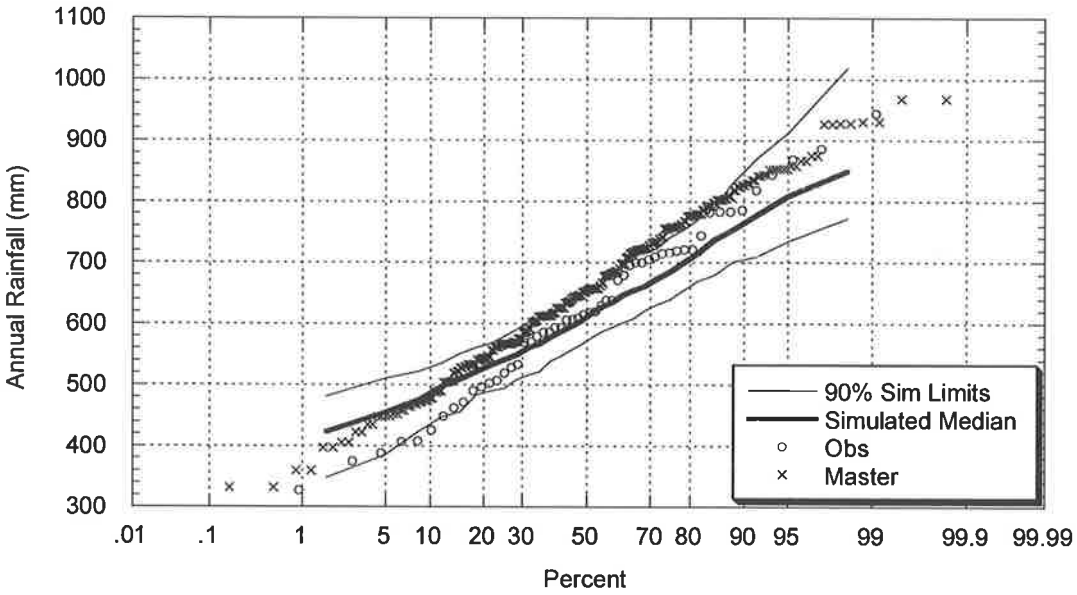

Figure C.3.9: Comparison between Observed and Target Simulated Annual Rainfall (Master – Melbourne; Target – East Sale)

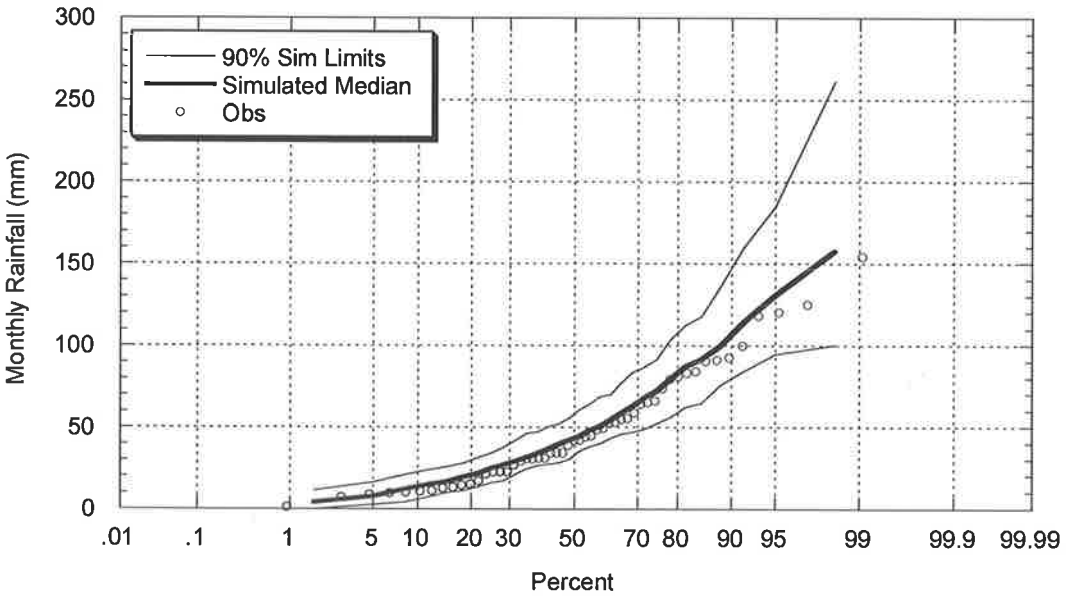

Figure C.3.10: Comparison between Observed and Target Simulated January Rainfall (Master – Melbourne; Target – East Sale)

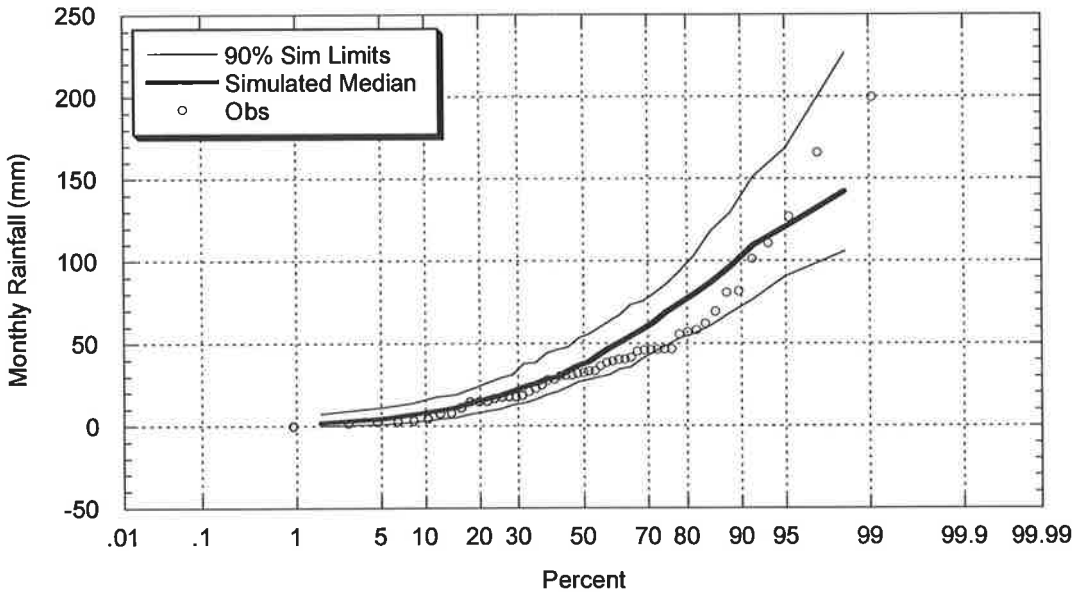

Figure C.3.11: Comparison between Observed and Target Simulated February Rainfall (Master – Melbourne; Target – East Sale)

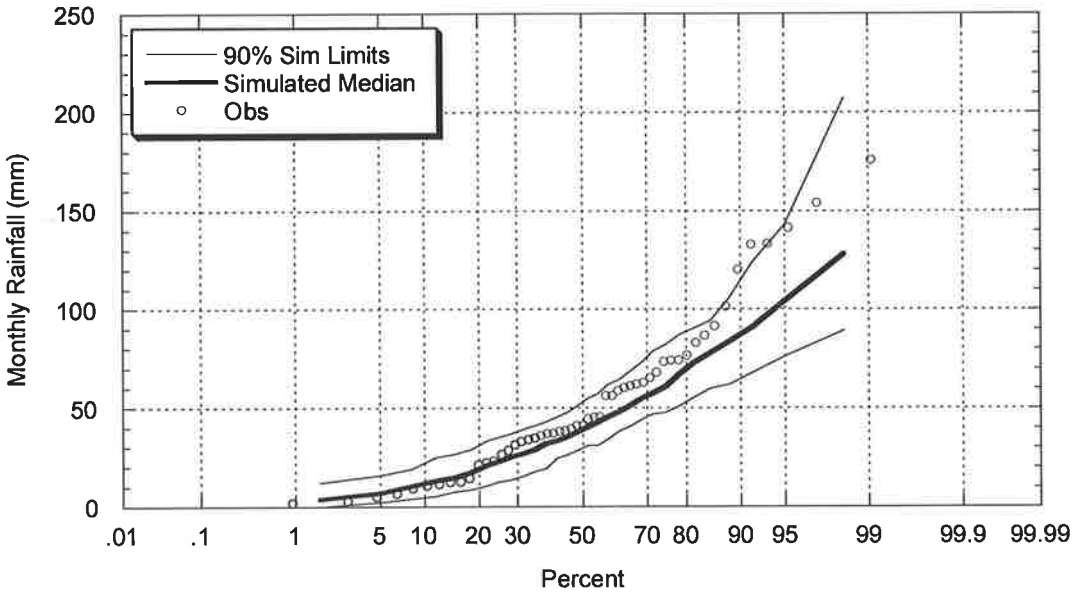

Figure C.3.12: Comparison between Observed and Target Simulated March Rainfall (Master – Melbourne; Target – East Sale)

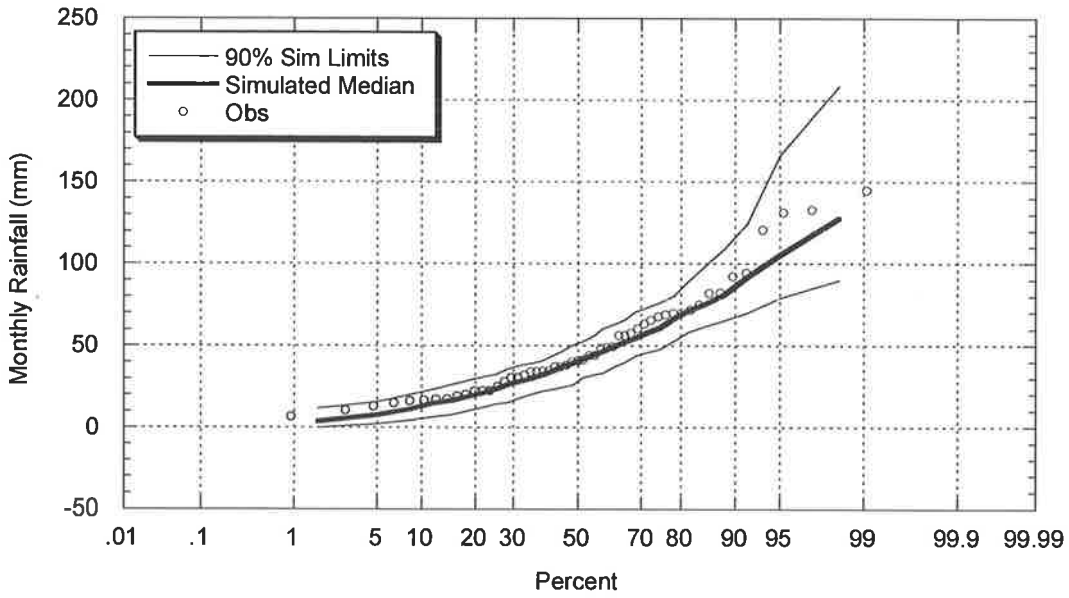

Figure C.3.13: Comparison between Observed and Target Simulated April Rainfall (Master – Melbourne; Target – East Sale)

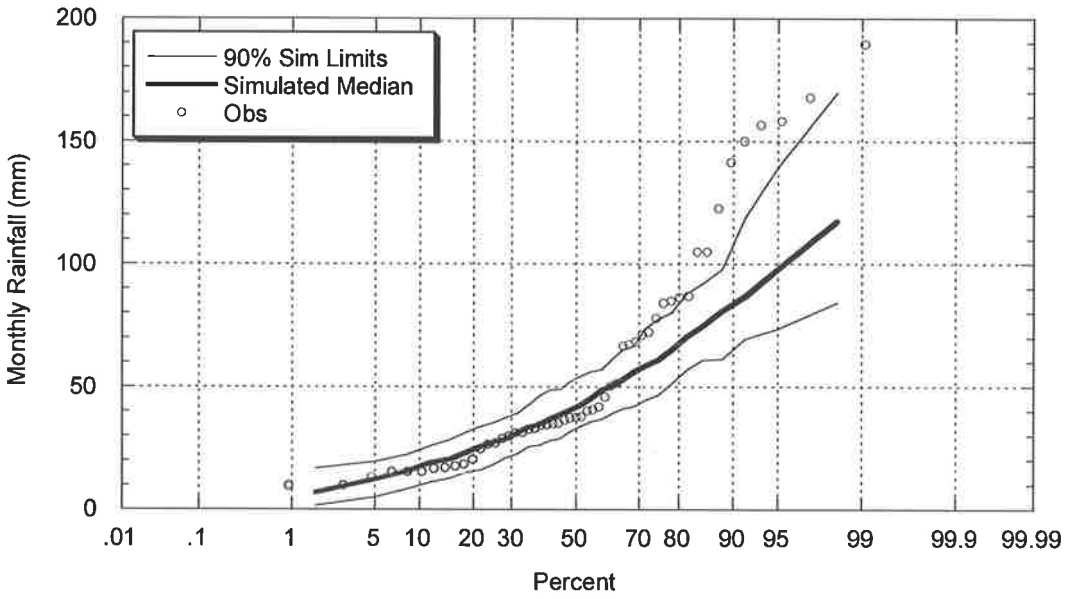

Figure C.3.14: Comparison between Observed and Target Simulated May Rainfall (Master – Melbourne; Target – East Sale)

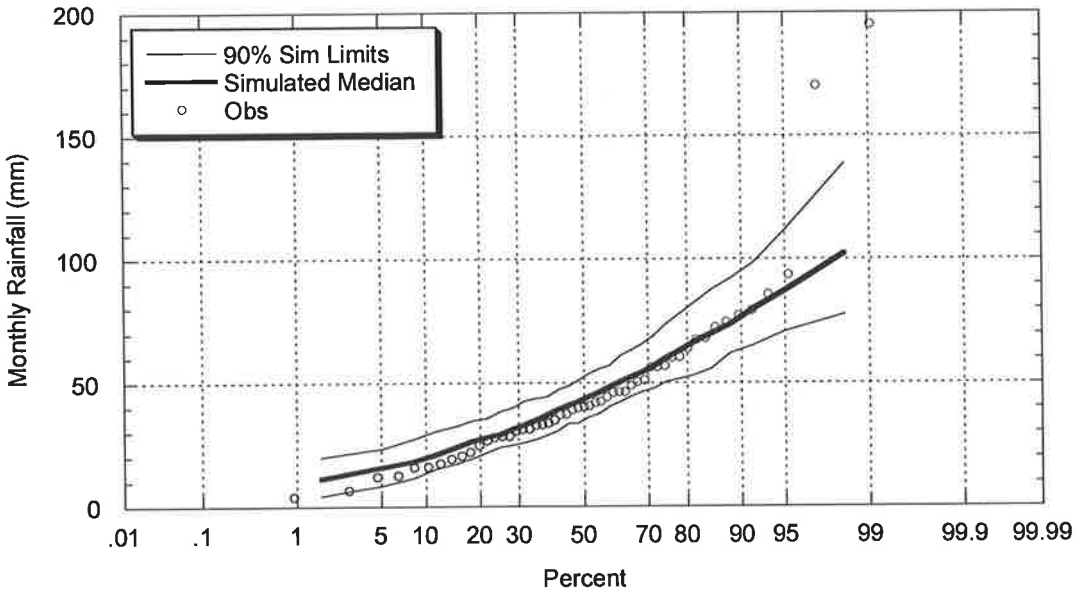

Figure C.3.15: Comparison between Observed and Target Simulated June Rainfall
(Master – Melbourne; Target – East Sale)

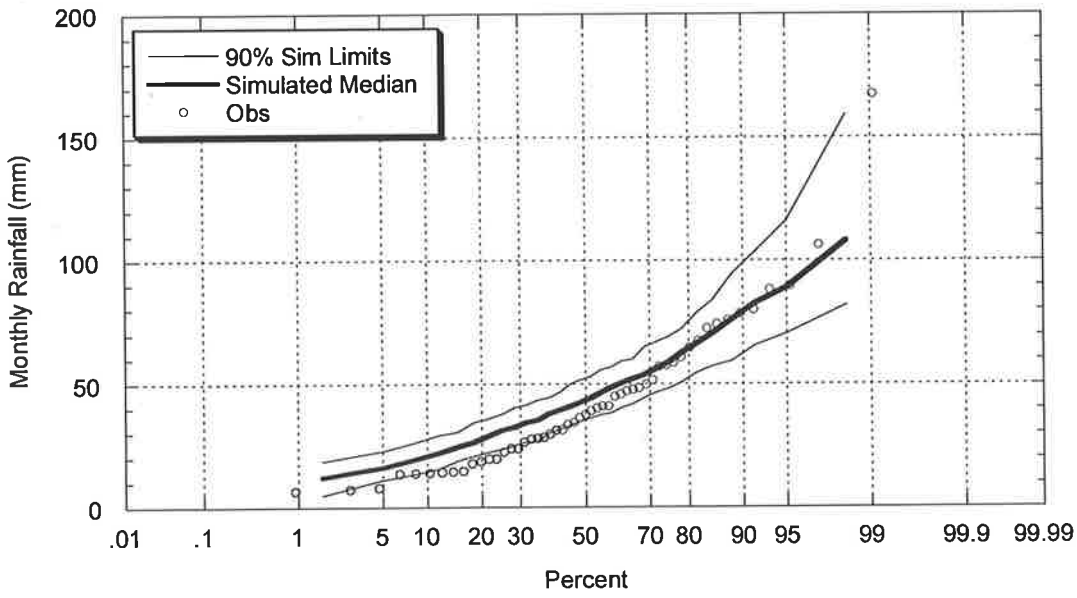

Figure C.3.16: Comparison between Observed and Target Simulated July Rainfall
(Master – Melbourne; Target – East Sale)

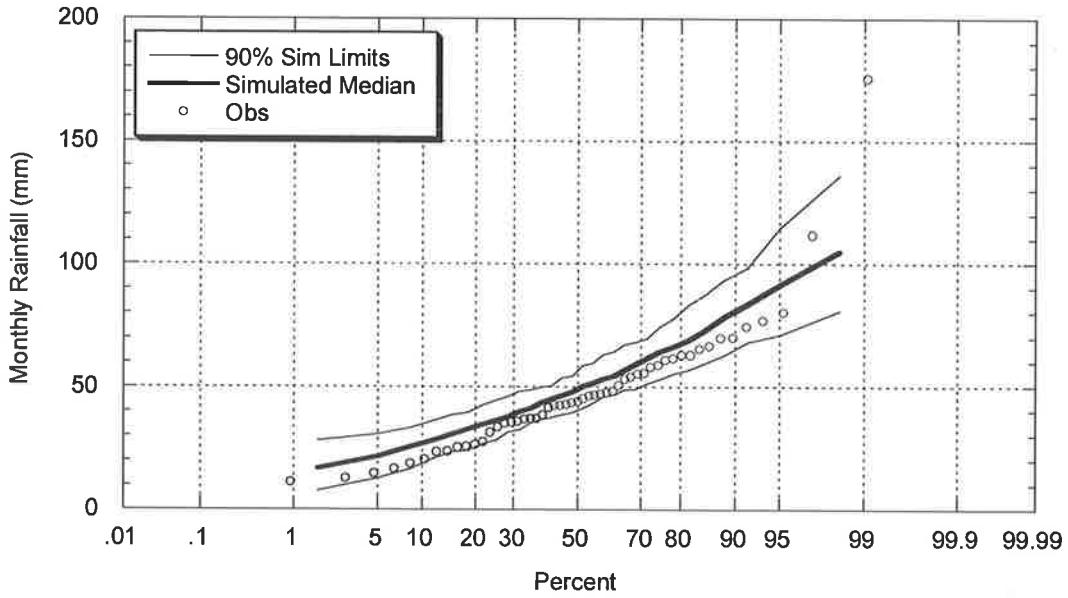

Figure C.3.17: Comparison between Observed and Target Simulated August Rainfall (Master – Melbourne; Target – East Sale)

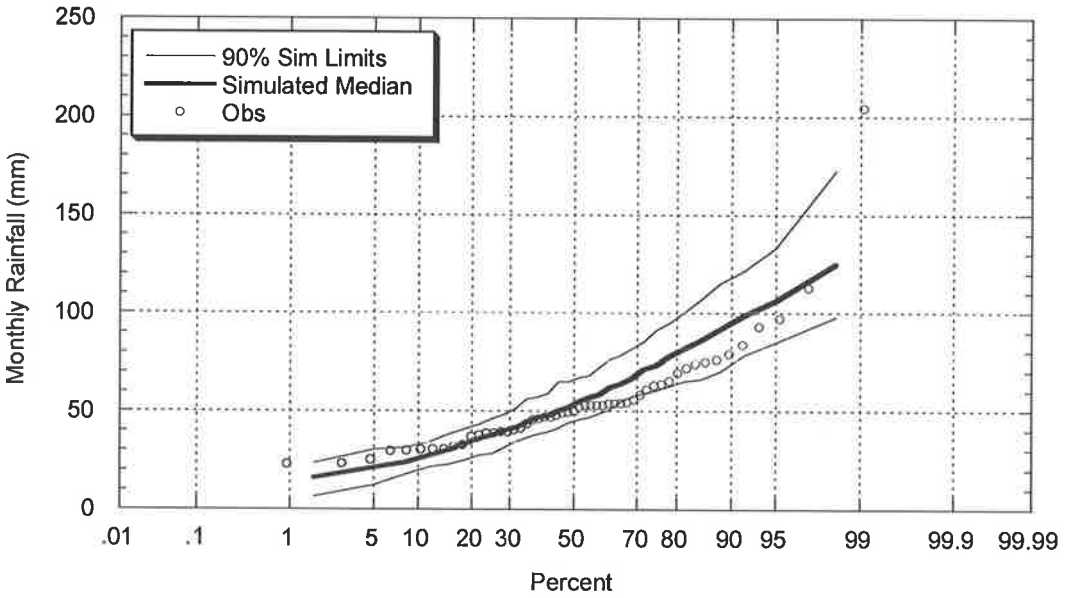

Figure C.3.18: Comparison between Observed and Target Simulated September Rainfall (Master – Melbourne; Target – East Sale)

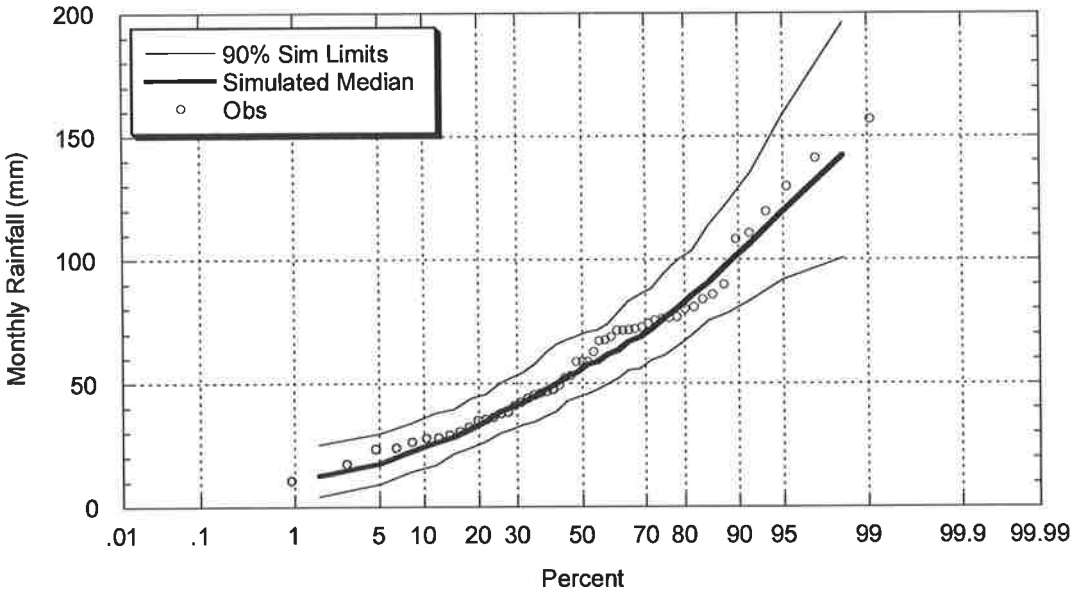

Figure C.3.19: Comparison between Observed and Target Simulated October Rainfall (Master – Melbourne; Target – East Sale)

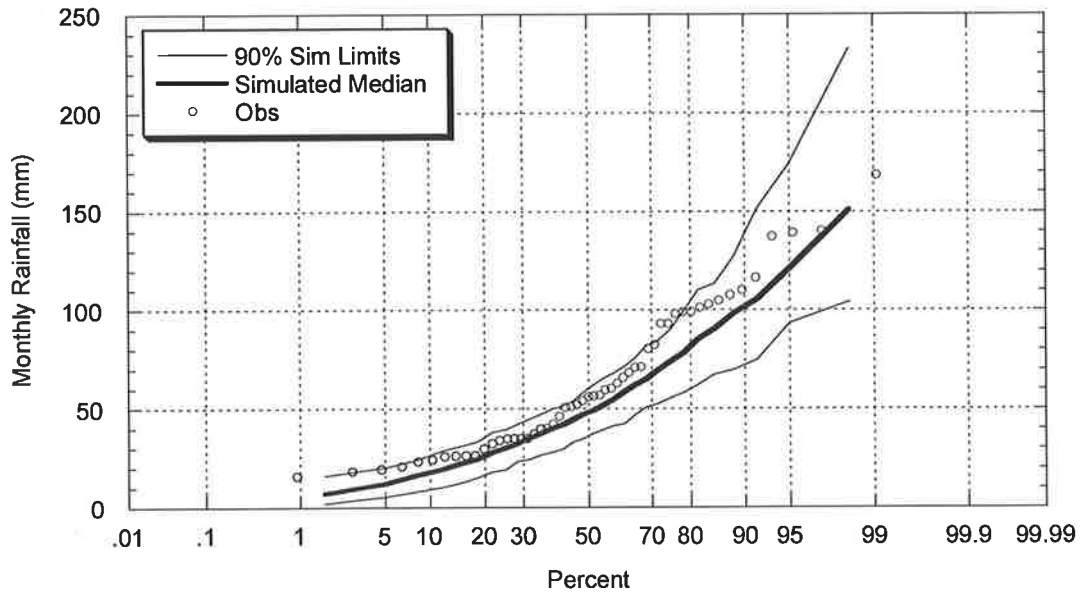

Figure C.3.20: Comparison between Observed and Target Simulated November Rainfall (Master – Melbourne; Target – East Sale)

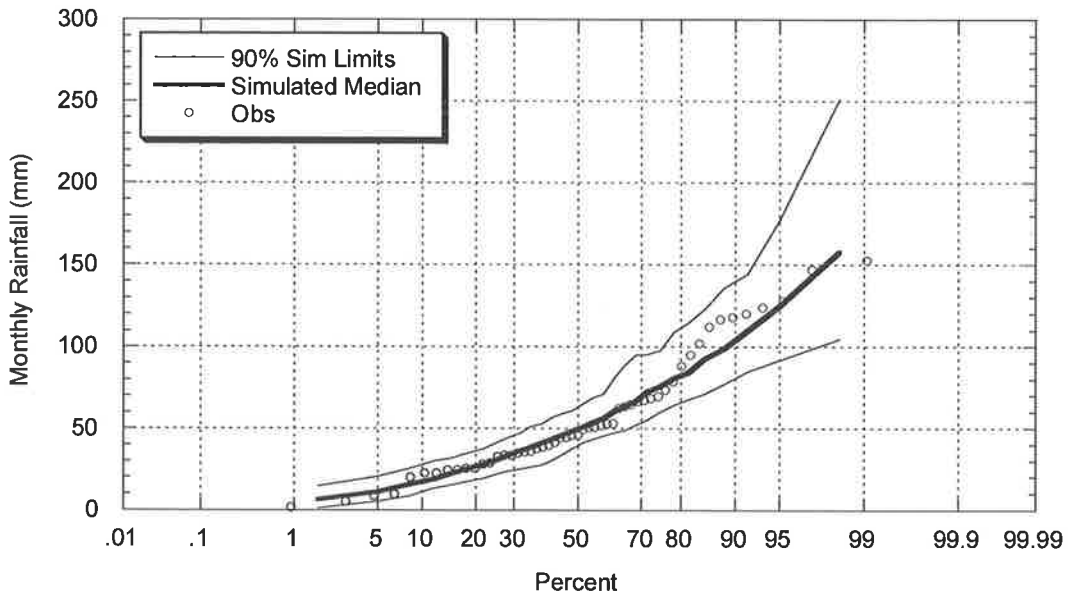

Figure C.3.21: Comparison between Observed and Target Simulated December Rainfall (Master – Melbourne; Target – East Sale)

C.3.1.4 Simulated and Observed Annual Intensity – Frequency – Duration

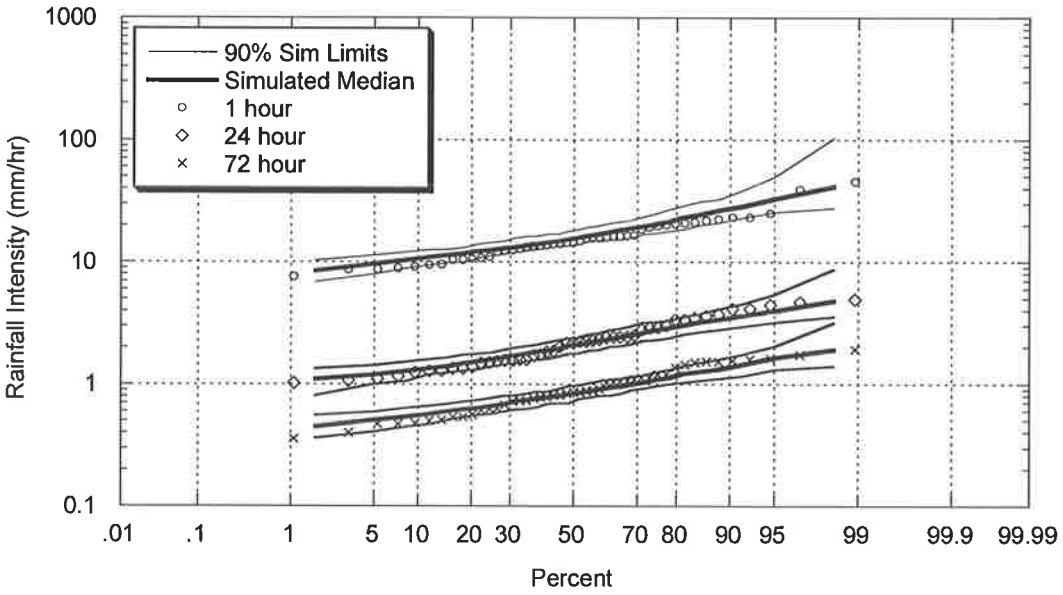

Figure C.3.22: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Melbourne; Target – East Sale)

C.3.2 Target – Ellinbank, Victoria (BOM# 85240)

C.3.2.1 Simulated and Observed Storm Event Characteristics

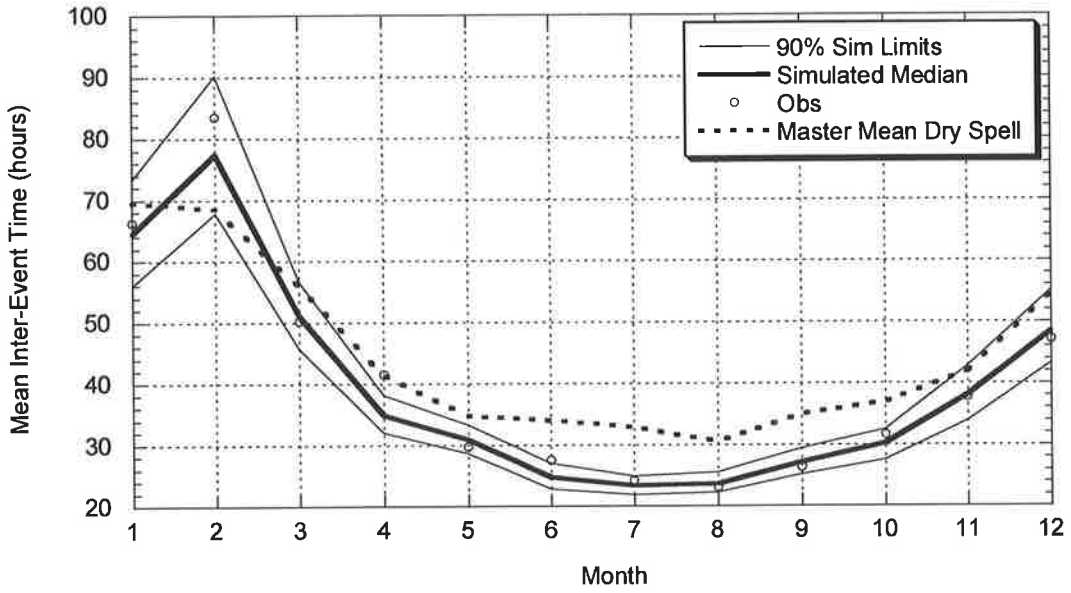

Figure C.3.23: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Melbourne; Target – Ellinbank)

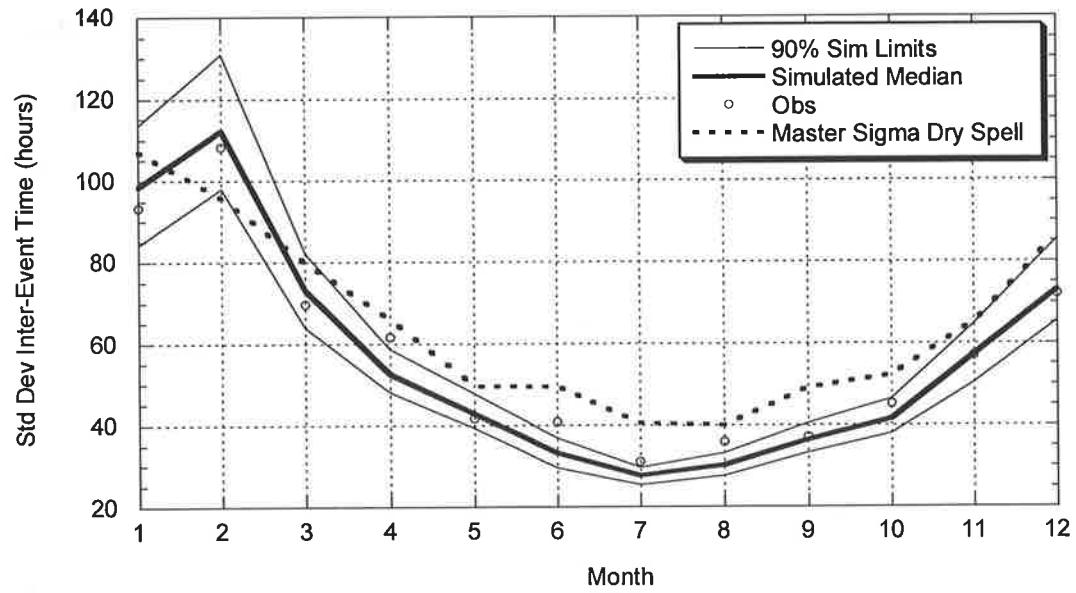

Figure C.3.24: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Melbourne; Target – Ellinbank)

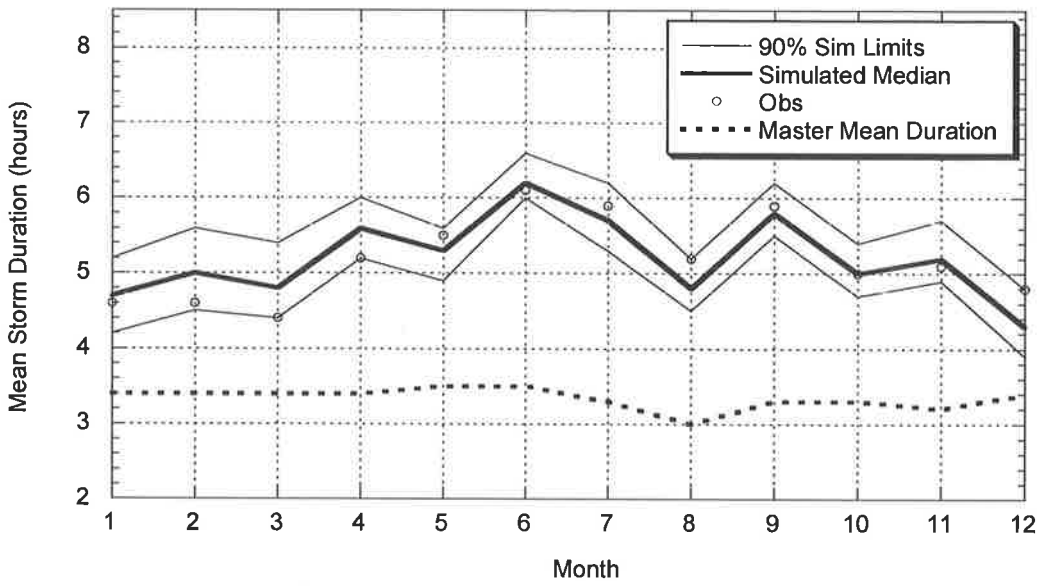

Figure C.3.25: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Melbourne; Target – Ellinbank)

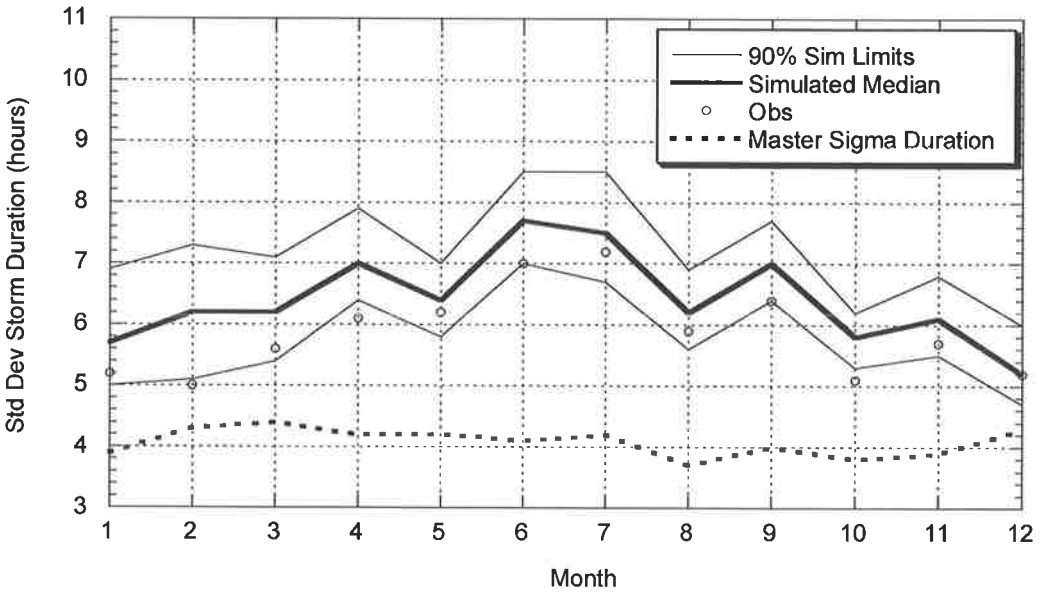

Figure C.3.26: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations (Master – Melbourne; Target – Ellinbank)

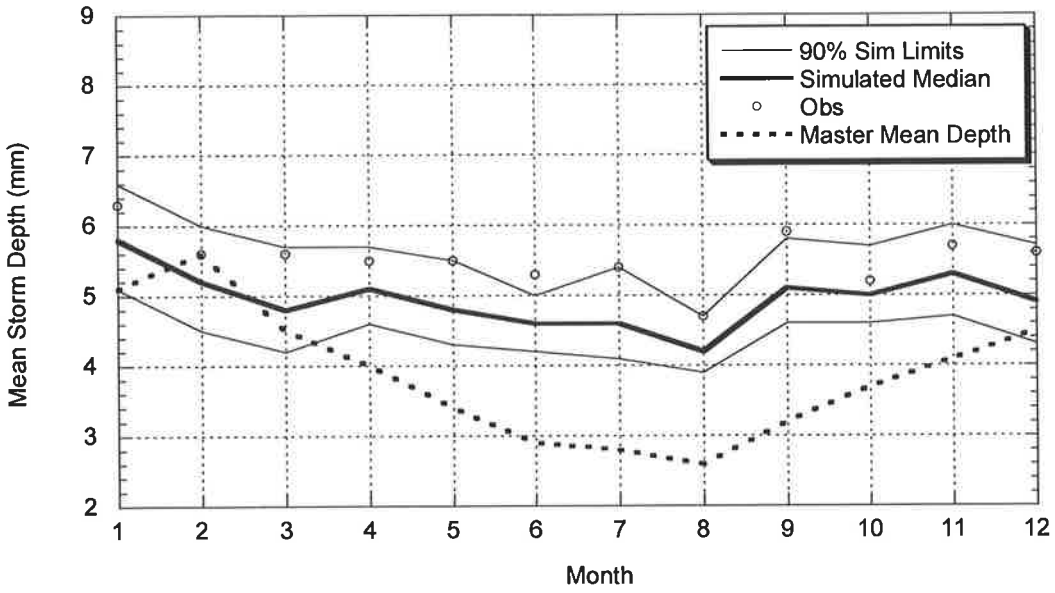

Figure C.3.27: Comparison between Observed and Target Simulated Average of Event Depths (Master – Melbourne; Target – Ellinbank)

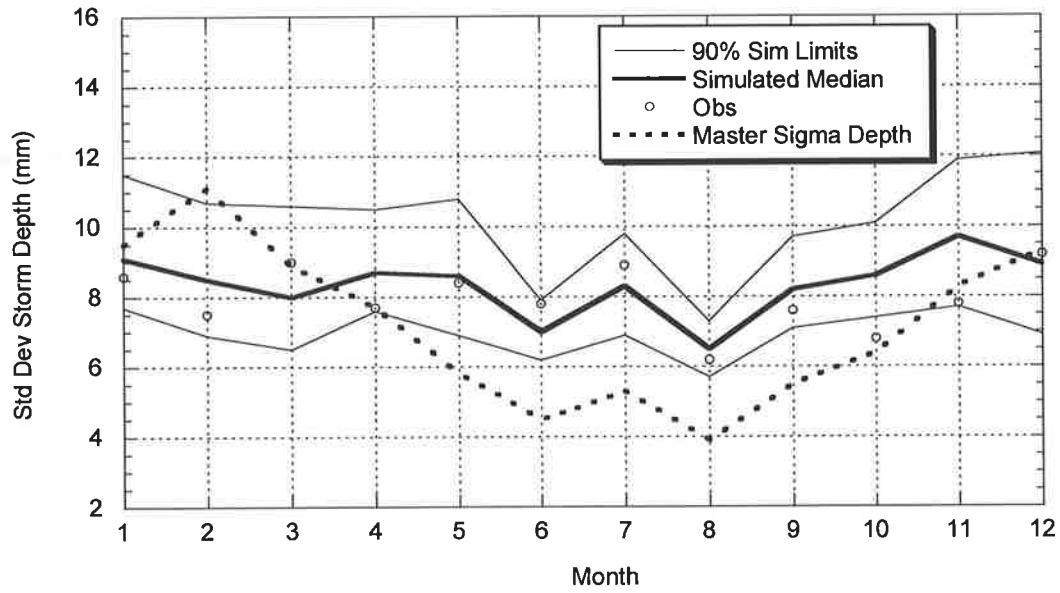

Figure C.3.28: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Melbourne; Target – Ellinbank)

C.3.2.2 Simulated and Observed Daily Statistics

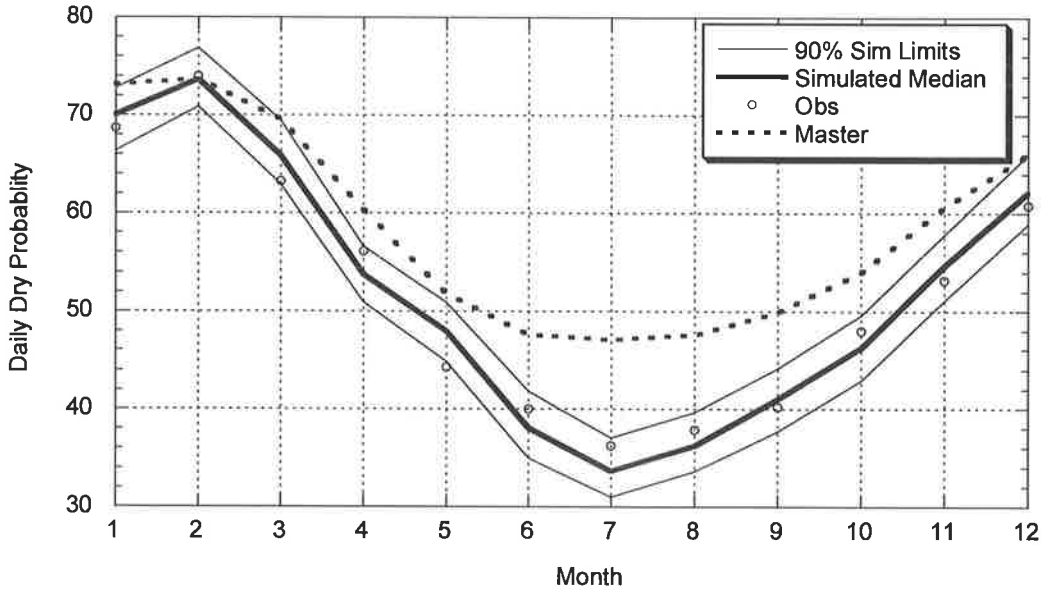

Figure C.3.29: Comparison between Observed and Target Simulated Daily Dry Probabilities (Master – Melbourne; Target – Ellinbank)

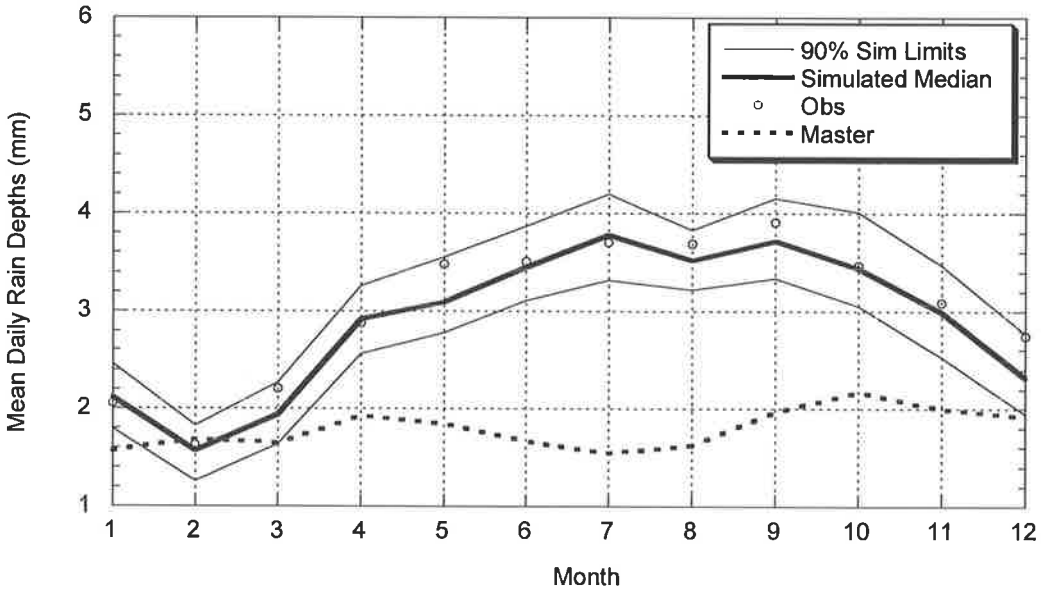

Figure C.3.30: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Melbourne; Target – Ellinbank)

C.3.2.3 Simulated and Observed Annual and Monthly Rainfall

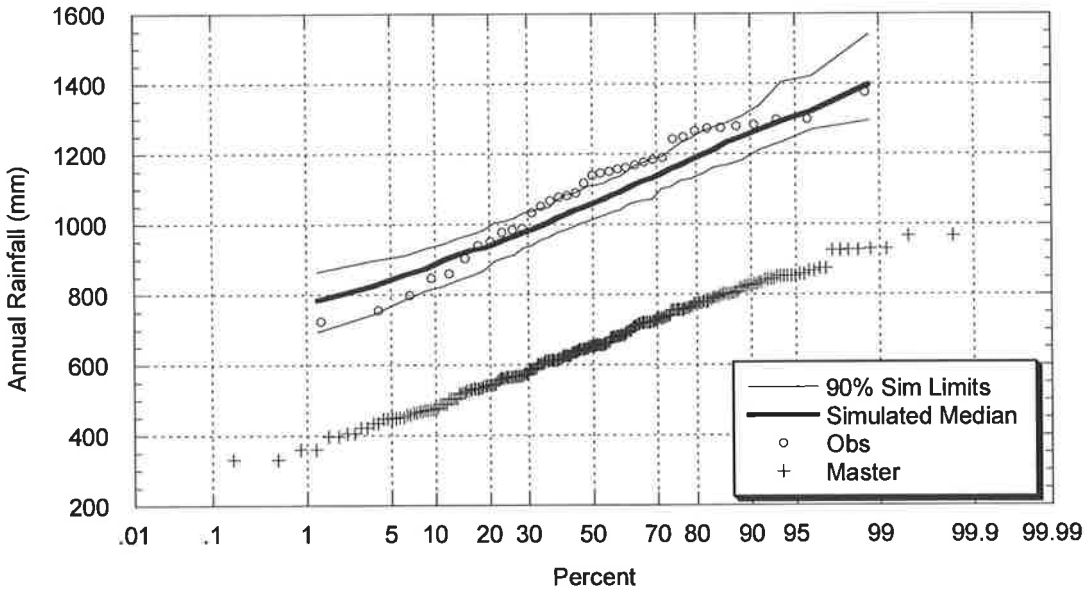

Figure C.3.31: Comparison between Observed and Target Simulated Annual Rainfall (Master – Melbourne; Target – Ellinbank)

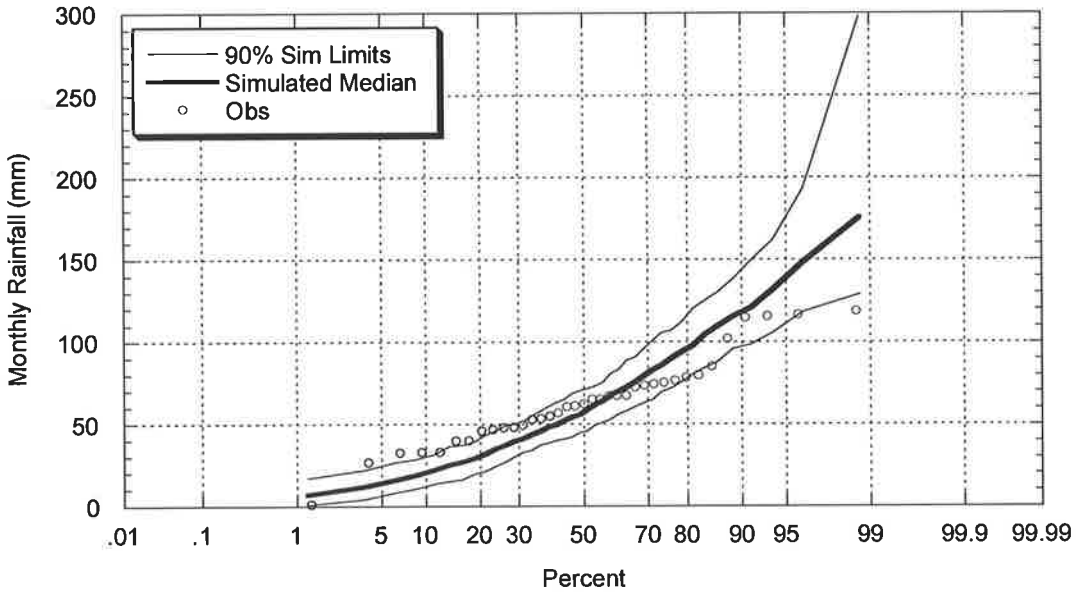

Figure C.3.32: Comparison between Observed and Target Simulated January Rainfall (Master – Melbourne; Target – Ellinbank)

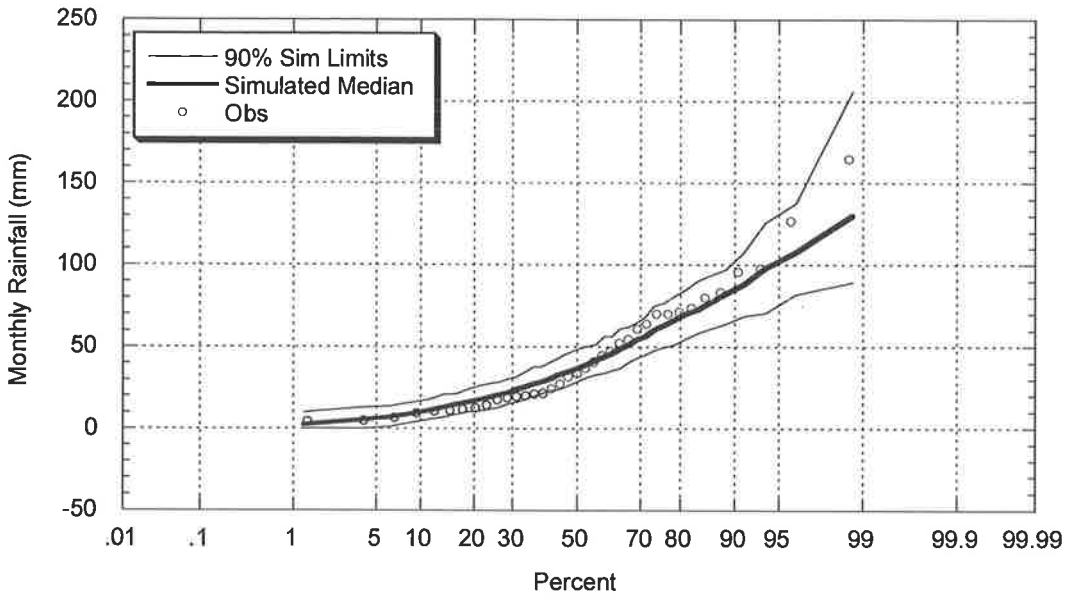

Figure C.3.33: Comparison between Observed and Target Simulated February Rainfall (Master – Melbourne; Target – Ellinbank)

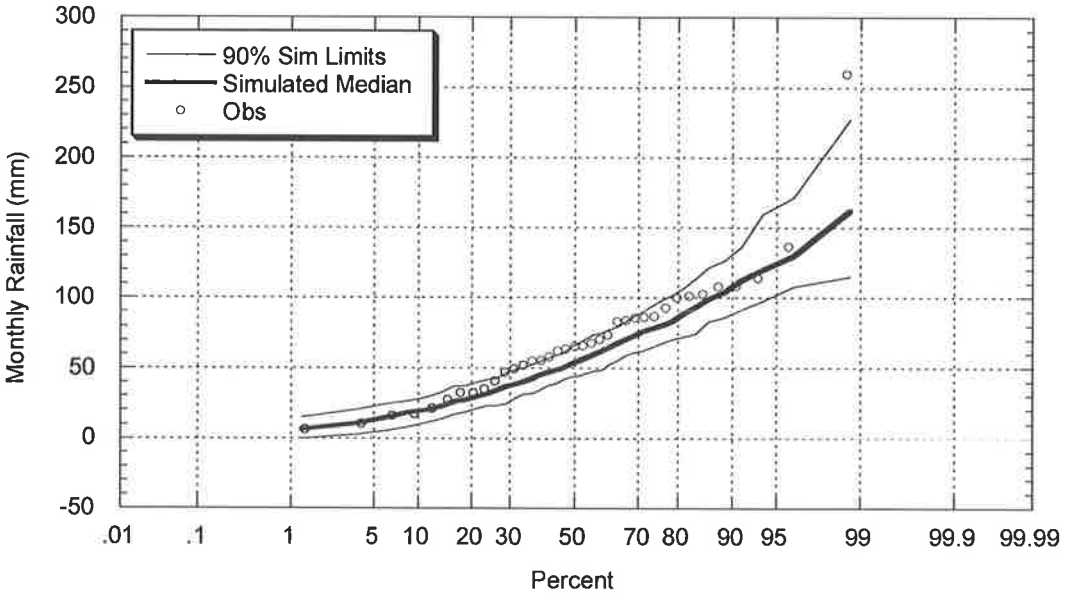

Figure C.3.34: Comparison between Observed and Target Simulated March Rainfall (Master – Melbourne; Target – Ellinbank)

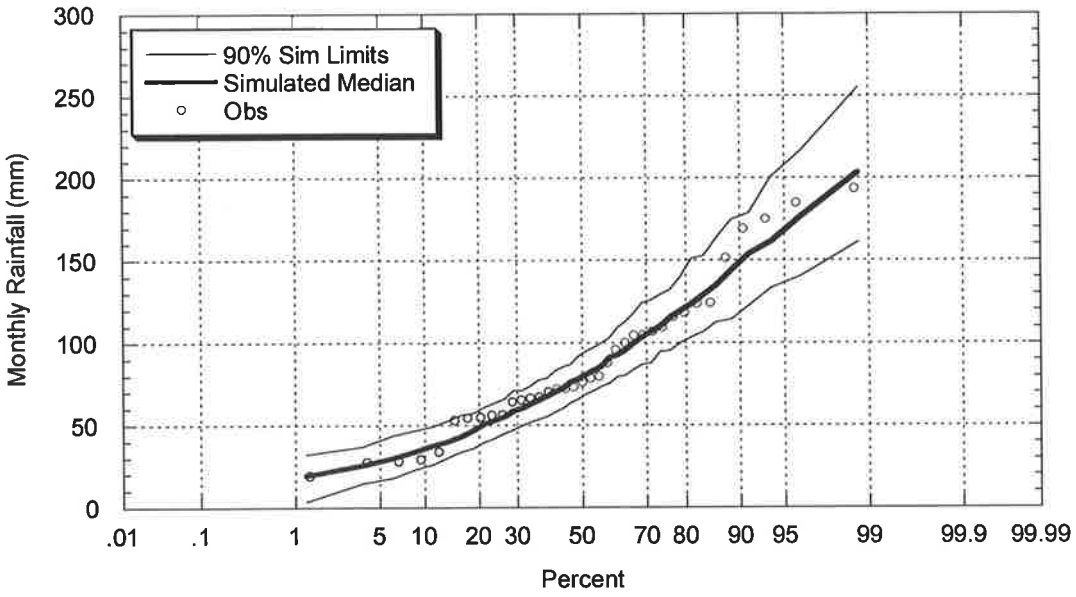

Figure C.3.35: Comparison between Observed and Target Simulated April Rainfall
(Master – Melbourne; Target – Ellinbank)

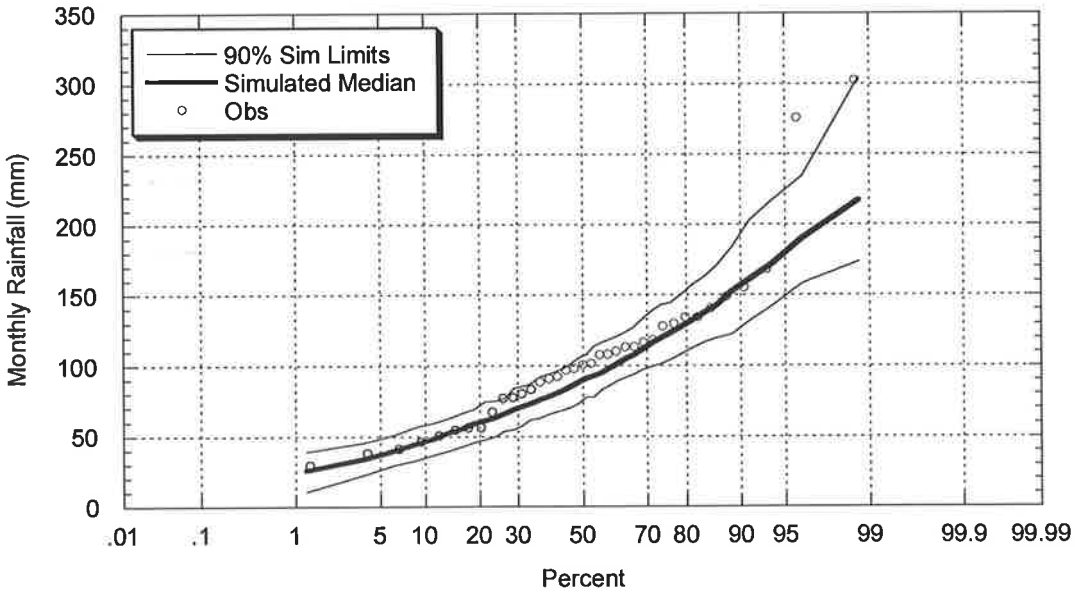

Figure C.3.36: Comparison between Observed and Target Simulated May Rainfall
(Master – Melbourne; Target – Ellinbank)

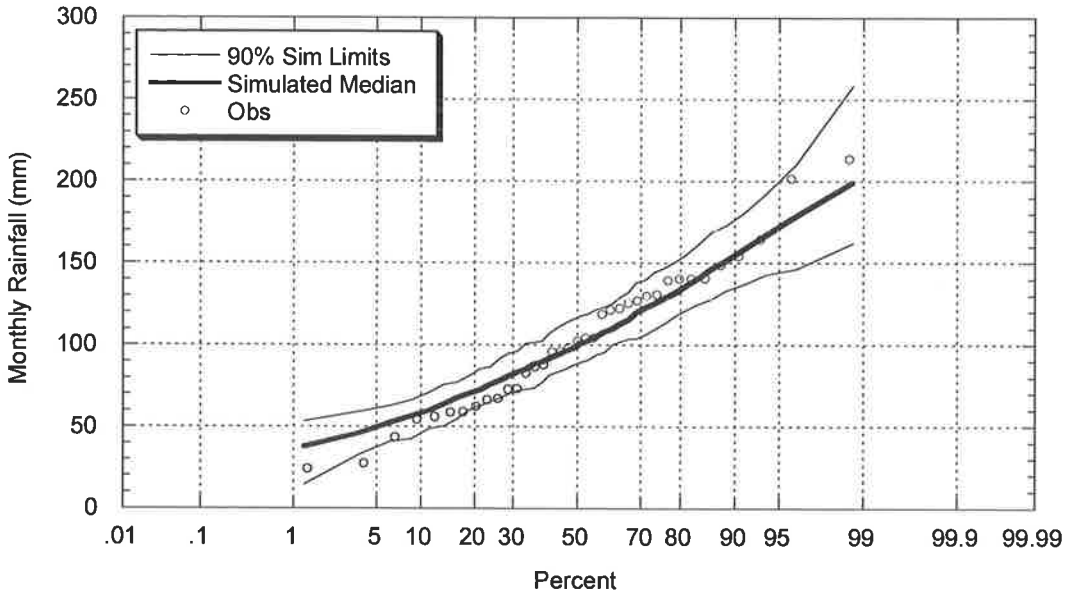

Figure C.3.37: Comparison between Observed and Target Simulated June Rainfall (Master – Melbourne; Target – Ellinbank)

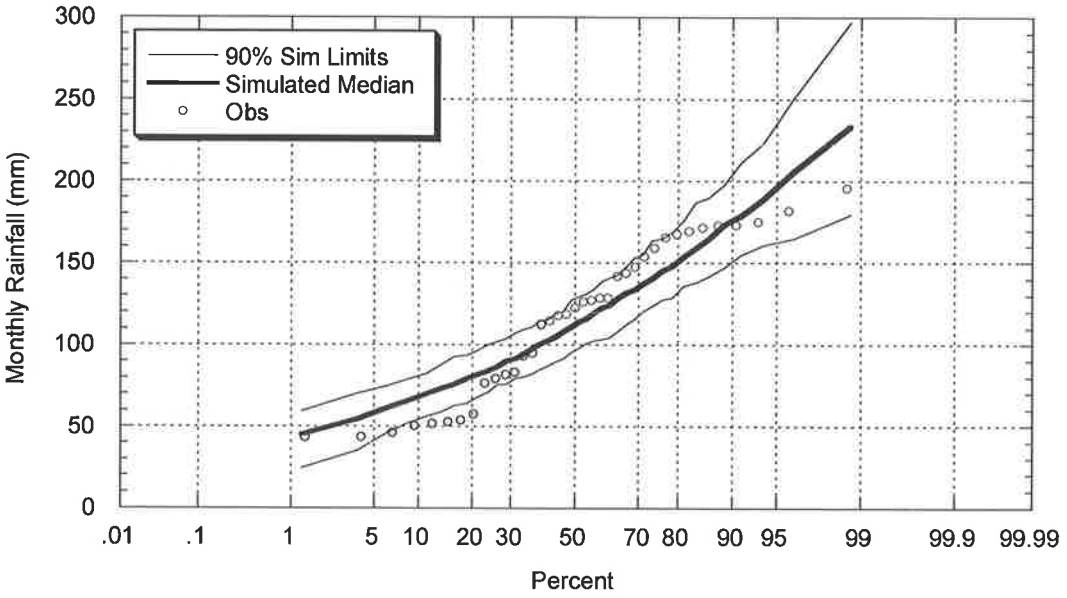

Figure C.3.38: Comparison between Observed and Target Simulated July Rainfall (Master – Melbourne; Target – Ellinbank)

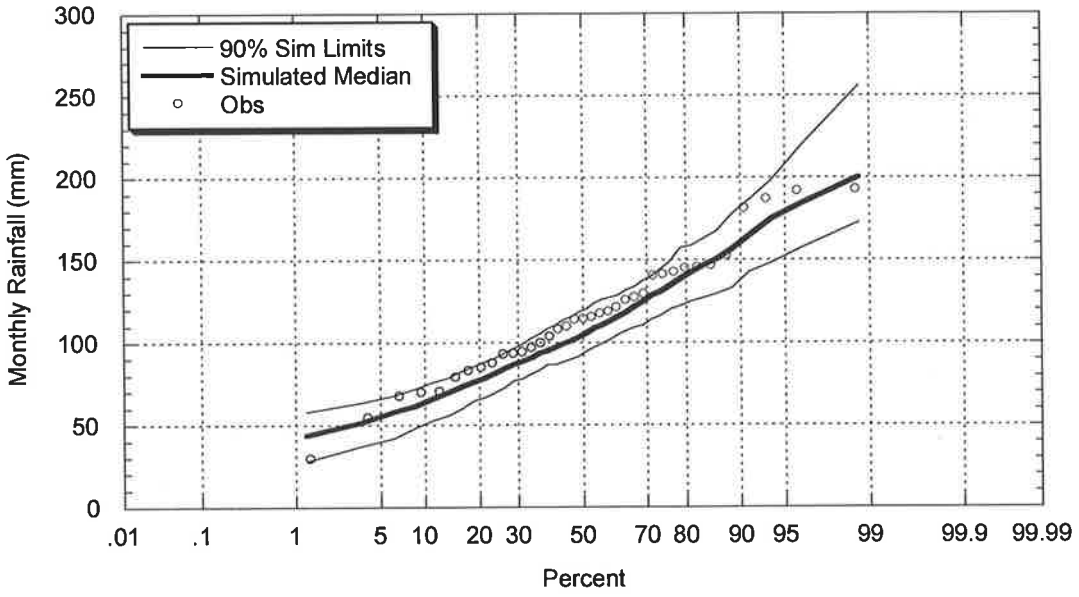

Figure C.3.39: Comparison between Observed and Target Simulated August Rainfall
(Master – Melbourne; Target – Ellinbank)

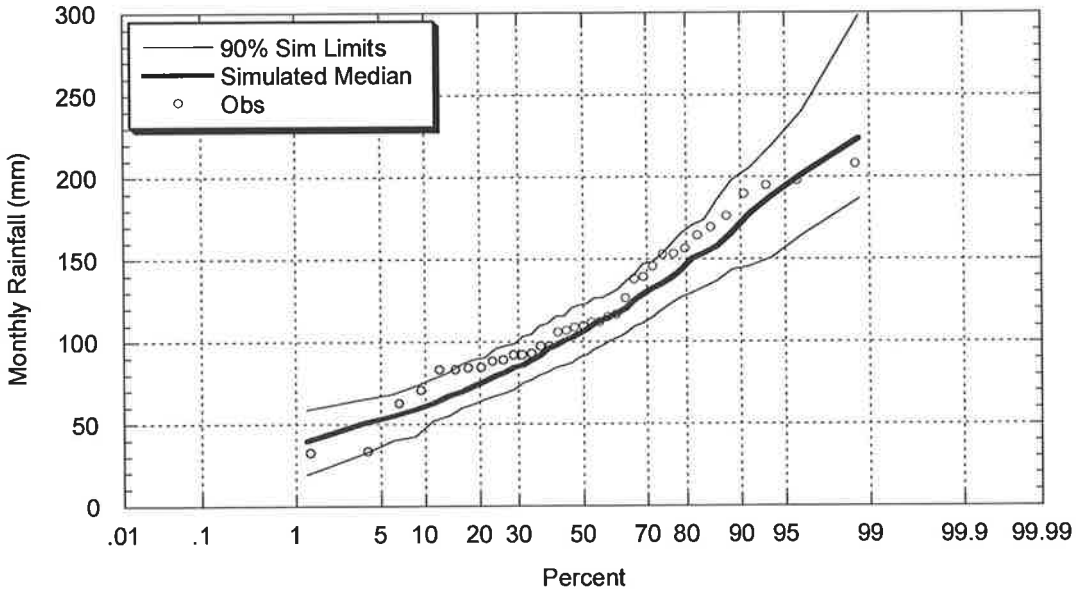

Figure C.3.40: Comparison between Observed and Target Simulated September Rainfall
(Master – Melbourne; Target – Ellinbank)

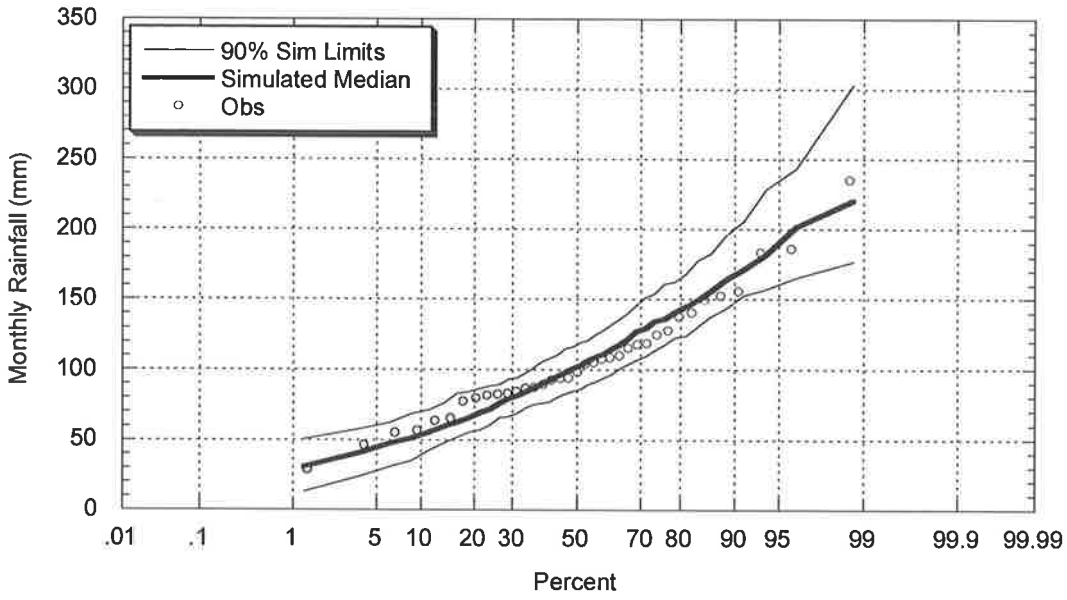

Figure C.3.41: Comparison between Observed and Target Simulated October Rainfall (Master – Melbourne; Target – Ellinbank)

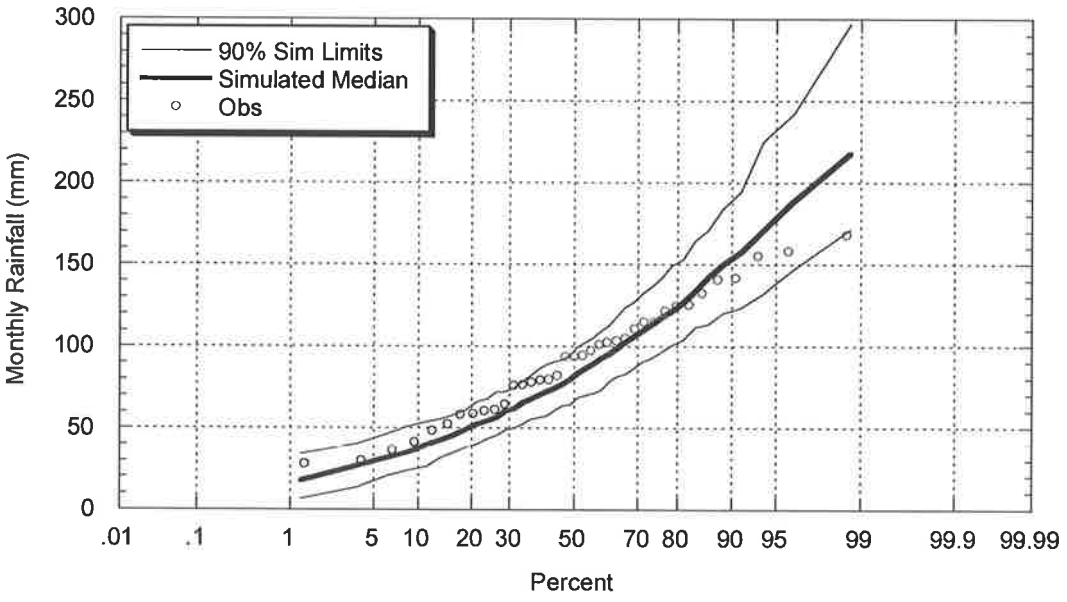

Figure C.3.42: Comparison between Observed and Target Simulated November Rainfall (Master – Melbourne; Target – Ellinbank)

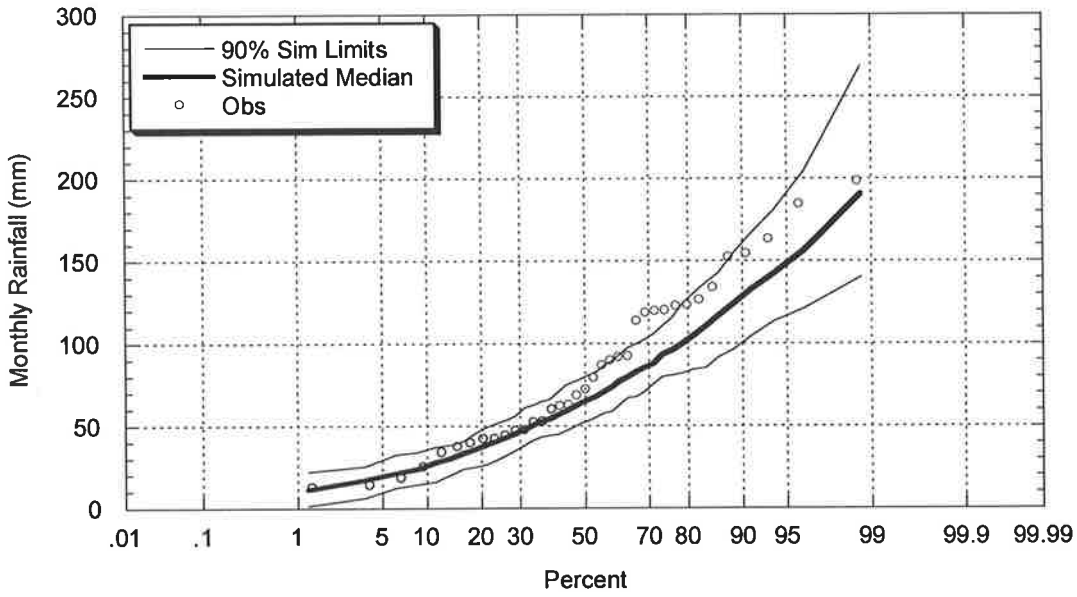

Figure C.3.43: Comparison between Observed and Target Simulated December Rainfall (Master – Melbourne; Target – Ellinbank)

C.3.2.4 Simulated and Observed Annual Intensity – Frequency – Duration

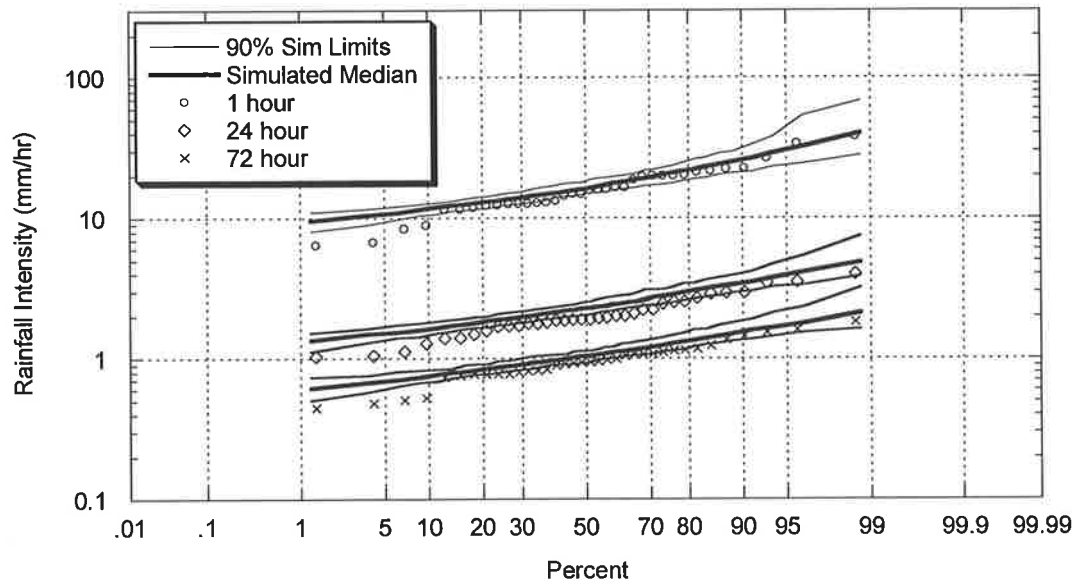

Figure C.3.44: Comparison between Observed and Target Simulated Annual Intensity Frequency Duration Relationship (Master – Melbourne; Target – Ellinbank)

C.3.3 Target – Laverton, Victoria (BOM# 87031)

C.3.3.1 Simulated and Observed Storm Event Characteristics

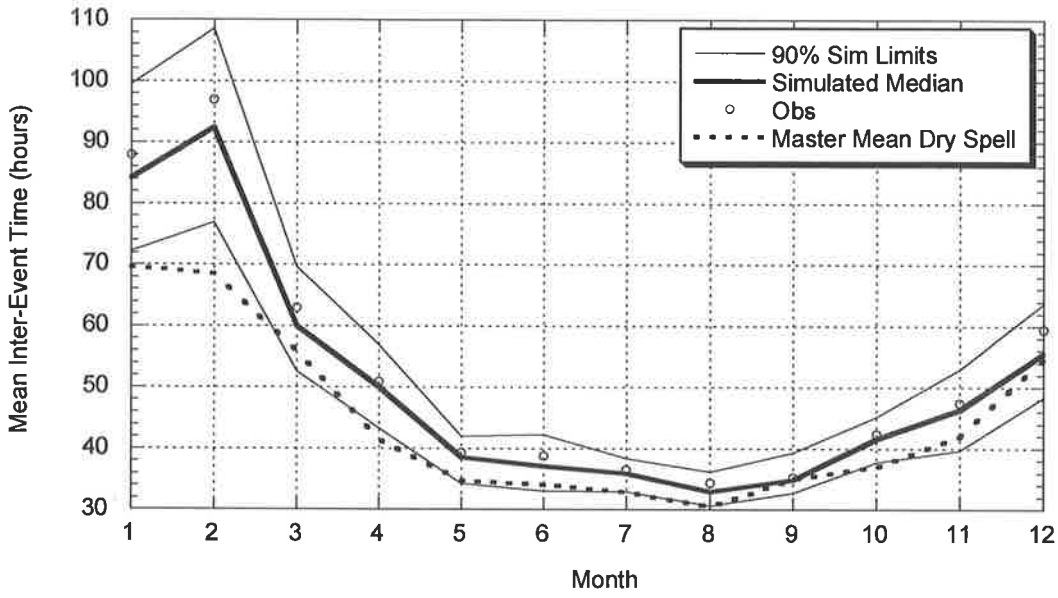

Figure C.3.45: Comparison between Observed and Target Simulated Mean of Inter-Event Times (Master – Melbourne; Target – Laverton RAAF)

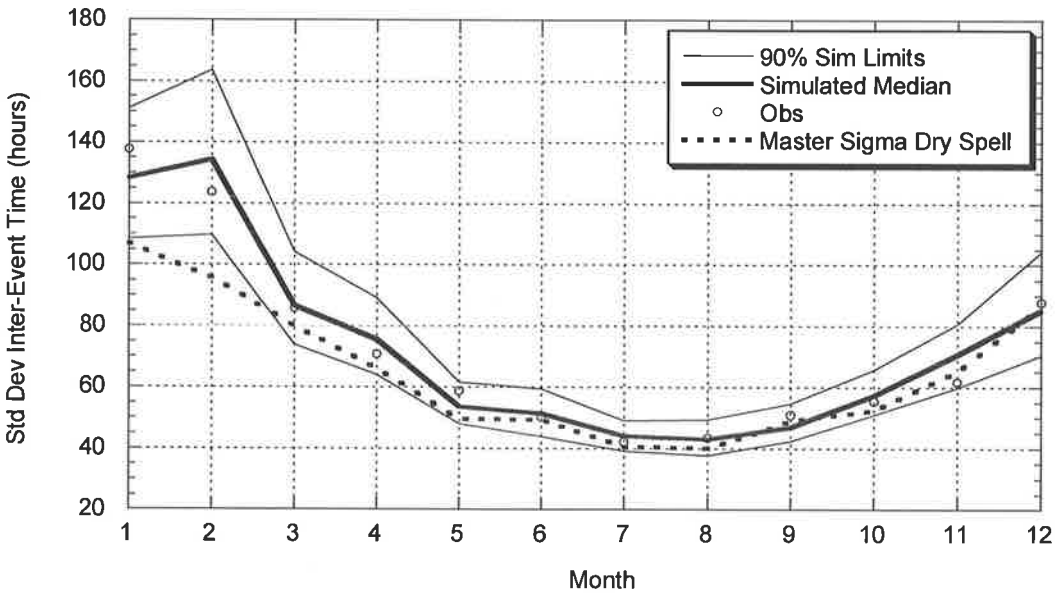

Figure C.3.46: Comparison between Observed and Target Simulated Standard Deviation of Inter-Event Times (Master – Melbourne; Target – Laverton RAAF)

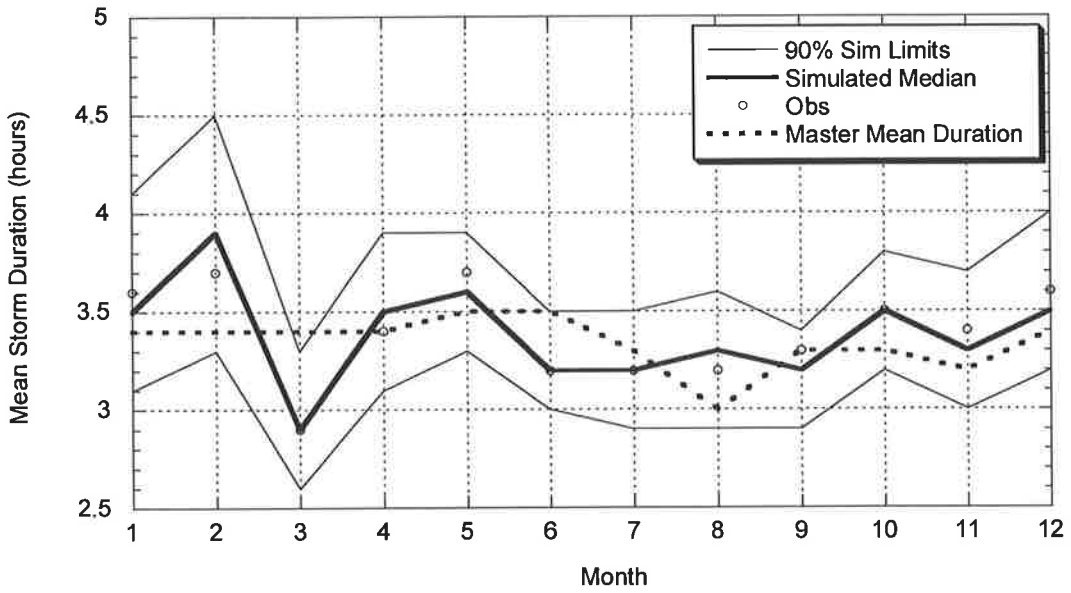

Figure C.3.47: Comparison between Observed and Target Simulated Mean of Event Storm Durations (Master – Melbourne; Target – Laverton RAAF)

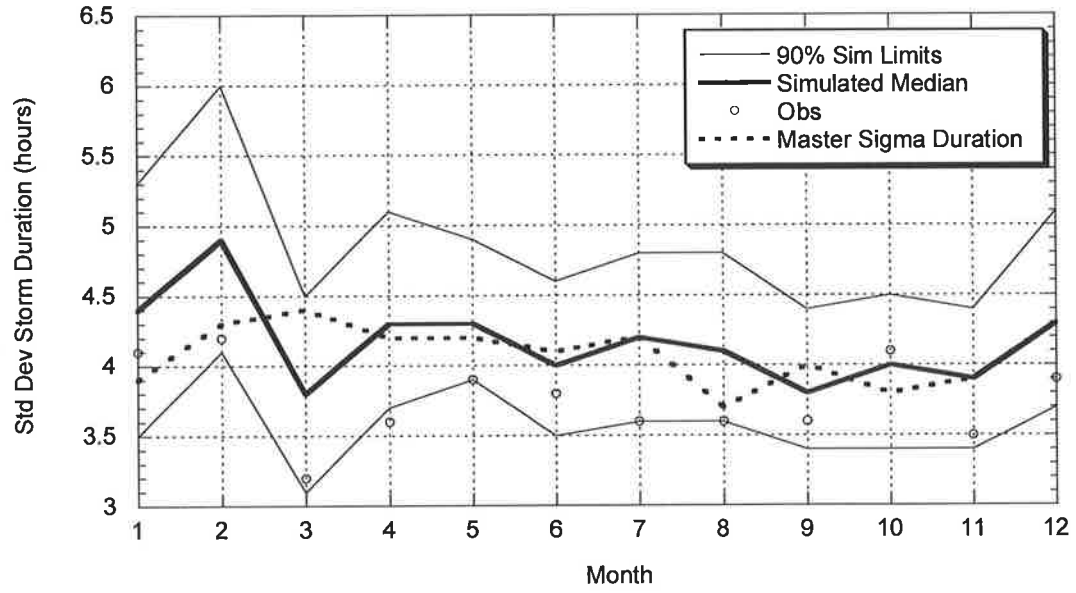

Figure C.3.48: Comparison between Observed and Target Simulated Standard Deviation of Event Storm Durations (Master – Melbourne; Target – Laverton RAAF)

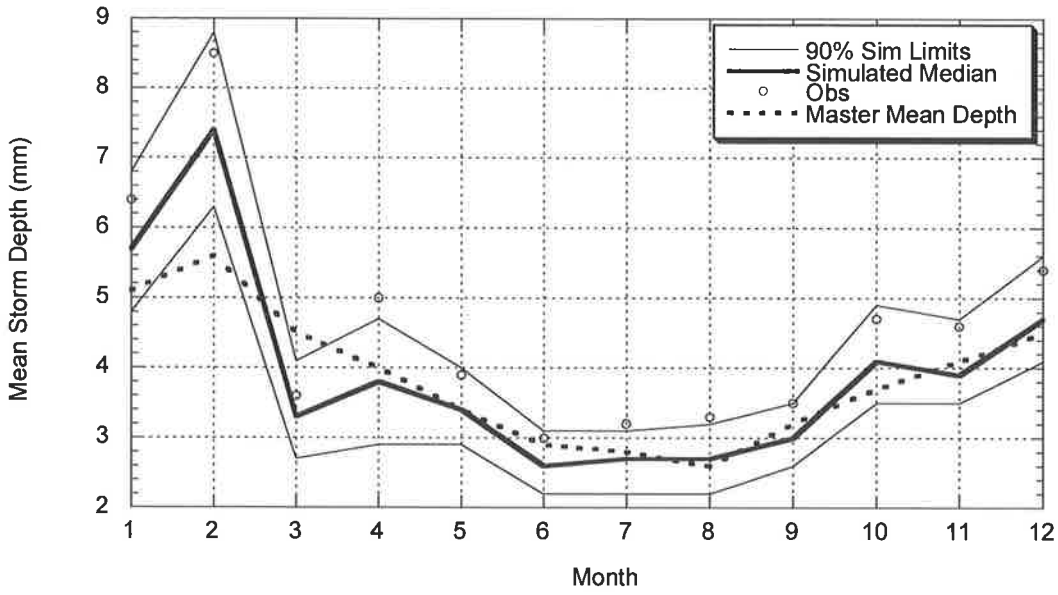

Figure C.3.49: Comparison between Observed and Target Simulated Average of Event Depths (Master – Melbourne; Target – Laverton RAAF)

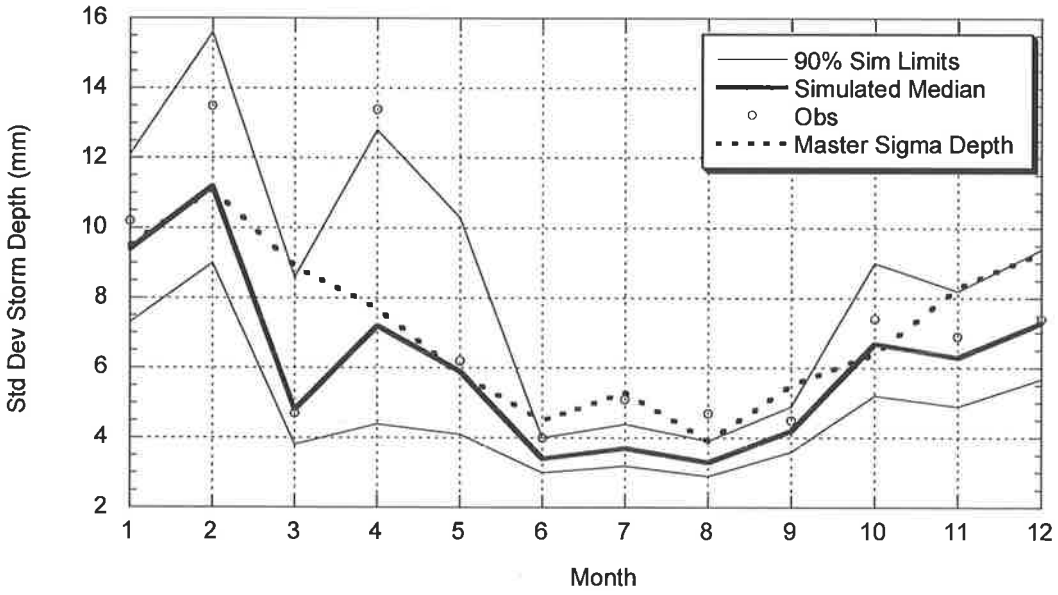

Figure C.3.50: Comparison between Observed and Target Simulated Standard Deviation of Event Depths (Master – Melbourne; Target – Laverton RAAF)

C.3.3.2 Simulated and Observed Daily Statistics

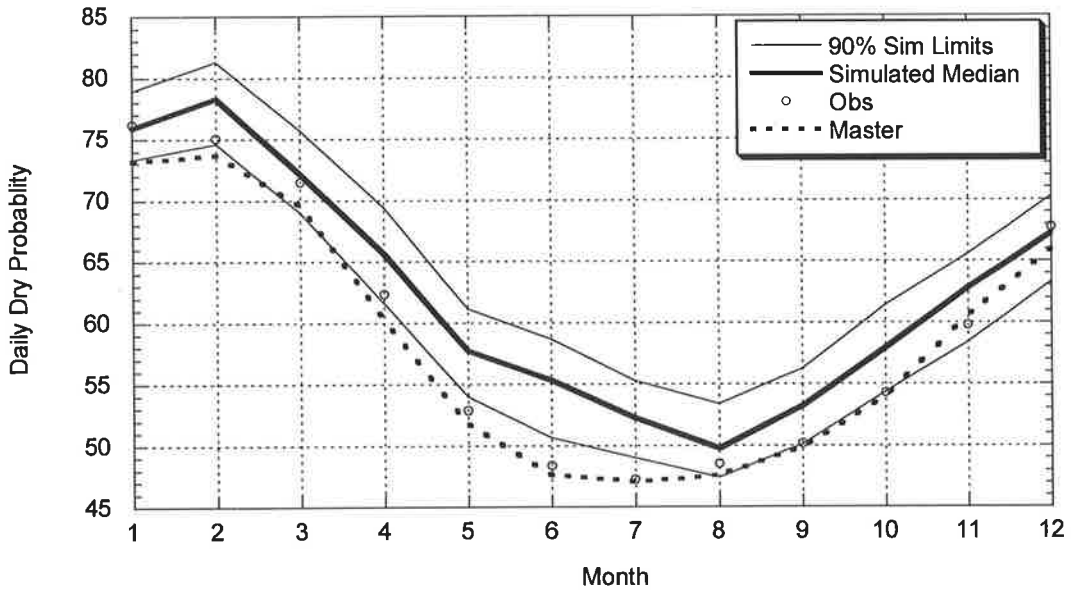

Figure C.3.51: Comparison between Observed and Target Simulated Daily Dry Probabilities (Master – Melbourne; Target – Laverton RAAF)

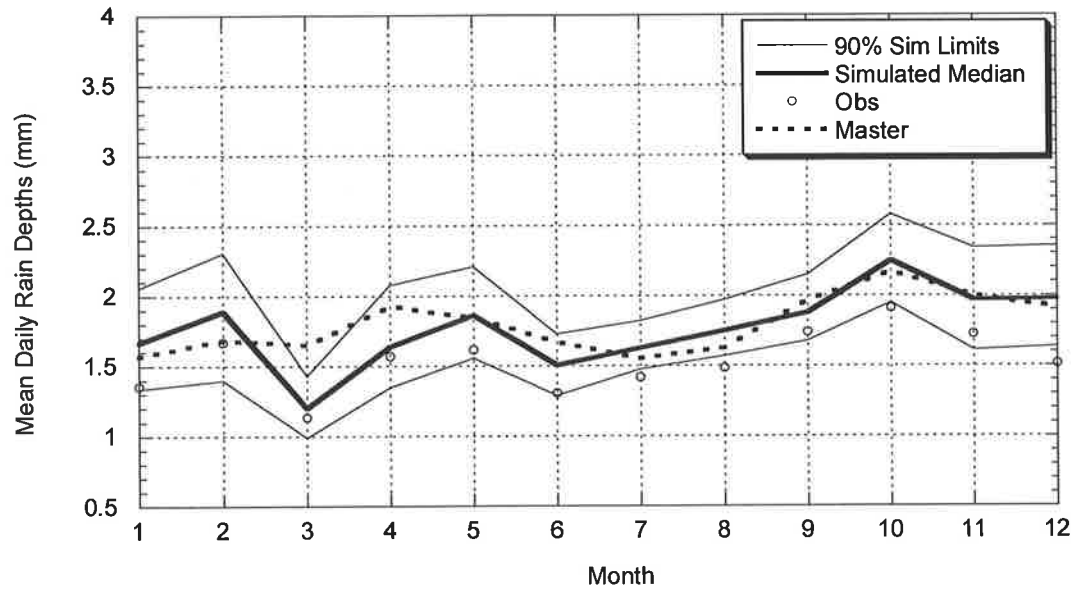

Figure C.3.52: Comparison between Observed and Target Simulated Daily Mean Depth (Master – Melbourne; Target – Laverton RAAF)

C.3.3.3 Simulated and Observed Annual and Monthly Rainfall

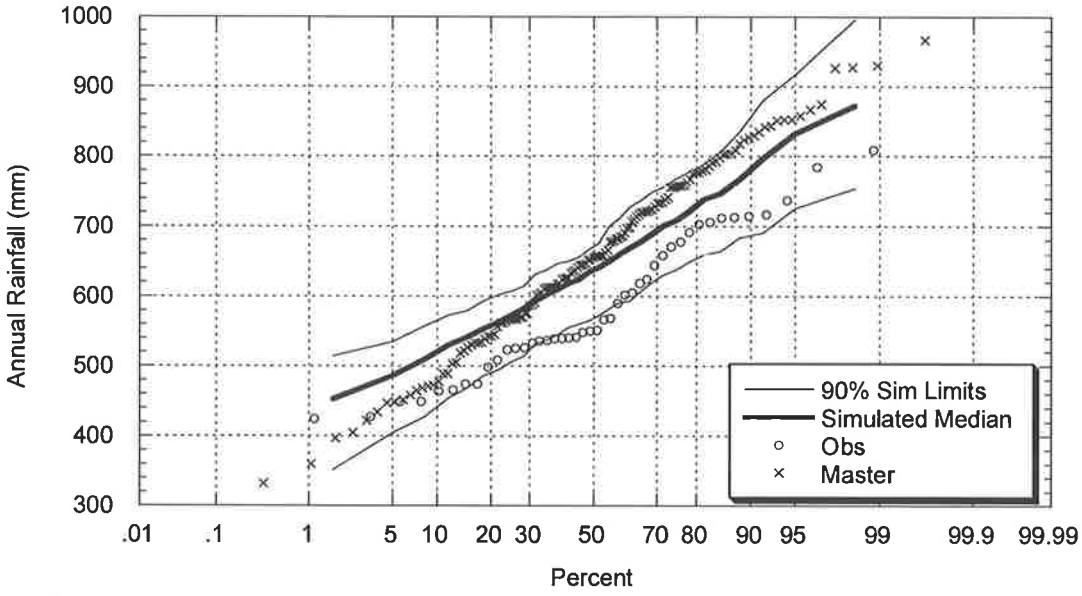

Figure C.3.53: Comparison between Observed and Target Simulated Annual Rainfall (Master – Melbourne; Target – Laverton RAAF)

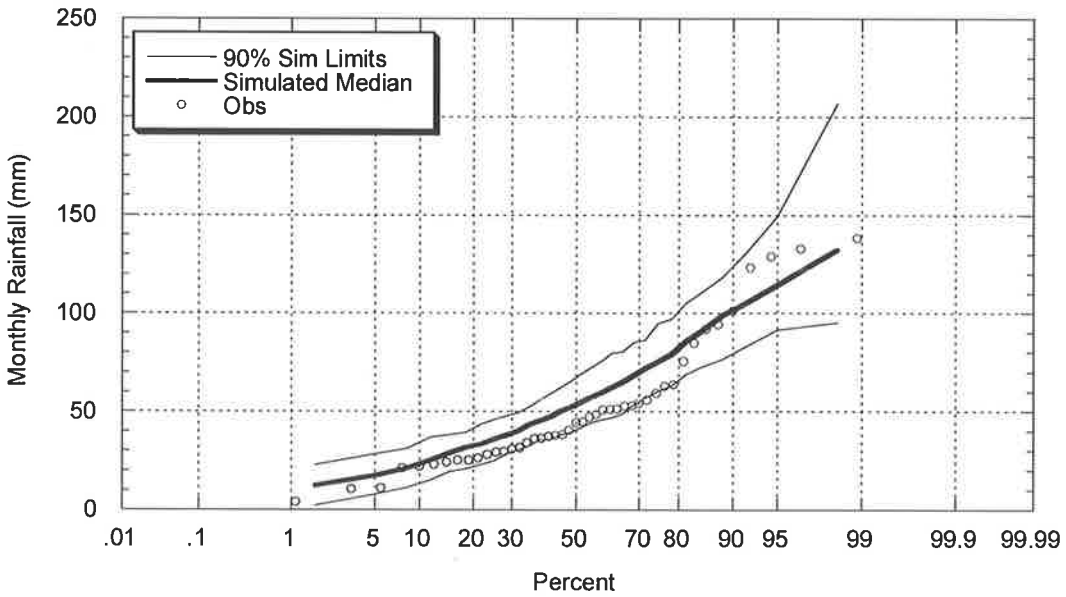

Figure C.3.54: Comparison between Observed and Target Simulated January Rainfall (Master – Melbourne; Target – Laverton RAAF)