



**Extreme Value Analysis of Non-Stationary
Processes – a Study of Extreme Rainfall Under
Changing Climate.**

By

Andrew J. Collier

*Submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy*

July 2010

This page is intentionally blank.

Abstract

The aim of this study has been to gain a greater understanding of the accuracy and levels of uncertainty associated with extreme rainfall event estimates, whilst considering both stationary and non-stationary processes (climate change).

This study started with the analysis and comparison of two extreme event fitting/estimation techniques: Linear Moments (L-Moments) and Maximum Likelihood Estimation (MLE) for the estimation of Generalized Extreme Value (GEV) distribution parameters. This thesis has found that MLE provides a number of advantages over L-Moments, especially when working with long or pooled data sets. These advantages include:

- The generation of confidence limits;
- Homogeneity testing; and,
- Trend detection / simulation.

However, the results of the analysis show that it is advisable to use L-Moments for single site analysis when the available data is less than 40 years in length. In this situation, L-Moments were found to produce less uncertainty.

Hosking and Wallis (1988) defined a method for the generation of synthetic data sets; this work has been reproduced and built upon as part of this thesis. Using this method it has been possible to gain insight on:

- Inter-site-dependence versus spatial separation (distance, km);
- The effects of inter-site-dependence on pooling groups;
- Regional correlation descriptors (level of dependence in a region);
- Synthetic data generation for regions with varying levels of dependence;
- Network Maximum (Netmax) Growth Curves; and,
- The effective number of sites in a defined region/pooling group.

This has been carried out using the 'R' statistical software/programming environment. Dales and Reed (1989), proposed the use of Netmax data (the largest value for one year across the network or pooling group) to increase the accuracy at the tail of an extreme event distribution by theoretically extending the curve. This hypothesis suggests that the separation between these two curves (the regional growth curve and the Netmax growth curve) is constant; allowing the Netmax curve to be translated and overlain on the regional growth curve. This study has found that the separation varies

with return period, implying that spatial correlation reduces (events become more independent) with increased rarity (or return period). However, these findings suggest complications with the use of Netmax data for the purpose of extending the regional growth curve.

In addition to the work detailed above, a method of trend detection in annual maximum rainfall has been demonstrated using synthetic data. Synthetic data has been used to enable control over the data, with this greater certainty and understanding in the results are achieved.

The same analysis was repeated on observed annual maxima for 1, 5 and 10 day durations, revealing evidence of trends, with stronger signals at higher durations. The trend was detected in the Location parameter, which relates to the mean. When using Synthetic data to understand the sensitivity of this test, it was found that the Location parameter required the weakest trend to be detected.

In summary this thesis has used synthetic data to gain a better understanding of:

1. Distribution fitting techniques;
2. Single site analysis;
3. Regional Analysis;
4. Spatial dependence; and,
5. Trend Detection.

All of the software that has been written as a result of this thesis to demonstrate the topics discussed, is included in Appendix 5, with explanations on the method of use. Should additional information be required, please contact Professor C. Kilsby at Newcastle University, who will forward on your enquiry.

Acknowledgements:

Supervisors:

C. KILSBY (School of Civil Engineering and Geosciences)

and

D. WALSHAW (School of Mathematics and Statistics)

Start Date: 1st October, 2003

End Date: July, 2010

Sponsors: EPSRC and Newcastle University

Without the ongoing support from Chris Kilsby and David Walshaw this thesis would not have been possible. I am very grateful to them both, especially during the first few years of my studies. I would also like to thank my wife and two children for their understanding and patience while I tried to write-up my findings and begin a new career during the years that followed.

Finally, I am extremely grateful to the Engineering and Physical Sciences Research Council (EPSRC) and Newcastle University for their financial support, without it I could not have contemplated undertaking this PhD thesis.

This page is intentionally blank.

Contents:

ABSTRACT	III
ACKNOWLEDGEMENTS:.....	V
1.0 - INTRODUCTION.....	3
1.1 - BACKGROUND.....	4
1.2 – IMPORTANT INFORMATION.....	6
1.2.1 - DATA QUALITY AND COVERAGE.....	6
1.3 – AIMS	8
1.4 - THESIS OVERVIEW	10
1.4.1 EXTREME RAINFALL VARIABILITY – A WIDER CONTEXT	11
2.0 - DAM SAFETY PRACTICE.....	17
2.1 – INTRODUCTION	17
2.2 - A SUMMARY OF THE 1975 RESERVOIRS ACT	17
2.3 - CAUSES OF FAILURES OR MAJOR INCIDENTS OF UK DAMS	18
2.4 - DESIGN FLOOD APPLIED TO SPILLWAYS PRIOR TO FLOOD STUDIES REPORT:	20
2.5 - DESIGN FLOOD FOR SPILLWAYS IN THE FUTURE	20
2.6 – MODELLING UN-GAUGED CATCHMENTS / CONTINUOUS SIMULATION	21
3.0 - CURRENT FLOOD ESTIMATION METHODS – UK AND EUROPE	25
3.1 - INTRODUCTION.....	25
3.2 - FLOOD ESTIMATION HAND BOOK (FEH)	25
3.2.1 - POOLING DATA	27
3.2.2 - GUMBEL REDUCED VARIATE – GRINGORTEN PLOTTING POSITION.....	27
3.2.3 - DEFINITION OF Y-SLICES.....	27
3.2.4 - NETWORK MAXIMUM POINTS	28
3.2.5 - FITTING THE GROWTH CURVE.....	28
3.2.6 - CONFIDENCE LIMITS FOR GROWTH CURVES	29
3.2.7 - TRENDS IN DATA	29
3.2.8 - REGIONAL FREQUENCY ANALYSIS.....	30
3.3 - DISTRIBUTION FITTING TECHNIQUES	30
3.3.1 - LINEAR MOMENTS (L-MOMENTS).....	30
3.3.2 - MAXIMUM LIKELIHOOD ESTIMATION (MLE)	32
A SIMPLE EXAMPLE OF MLE	33
3.4 – PROBABLE MAXIMUM PRECIPITATION	36
3.4.1 - THE THEORY OF PROBABLE MAXIMUM PRECIPITATION (PMP)	36
3.4.2 - THE PROBLEM WITH PMP ESTIMATES	37
3.4.3. STORM MODEL APPROACH.....	38
3.4.4 MAXIMISATION AND TRANSPOSITION OF ACTUAL STORMS	39
3.4.5 - GENERALIZED METHOD	43
3.5 - RESERVOIR FLOOD ESTIMATION, PROCEDURES AND DEVELOPMENTS IN OTHER EUROPEAN COUNTRIES; 10,000-YEAR RETURN PERIOD ESTIMATES	44
3.5.1 - INTRODUCTION	44

3.5.2 - FINLAND	44
3.5.3 – SWEDEN	45
3.5.4 – NORWAY	45
3.5.4 – UNITED STATES OF AMERICA	46
3.6 – SUMMARY.....	47

4.0 – THE STATIONARY MODEL.....51

4.1 - INTRODUCTION.....	51
4.2 - COMPARISON OF TWO DISTRIBUTION FITTING TECHNIQUES: L-MOMENTS AND MAXIMUM LIKELIHOOD ESTIMATES (MLE)	53
4.3 - METHOD FOR SINGLE SITE DISTRIBUTION FITTING TECHNIQUE COMPARISON:	54
4.4 – SUMMARY.....	58

**5.0 – SYNTHETIC DATA GENERATION FOR A REGION OF N-SITES WITH
SPATIAL INTER-SITE-DEPENDENCE.....61**

5.1 - INTRODUCTION.....	61
5.2 – SYNTHETIC DATA GENERATION.....	62
5.3 – METHOD.....	62
5.4 - A SUMMARY OF THE METHODOLOGY (WITH EXAMPLES)	65
5.5 - THE GENERATION OF A MATRIX OF SITE-TO-SITE SEPARATION, IN KILOMETRES:	66
5.5.1 – RESULTS AND DISCUSSION	68
5.6 - NETMAX CONCEPT AND LN(NE), THE EFFECTIVE NUMBER OF SITES	77
NETWORK MAXIMUM VALUES (NETMAX)	77
THE EFFECTIVE NUMBER OF SITES	77
5.7 – LN(NE) – EFFECTIVE NUMBER OF SITES FOR OBSERVED 1, 2, 5 AND 10 DAY ANNUAL MAXIMA DATA. REGIONAL CORRELATION / DEPENDENCE DATA FOR REGIONS WITHIN GREAT BRITAIN.....	81
5.7.1 – RESULTS FOR THE EFFECTIVE NUMBER OF SITES FOR REGIONS WITHIN GREAT BRITAIN	81
5.8 – HOMOGENEITY TESTING	85
5.8.1 – INTRODUCTION	85
5.8.2 – EXISTING TECHNIQUES.....	85
5.8.3 – POSSIBLE FUTURE TECHNIQUE.....	85
5.8.4 – METHOD DESCRIPTION.....	86
5.8.4.1 - EXAMPLE	87
5.8.5 – DISCUSSION.....	88
5.9 – SUMMARY.....	92

**CHAPTER 6 – USING A NON-STATIONARY MODEL TO REPRESENT AND TEST
FOR A TREND IN ANNUAL MAXIMA RAINFALL.....96**

6.1 – INTRODUCTION	96
6.2 – NON-STATIONARY GEV DISTRIBUTION	96
6.2.1 – GEV PARAMETERS	98
6.3 - GENERATION OF A TREND	100
6.4 – TREND DETECTION	101
6.4.1 - METHOD	101
6.4.2 – TREND DETECTION USING A TWO PARAMETER, GUMBEL (K=0) DISTRIBUTION	103
6.4.3 - SENSITIVITY TESTING OF PARAMETERS	105
6.4.4 – TREND DETECTION USING A 3 PARAMETER GEV (K≠0) DISTRIBUTION	110

6.4.5 - EQUIVALENCE OF CHANGES IN GEV PARAMETERS ON QUANTILE ESTIMATES	118
6.5 – TREND DETECTION IN OBSERVED ANNUAL MAXIMA DATA.....	121
6.6 – SUMMARY.....	126
<u>CHAPTER 7 – SUMMARY OF THESIS</u>	<u>130</u>
7.1 – SUMMARY OF CHAPTERS	130
7.2 – MAIN OUTCOMES	131
7.3 SUMMARY OF METHODS DEVELOPED TO ACHIEVE THE MAIN OUTCOMES	133
7.4 – FUTURE WORK	134
7.4.1 – INTRODUCTION	134
7.4.2 – AREAS TO BE CONSIDERED FOR FUTURE WORK	134
7.4.3 – FUTURE WORK EXPLAINED	134
<u>REFERENCES</u>	<u>138</u>
<u>TABLE OF APPENDICES:</u>	<u>146</u>
APPENDIX 1 – REGIONAL CORRELATION AND EFFECTIVE NUMBER OF SITES, RESULTS	148
APPENDIX 2 - REGIONAL HOMOGENEITY RESULTS	183
APPENDIX 3 – SAMPLE OF ‘R’ ROUTINE FOR HOMOGENEITY TESTING	195
APPENDIX 4 - STATISTICAL TERMINOLOGY AND CONCEPTS.....	199
1. BASICS OF THE TECHNIQUE.....	199
2. APPLICATIONS	199
3. USES.....	199
4. LAWS OF PROBABILITY	200
5. TERMINOLOGY AND CONCEPTS	200
6. STATISTICAL PARAMETERS.....	201
APPENDIX 5 – COMPLETE ROUTINES FOR USE WITHIN ‘R’	205
A5.1 - INTRODUCTION	205
A5.2 – EXPLANATION OF ROUTINE 1.....	205
A5.3 - ROUTINE 1.....	207

This page is intentionally blank.

Chapter 1

This page is intentionally blank.

1.0 - Introduction

According to Thompson and Perry (1998), floods caused the highest percentage of deaths (26%) among all natural hazards evaluated around the world during the period 1963 – 1992. With a total of 32% they also have the leading position in total significant damage and the second position (32%, after droughts) in the total number of people affected.

Climate model integrations predict increases in both frequency and intensity of heavy rainfall in the high latitudes of the Northern Hemisphere under enhanced greenhouse conditions (Jones and Reid, 2001; Palmer and Räisänen, 2002). These projections are consistent with recent increases in rainfall intensity seen in the UK (Osborn et al., 2000; Fowler and Kilsby, 2003a,b), Europe (Brunetti et al., 2000; Frei and Schär, 2001) and worldwide (e.g. Karl and Knight, 1998; Iwashima and Yamamoto, 1993; Zhai et al., 1999), although it is not possible to relate one to the other as cause and effect (Fowler et al., 2004).

It is important to determine whether these increases are due to natural variation, or whether they are part of a trend or change in the climate. The impact of such changes is more wide reaching than the obvious increased risk of flooding to floodplain areas. This impact extends for example to structural calculations for bridges, dams and flood defences, to list just a few. The proper estimation of design values requires that these data series from which the probability distribution parameters are to be estimated, come from independent and identically distributed (i.i.d) observations. The proper assessment of risk factors for a designed structure requires that the statistical inference has also to be valid during the projected life span of the structure. This requires the conditions (e.g. climate) under which the inferences are made, to remain constant in the future.

Recent extreme rainfall events in the UK have characteristically been extended over several days, with unremarkable one-day totals (Fowler et al., 2004). This study has revealed evidence of trends, with stronger signals at higher durations, meaning multi-day events.

1.1 - Background

For some time now, the Association of British Insurers has been considering the withdrawal of flood insurance from 10% of UK properties, worth some £200 billion, considered to have inadequate flood defences after 31st December 2002, and there has been a similar insurance response to flood hazard globally (Crichton, 2002). In addition to the existing, ‘known’ or observed problem, hydrologists and engineers have the unenviable task of estimating the depths of extreme, rare rainfall and floods, whilst making provision for ‘climate change’. Calculating the rainfall depths, flood volumes and river flow rates, which lead to the associated defences / structures which must not fail during these extreme conditions, has been the motivation and inspiration for many analysis techniques and methods of extrapolation; for example Probable Maximum Precipitation (PMP) and the Probable Maximum Flood (PMF).

Many structures are designed to withstand or pass a rare event, for example: drainage systems, bridges, dams and flood defences. The ‘rare’ design event may have a return period of 40 to 100-years for standard assets, or in excess of 1,000-years for key structures. Often the design criteria are based on the acceptability of failure or the cost/benefit ratio. For example, urban flood defences may be designed for the 50 year return period event, while a large class ‘A’ reservoir will be designed for the 10,000-year event. To put that into context, a nuclear reactor is designed to withstand a seismic event with a return period of 10,000-years.

The rarity of the event may be assessed using a long record. This does not imply that it will be 50 years until the next 50 year event, but over a sufficiently long period of time this event will approximate to a 50 year recurrence interval. This statement assumes that there is stationary underlying distribution to the data, and that any changes observed in the data, are natural fluctuations within an unchanging envelope of variability (Milly et al 2008). “This is a foundational concept of hydrological analysis and engineering. It implies that any variable (e.g., annual streamflow or annual flood peak) has a time-invariant (or 1-year-periodic) probability density function (pdf), whose properties can be estimated from the instrument record. Under stationarity, pdf estimation errors are acknowledged, but have been assumed to be reducible by additional observations, more efficient estimators, or increased regional

data. The pdfs, in turn, are used to evaluate and manage risks to water supplies, waterworks, and floodplains” (Milly et al 2008). “In view of the magnitude and ubiquity of the hydroclimatic change apparently now under way, however, we assert that stationarity is dead and should no longer serve as a central, default assumption in water-resource risk assessment and planning. Finding a suitable successor is crucial for human adaptation to changing climate” (Milly et al 2008).

This thesis will focus on the very extreme events required for the design of dam spillways, for example. Currently, large (Class A) reservoirs are designed to pass a Probable Maximum Flood (PMF) or 10,000-year flood. The question being asked by government organisations and insurance companies is: what will the precipitation depth of these events be like in the future, for example in 50 or 100-years time? If the associated peak flow rate increases (in association with climatic changes), then the spillway must be modified to avoid damage to the reservoir and possible failure, but by how much must the spillway capacity be increased and by when?

One way of making ‘predictions’ is to use statistical analysis. Using available rainfall data and extrapolating from this data, the probability of an event taking place can be calculated – for 179 rain gauges throughout Great Britain with data from 1960 to 2000. A map of the rain gauges (figure 1.2a) is included in section 1.2.

This analysis will be expanded upon in the appropriate chapters, giving an explanation of each technique that has been reviewed.

Using statistical analysis, it is possible to provide varying return period estimates, for example 100, 1,000 or 10,000-year return period events. However, there are a number of problems associated with generating extreme rainfall estimates:

1. Extreme events are by definition rare and available rainfall records are usually short;
2. Incomplete data sets;
3. Errors in the recorded data set and error filtration / correction that can remove real extreme events;
4. Spatial dependence, when pooling sites in a region to effectively increase the length of the data set; and,
5. The existence of trends in time series of rainfall.

Of the 5 problems listed above, it is not possible to identify one of these as being the most significant, as clearly each has a significant impact on the accuracy of generating

extreme rainfall estimates. The estimation of extreme rainfall in relation to reservoir risk assessment is conventionally carried out using the following techniques:

- Probable Maximum Precipitation (PMP), see chapter 3 for an explanation of how this can be calculated.
- Statistical analysis of available rainfall data to generate a non-exceedence probability of $p=0.9999$ or a 1 in 10,000-year estimate; put a slightly different way the probability of an event equalling or exceeding the 10,000-year event is $p \leq 0.0001$.

As stated in the abstract, this study has used synthetic data to gain a better understanding of:

1. Distribution fitting techniques;
2. Single site analysis;
3. Regional Frequency Analysis (RFA);
4. Spatial dependence; and,
5. Trend detection.

Synthetic data has the advantage of being from a known distribution, and the presence of a trend, for example, can be controlled and its impact observed. Having control over the data enables conclusions to be drawn from the findings and this then greatly aids the interpretation is actual rainfall data when analysed.

1.2 – Important information

The reader should note the following: unless stated otherwise, much of the work within this document is carried out using ‘synthetic data’ generated from a known distribution, and not observed or recorded data. The term ‘synthetic data’ is used here to represent data generated using Monte Carlo simulation (random number generator). This approach has been chosen, so as to have prior knowledge of the desired or true result. This allows analysis and conclusions to be drawn at each stage during this piece of research.

1.2.1 - Data quality and coverage

Where observed data has been used in this thesis, this is the same data set used by Fowler and Kilsby (2003a). This in turn is based on an original data set from 1961 to 1995, which was subsequently updated by Fowler and Kilsby (2003a) to extend the

records from 1961 to 2000. The initial and subsequent data was extracted from the archives of the British Atmospheric Data Centre (BADC, www.badc.rl.ac.uk) (Fowler and Kilsby, 2003a, Osborn and Hulme, 2002), with selection criteria including the requirement for a reasonable spatial and temporal coverage of Great Britain, as well as record length and completeness (Osborn and Hulme). All of the 110 stations selected for the initial period of 1961 to 1995 had complete or nearly complete data for this period. Sites were then added to this initial data set to ensure that each of the eight regions contained twenty or more stations. Wigley et al. (1984) defined nine regions, but only eight have been used in this thesis. However, all nine are illustrated in Figure 1.2b on the following page.

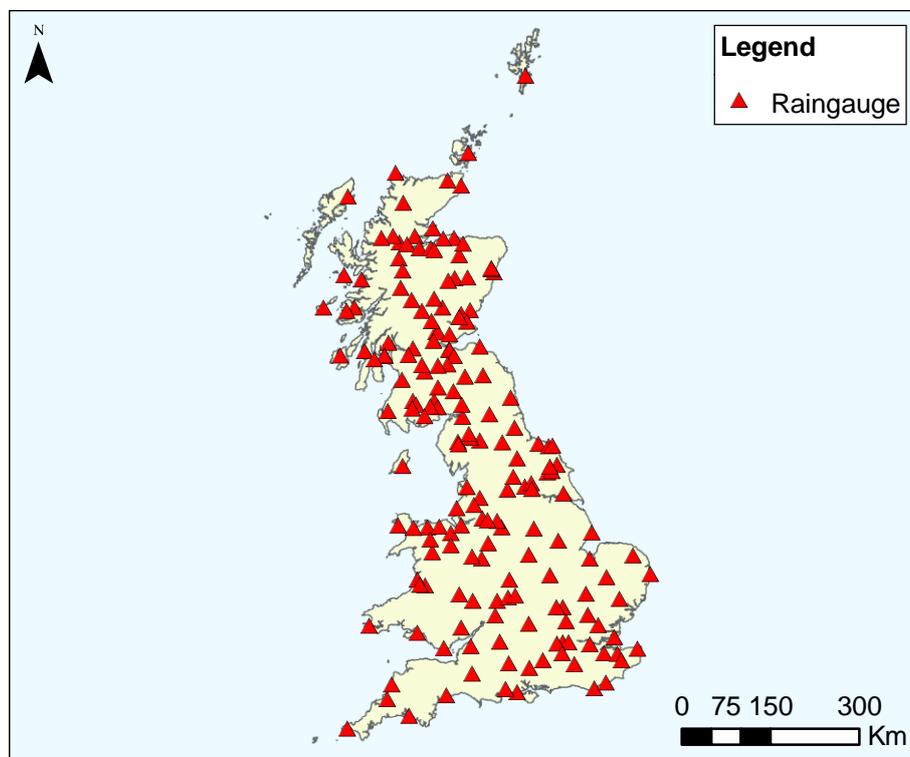


Figure 1.2a: Map showing the location of the 179 rain gauges used in this thesis.

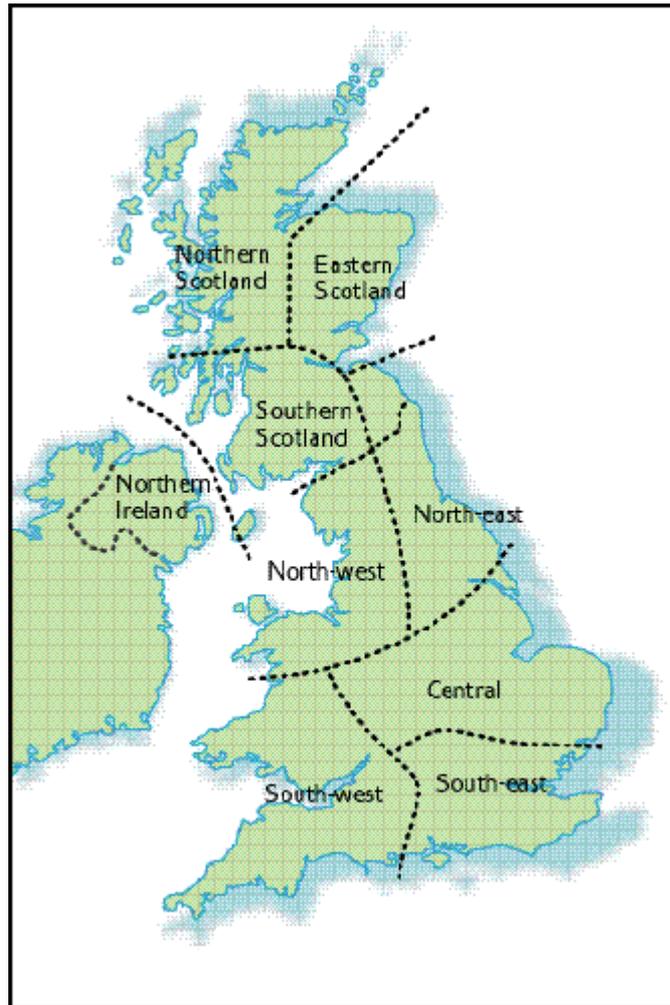


Figure 1.2b: Map showing the regions as defined by Wigley et al (1984). The northern Irish region (and rain gauges contained within) has not been included in this thesis.

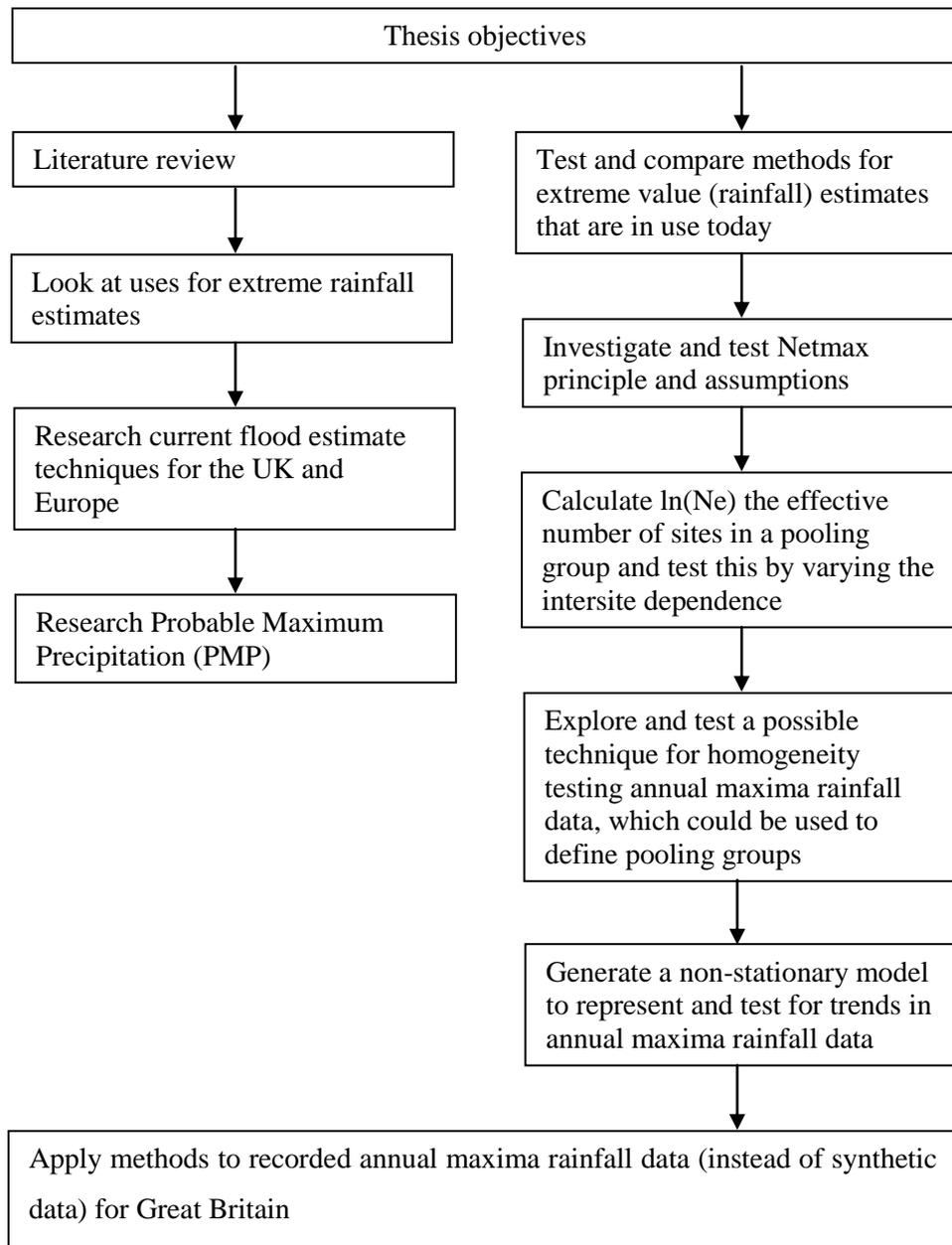
Wigley (1984) reports that his analysis of rainfall data (1861–1970) points to strong geographical and topographical control, that results in England and Wales being divided into five coherent sub-regions and a further three in Scotland. These were defined using a regression technique, that was developed to produce homogeneous area-average precipitation series for England and Wales using the longest site precipitation records available and maintaining even spatial coverage.

1.3 – Aims

Given the need for extreme rainfall estimates, i.e. greater than the 1 in 1,000-year return period event, this thesis aims to:

- Look at uses for extreme rainfall estimates, i.e. dam safety practice (reservoir design);
- Better understand the methods for extreme value (rainfall) estimates that are in use today;
 - Understand the limitations of these methods.
- Look at and understand alternative (not statistically based) methods of producing extreme value rainfall estimates – Probable Maximum Precipitation (PMP);
- Look at current flood estimation techniques for the UK and Europe;
- Carry out a comparison of two Generalised Extreme Value (GEV) distribution fitting techniques, for the stationary model (which assumes there is no [climatic] trend in the data);
- Use a method for synthetic data generation, to look into and gain greater understanding on:
 - The Netmax concept and $\ln(N_e)$, where $\ln(N_e)$ is related to the effective number of sites in a pooling group based on varying (artificially introduced) intersite dependence;
 - A possible technique for homogeneity testing, which could be used to define pooling groups;
- Generate a non-stationary model to represent and test for a trend in Annual Maxima rainfall – using synthetic data.
- Apply this method to recorded annual maxima rainfall data for Great Britain.

This workflow is summarised in the flow chart below.



1.4 - Thesis Overview

Chapter 1, this chapter, introduces the problem and sets out the aims of this thesis. Chapters 2 and 3 look at the theory behind dam safety practice and introduce the flood estimation methods currently used within the UK and Europe. Chapter 3 also looks at Probable Maximum Precipitation (PMP), which is one method of estimating the most extreme precipitation likely to fall at the site of interest. Chapter 4 considers two distribution fitting techniques and demonstrates the strengths and weaknesses of each.

For single site analysis, **chapter 4.3** demonstrates that L-Moments (one of the two distribution fitting techniques) are the preferred choice for short time-series (less than 40 years), typically meaning single site analysis. Maximum Likelihood Estimates (the other distribution fitting technique) was shown to outperform L-Moments for longer time-series, or when using Regional Frequency Analysis (RFA), with a combined record length greater than 40 years.

Regional Frequency Analysis (RFA) pools data from multiple sites. This method is frequently used to increase the length of the series. The main problem with this approach is spatial dependence (correlation) between sites. This reduces the effective length of the pooled time-series; but by how much? Chapter 5 answers these question using:

Chapter 5.2 - A multivariate normal model has been used to model the dependence structure between sites and to observe the impact on confidence intervals for fixed quantile estimates;

Chapter 5.6 - The same model was used to test the Network Maximum (Netmax) theory. This theory assumes that the separation between the Netmax growth curve and the regional growth curve is proportional to the number of sites in the region; it also assumes that the separation is constant; and,

Chapter 6.4 - Trend Detection: Using the assumption that trends can be detected in Annual Maxima time series of rainfall and hence the Generalized Extreme Value (GEV) distribution. Tests were carried out by allowing the fitting of a time dependent covariate to each of the GEV parameters.

1.4.1 Extreme rainfall variability – a wider context

The variability of extreme rainfall including trends over the observed period and possible future changes is important for many applications ranging from water resources to flood risk management. This thesis focuses on engineering design requirements and methods adopted by engineers and hydrologists but valuable information can be obtained from areas of study beyond this highly specialised area. This literature review has therefore covered the following areas of research that are considered to be particularly relevant:

- Stationarity and trends in rainfall intensity

- Spatial dependence

Stationarity and Trends in Rainfall Intensity

Current techniques for the fitting of distributions to extreme event data typically rely upon an assumption of stationarity, meaning that there is no underlying trend in the data and that variations in the data are from natural fluctuations in the climate. This is a foundational concept of hydrological analysis and engineering. This assumption is increasingly questioned as there is a growing consensus amongst the scientific community that variations in observed data are not entirely due to natural variations but indicate the presence of an underlying trend: “In view of the magnitude and ubiquity of the hydroclimatic change apparently now under way, however, we assert that stationarity is dead and should no longer serve as a central, default assumption in water-resource risk assessment and planning” (Milly *et al.* 2008).

Trends (usually increases) in rainfall intensity have been investigated in many parts of the world. They have been detected in UK data (Osborn *et al.*, 2000; Fowler and Kilsby, 2003a,b, Ekström *et al.*, 2005, Maraun *et al.*, 2008), Europe (Brunetti *et al.*, 2000; Frei and Schär, 2001, Fowler *et al.*, 2007) and worldwide (e.g. Karl and Knight, 1998; Iwashima and Yamamoto, 1993; Zhai *et al.*, 1999).

Many authors have recognised the importance of extreme rainfall and the likelihood that climatic changes are occurring and that we need to be able to identify by how much they have already changed and by how much they may change in the future. For this reason, climatologists have explored various ways of:

- Extracting,
- Examining and,
- Interpreting the data.

Maraun *et al.* (2008) take the non-parametric method described by Osborn *et al.* (2000) where each daily rainfall is assigned to one of ten categories based on its amount. Each category makes up 10% of the total rainfall amount for this month. The analysis (via trend fitting, principal component analysis and area averaging) is then carried out using these category time series. Particular attention is given to the

category containing the highest daily totals. This method disregards absolute precipitation values or an explicit analysis of the annual cycle, instead focusing on the precipitation intensity distribution (Maraun et al., 2008). Using this approach, seasonal trends have been detected in rainfall data across the UK.

Engineers most often extract annual maximum values from rain gauge records, and fit distributions to estimate rainfall events of the desired return period. This approach is essentially the basis for this thesis, with the addition of spatial pooling to extend the observed data set.

Where trends have been found in rainfall, it has been predicted that this will typically lead to an increase in rainfall intensity. Groisman et al (1999) report that changes in mean precipitation totals tend to have the most influence on the heavy precipitation rates. This scenario gives changes in heavy rainfall which are comparable to those observed and are consistent with the greenhouse-gas-induced increases in heavy precipitation simulated by some climate models for the next century (Groisman et al, 1999, Haylock et al, 2006, Hegerl et al, 2004).

Spatial Dependence

This thesis has investigated the impact of spatial dependence upon the quality of the data in a pooling group. A pooling group is a selection of rain gauge sites that have been pooled to produce a larger dataset. If the data is correlated to some extent, then the effective size of the pooled series is reduced. A number of other studies have investigated spatial dependence (correlation between sites) at the daily time-step and have also shown the expected decline from high correlation of nearest neighbours to the low correlations of distant sites (Wilby *et al.*, 2003, Osborn *et al.* 1997, Buishand and Brandsma, 2001). They also show that the structure of the observed decay differs between regions, as found by this thesis.

The objective of investigating spatial dependence was to understand the impact it had upon the effective size of the data set in a pooling group. Large pooling groups are desirable as they can reduce the uncertainty in extreme event estimates by providing a larger sample of independent data. A concept which will be explored later is that with very large pools of data, there could be a limiting distribution for the most extreme

events. Wilson and Toumi (2005) investigate the possibility of a fundamental probability distribution for heavy rainfall and find that little is known about the physical limits of heavy rainfall. From a physics standpoint, they propose a mean value for the shape parameter for an extreme value distribution for UK rainfall. If this value were to be used instead of empirically estimated values, radical changes in return period event estimates would be found.

Chapter 2

This page is intentionally blank.

2.0 - Dam Safety Practice

2.1 – Introduction

Although this study does not look exclusively at reservoir safety, reservoirs are one of the few structures that require such extreme estimates. As such, additional research has been carried out to appreciate the association between extreme events and dam failure.

At least 60 catastrophic dam failures (Wright, 1994) have occurred in the UK (Binnie and Partners, 1986) of which four are known to be due to over-topping. Between one quarter and one third of all dam failures are due to overtopping, which is in turn due to inadequate spillway design (Gruner, 1963).

2.2 - A Summary of the 1975 Reservoirs Act

The Reservoirs Act of 1975, which replaced earlier similar legislation (The Reservoirs Safety Provisions Act 1930), was set up to promote the safety of large raised reservoirs. These are defined as retaining more than 25,000m³. This is approximately a football pitch 4.27m deep. A small reservoir therefore, retains a volume of less than 25,000m³.

The volume is measured above the lowest point of naturally occurring ground level - i.e. the level to which the reservoir could drain if it were to fail. A natural depression does not count, unless the water could drain out by gravity.

The regulations require that any reservoir within the scope of the Act may only be designed, or its construction supervised, by an engineer on the appropriate panel. Following construction, another panel engineer must inspect the reservoir within three years. During the life of the structure, a member of the Supervising Engineers panel must be retained to carry out regular inspections, typically every year. An engineer from the appropriate panel must inspect at periods to be advised, but not less than every ten years or when requested by the supervising engineer. The inspecting engineer may instruct that work be carried out for the safety of the reservoir, and this instruction has the force of law.

The initiating mechanism for a specific incident or failure is frequently masked by a progression from one mechanism to others, for example: internal erosion can lead to local depression, which in turn could lead to overtopping by an extreme flood event [CIRIA C542, p 19].

This study is most interested in overtopping, which accounts for approximately 30% ($\pm 5\%$) of major incidents / failures. When looking at these figures, it is important to note that some 85% of UK dams are earth-fill embankments [CIRIA C542, p 24].

Overtopping can be initiated by many mechanisms such as [CIRIA C542, p 22]:

- Inadequate spill weir
- Blockage of weir/spillway
- Lack of freeboard
- Local settlement
- Excessive pumping into reservoir
- Inappropriate design flood inflow

The last mechanism listed is inappropriate design flood inflow. This is of particular interest as the majority of UK dams have been designed to withstand either the Probable Maximum Flood (PMF) or the 10,000-year return period flood event. If we take into consideration that the median age of dams in the UK is of the order of 105-110 years [CIRIA C542, p 24] it is understandable that questions are now being asked - questions such as how climate change might impact on these estimates and what this will mean for dam safety in the future.

The overflow capacity of many reservoirs has been increased in recent years following the increases in recommended design floods. Works to increase overflow capacity include the following:

- Modification of the existing weir/channel/stilling basin;
- Construction of new overflow works to replace or supplement the existing ones;
- Construction of an auxiliary overflow with its crest at a level above top water level which operates infrequently; grass or reinforced grass is commonly used in such instances;

- Improvements to the crest and/or downstream face of a dam so that rare overtopping is tolerable in accordance with Floods and Reservoir Safety (ICE, 1996);
- Raising the crest level of the dam;
- Constructing or raising a wave wall; and,
- Lowering the top water level.

In many instances a combination of some of the above works are used to provide the necessary overflow capacity [CIRIA C542, p 88].

Many recent small dams have been designed to overtop on rare occasions with flow passing over an engineered auxiliary overflow. Older dams often show signs of frequent uncontrolled overtopping yet have functioned satisfactorily for many decades. The behaviour under overtopping conditions depends on many factors and influences and their complex interaction, but it is evident that the vast majority have been able to withstand such flows without significant damage. Furthermore, it appears that the instances of overtopping may, in some instances be more frequent than envisaged with some dams overtopping as often as annually [Reservoir Safety and the Environment, p 260].

2.4 - Design Flood Applied to Spillways Prior to Flood Studies Report:

It would appear that with the state of knowledge in 1930 and for years afterwards, spillways were designed to pass a flood which the engineer, based on his own experience, considered to be a maximum for the catchment. No doubt the engineer had his own empirical formula or method for assessing design floods. On the whole, as these were used by engineers with considerable experience, they seem to have worked very well, possibly aided by additional freeboard and the fact that reservoirs are usually drawn down in the summer when thunderstorms are most likely [K. T. Bass, 1975].

2.5 - Design Flood for Spillways in the Future

Following the publication of the Flood Studies Report (FSR) and then the Flood Estimation Handbook (FEH), it is possible to make an estimate of floods having

return periods of up to 1 in 2,000 years. Furthermore, estimates of the Probable Maximum Flood (PMF) based on Probable Maximum Precipitation (PMP) may be obtained. Typically this is carried out using a rainfall-runoff method and calibrated using observed storm events.

2.6 – Modelling un-gauged catchments / Continuous simulation

River flood frequencies at un-gauged sites across Britain can be estimated using continuous simulation. Rainfall - Runoff modelling is undertaken for a set of catchments for which flow data are available to allow calibration of model parameters. These parameter values are then related to more widely obtainable 'catchment property' data which are available across the zone of concern for which the final methodology is required. These catchment properties are used to define model parameters for un-gauged sites and the runoff model(s) then run for these sites to derive a flow time series from which flood characteristics and statistics can be drawn. If, in addition, the runoff models can be driven by long rainfall time series (observed or generated), it is possible to extend estimation of floods to higher recurrence intervals than those warranted by calibration period data (Calver et al). The generated rainfall data can be derived using stochastic processes and therefore can simulate very large events, of which allowance can be made for future scenarios, such as increasing intensity and the inclusion of an underlying trend in the generation of the data. A schematic is provided on the following page that outlines this process (figure 2.6).

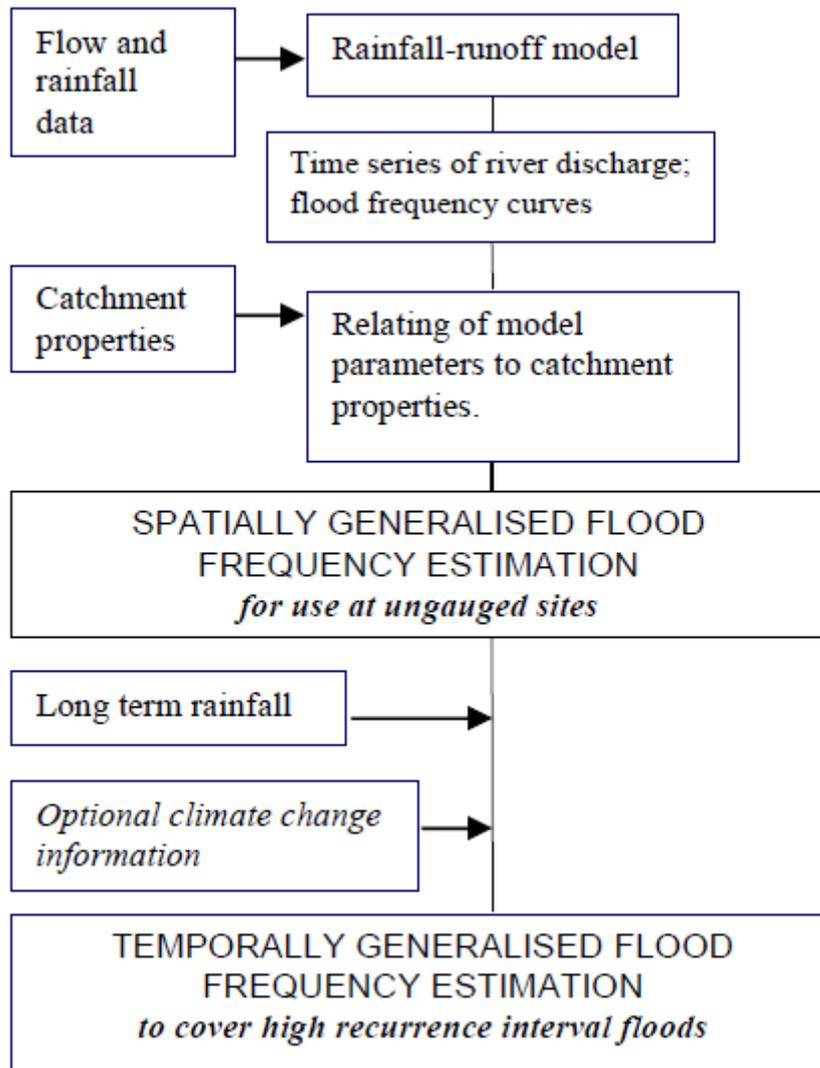


Figure 2.6: Schematic representation of continuous simulation

Chapter 3

This page is intentionally blank.

3.0 - Current Flood Estimation Methods – UK and Europe

As discussed in the Institute of Hydrology's 1999, Flood Estimation Handbook (Volumes: 1 and 2), there are a number of flood estimation methods in use within the UK and Europe. These methods are discussed within this chapter.

3.1 - Introduction

There are a number of limitations associated with rainfall run-off models. However, their advantage in comparison with river flow records is that rainfall records are usually longer and more numerous, further, there tend to be more rain gauges than river stage recorders. The rainfall depths associated with a certain return period event need to be successfully estimated to aid many bridge, culvert, drainage and reservoir spillway calculations, although the methods used to achieve these estimations varies, having evolved over time. This chapter aims to introduce the reader to some of the techniques widely in use today.

3.2 - Flood Estimation Hand book (FEH)

The Flood Studies Report (FSR) has provided the “state of the art” method for flood and rainfall estimation in the UK for almost 25 years before 1999. However, during this time, the guidance document contained within the report was reviewed and revised. Difficulties then arose because not all users were aware of these revisions and the guidance lost its value.

In 1999, the FEH handbook was introduced. This offered a more robust method, which contained datasets in a digital format. The documents and datasets contained within the handbook provide users with a cohesive set of procedures for flood frequency and rainfall estimations and makes the user aware of the uncertainty associated with estimation.

The basis of the FEH rainfall frequency analysis is formed by Annual Maximum (AM) rainfall data. AM are the largest rainfall observations at each site, for that year of record. The key components of the analysis are:

- the index variable, R_{med} (the median of the AM rainfall for a single site) and

- the growth curves (the distributions associated with AM rainfall that have been standardised by R_{med}).

The FORGEX (FOcused Regional Growth curve Extension) method was used to generate rainfall growth curves. These growth curves as well as the R_{med} values for durations ranging from 1 hour to 8 days were mapped across a 1km grid. Combining the R_{med} values with the associated Growth curves allowed a Depth Duration Frequency (DDF) model to be produced. An explanation of how to use this method is contained within Chapter 2 of the FEH. All results in the FEH are given as fixed duration events and should be converted to sliding duration events using table 2.1, page 8, Volume 2 of the FEH.

Catchment wide analysis is carried out by calculating a weighted average of point DDF values within the defined catchment. The catchment average rainfall depth is then calculated by applying an areal reduction factor. This is based on the assumption that, especially for extreme storms, the rainfall is not uniform across the whole catchment. The areal reduction factors are the same ones used in the FSR and can be found in figure 3.1, page 10, Volume 2 of the FEH.

The FEH method for obtaining growth curves for annual maxima rainfall can be summarised as follows:

- The median of the at-site annual maxima, R_{med} , is used as the index variable;
- Individual durations are treated separately in the construction of growth curves;
- Annual maxima values are pooled from a network of gauges which expands with return period, giving preference to local data;
- Shifted network maximum rainfalls account for inter-site-dependence in rainfall extremes;
- The growth curve is then extended to provide a longer return period;
- To avoid an explicit distributional assumption, the growth curve is comprised of linear segments on a Gumbel scale.

FORGEX is an empirical, graphical method in that it plots points on a rainfall-return period scale and then fits a line through the points. This is in contrast to the approaches that fit assumed distributions using methods such as maximum likelihood or L-Moments.

3.2.1 - Pooling data

Unlike the station year method, the FORGEX technique does not simply group sites and or concatenate their associated data series; instead each series of AM data is effectively plotted on a Gumbel reduced variate scale, so that data from different sites is superposed.

3.2.2 - Gumbel reduced variate – Gringorten plotting position

Annual maxima from individual records, with a minimum length of 10 years, are ranked and allocated plotting positions on a Gumbel reduced variate scale. Following established practice (Shaw, 1994), the Gringorten plotting position formula is used:

Equation 3.2.2.1:
$$F(i) = \frac{(i - 0.44)}{(N + 0.12)}$$

Where $F(i)$ is the non-exceedance probability, 'i' the rank in increasing order, and N the number of annual maxima. The Gumbel reduced variate 'y' is defined by:

Equation 3.2.2.2:
$$y = -\ln(-\ln F)$$

3.2.3 - Definition of y-slices

For the FORGEX method and therefore FEH rainfall return period estimates in Great Britain, each rain gauge network in the hierarchy (based on separation from y the focal point) is associated with the definition of the growth within a particular y-slice. The y-slices have width 1.0 on the Gumbel reduced variate scale, and the first one ends at $y=0.3665$ which is the position of the median ($T=2$ years). Pooled data points are plotted within the jth network to ensure that such data are used in preference to data from further-a-field.

Larger networks include more long-record stations, and thus provide pooled data points that plot in y-slices that correspond to rarer events. However, there are few sites in the UK with records longer than about 100-years. This means that pooled points alone cannot define the growth curve beyond about the fifth y-slice.

3.2.4 - Network maximum points

The network maximum (Netmax) series is defined as the annual maximum series of the largest standardised value recorded by the network of rain gauges. There is one Netmax value for each year of record, across the selected network of rain gauges.

Dales and Reed (1989) showed that the distribution of the network maximum from N independent and identically distributed (iid) Generalised Extreme Value (GEV) distributions lies exactly $\ln(N)$ to the left of the regional growth curve on a Gumbel reduced variate scale, and Reed and Stewart (1994) note that this result is not restricted to the GEV; Figure 5.6.1 – Illustration of the Netmax principle, as explained by Dales and Reed.

In practice, because of inter-site dependence in annual maxima, the Netmax growth curve is found to lie a shorter distance to the left. Dales and Reed label this distance $\ln(N_e)$ terming N_e the effective number of independent gauges.

Thus spatial dependence can be assessed from the relationship between typical and network maximum growth curves. Conversely, the fitting of the regional growth curve can be aided by information on spatial dependence. If an estimate of N_e is available, the top part of the Netmax series can provide valuable information to guide the extension of the regional growth curve to long return periods. N_e could of course simply be estimated from the separation between typical and Netmax growth curves, but a more reliable estimate would combine results from many growth curve analyses.

3.2.5 - Fitting the Growth Curve

The rainfall growth curve is represented by a concatenation of linear segments on the Gumbel reduced variate scale. Because of the standardisation by the median, the growth curve is constrained to take the value 1.0 at a return period of 2 years. Thus fitting the growth curve involves only determining the gradient of each segment. The rules defining the segmentation of the growth curve are explained by Reed et al. (1999).

The growth curve is fitted jointly to pooled and network maximum points by a least-squares routine, which has been adapted to encourage smoothness, i.e. avoid large changes in gradient between adjacent segments.

3.2.6 - Confidence Limits for growth curves

As discussed in Volume 2 of the FEH, an indication of the range of values in which the true growth rate is expected to lie is offered by confidence limits. These indicate the degree of uncertainty in growth rates caused by the limitations of the sample size, but make no account for sources of error such as gauging inaccuracies. The true growth rate could only be known if we had an infinitely long record of rainfall, in which case we could derive the underlying population of annual maxima (assuming no climate change).

Confidence limits are achieved within FEH by the use of ‘bootstrapping’. This method is based on the generation of many re-samples selected from the original sample. Using the FORGEX method, distributions are fitted to the re-sample. This approach is repeated 199 times, with the 5th and 195th values are used to give the 95% confidence limits.

3.2.7 - Trends in Data

Volume 2 of the FEH also discusses that whilst several studies have examined trend in flood frequency, weather types or monthly and annual total rainfalls, there has been little investigation of trend in UK rainfall extremes. Dales and Reed (1989) found no obvious trend in annual maximum 1-day rainfalls standardised by SAAR₄₁₇₀, where SAAR is the Site Annual Average Rainfall. Their study was based on data from 1870 to 1980, using a large number of gauges in England and Wales. Others have found shifts in the frequency of heavy 1-day rainfalls in some areas. For example Perry and Howells (1982) suggested that the frequency of heavy daily rainfall in south Wales has increased through this century.

Volume 2 of the FEH then shows the mean 1-day annual maxima from 1900 to 1990 for 38 rain gauges across the country, most years have close to 38 years of data. There is a substantial year-to-year variation in the mean, but no evidence of an overall trend.

3.2.8 - Regional Frequency Analysis

Regional Frequency Analysis uses standardised annual maxima data from several sites within a region. Spatial samples of data (the at site record) are joined or pooled in substitution of long temporal data sets, which do not exist in the region of interest. It is assumed that the data sets are homogeneous and independent. This method of pooling data is referred to as the ‘station-year method’. However, due to inter-site-dependence, the effective record length is less than the total number of annual maxima in the pooled data set. The effective number of sites for this method is defined as N_e , which can be thought of as the effective number of independent sites that would generate a data set of the same length.

3.3 - Distribution fitting techniques

3.3.1 - Linear Moments (L-Moments)

L-Moments are defined as linear combinations of expected values of order statistics of a variable and are estimated from samples using functions of weighted means of order statistics. The advantages of L-Moments over classical moments are:

- Able to characterise a wider range of distributions;
- More robust to the presence of outliers in the data when estimated from a sample; and,
- They are less subject to bias in estimation and approximate their asymptotic normal distribution more closely

[Hosking, 1990].

Definitions of L-Moments, L-skewness and L-kurtosis:

Given X a random variable with density function f and $(X) < 1$.

L-Moments are defined as:

$$L1 = E(X_{1:1})$$

$$L2 = \frac{1}{2} E(X_{2:2} - X_{1:2})$$

$$L3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3})$$

$$L4 = \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$$

Where:

- $L1$ is a measure of location,
- $L2$ is a measure of spread,
- $L3$ and $L4$ are ratios that measure skewness and kurtosis, respectively, and
- $X_{(i:n)}$ denotes the i th order statistic in a sample of size n .

The ratios that measure L-skewness and L-kurtosis are:

$$\tau_3 = \frac{L3}{L2} \text{ and } \tau_4 = \frac{L4}{L2},$$

where τ_3 is the measure of L-skewness and τ_4 is the measure of L-kurtosis.

The L-Moments of a random variable X exists if X has finite mean. A distribution may be specified by its L-Moments even if some of its classical moments do not exist [Hosking, 1990].

Estimation of L-Moments from a sample:

L-Moments are estimated from samples using functions of weighted means of order statistics. The L-Moments and ratios of L-Moments are estimated by:

$$l_1 = \bar{x}$$

$$l_2 = 2w_3 - l_1$$

$$l_3 = 6w_3 - 6w_2 + l_1$$

$$l_4 = 20w_4 - 30w_3 + 12w_2 - l_1$$

$$\tau_3 = \frac{l_3}{l_2}$$

$$\tau_4 = \frac{l_4}{l_2}$$

where:

$$w_2 = \frac{1}{n(n-1)} \sum_{i=2}^n (i-1)x_{i:n}$$
$$w_3 = \frac{1}{n(n-1)(n-2)} \sum_{i=3}^n (i-1)(i-2)x_{i:n}$$
$$w_4 = \frac{1}{n(n-1)(n-2)(n-3)} \sum_{i=4}^n (i-1)(i-2)(i-3)x_{i:n}$$

Where \bar{x} is the sample mean.

3.3.2 - Maximum Likelihood Estimation (MLE)

Introduction:

The maximum likelihood estimator (MLE) is currently one of the most common parameter estimation procedures. The parameters are computed, either through exact formulas or numerical techniques. Whichever technique is chosen, the objective is to maximize the likelihood function. To understand the likelihood function, it is necessary to understand the concept of likelihood. What follows is an example taken from the following website: http://statgen.iop.kcl.ac.uk/bgim/mle/sslike_3.html

If the probability of an event X dependent on model parameters p is written

$$P(X | p)$$

then we would talk about the likelihood

$$L(p | X) = P(X | p)$$

that is, the likelihood of the parameters given the data.

The argument for using probability, only work for discrete data – where outcomes have a non-zero probability of occurring. For continuous data we use $L(p | X) = P(X | p)$, where P is the probability density.

For most models, it has been shown that certain data are more probable than other data. The aim of maximum likelihood estimation is to find the parameter value(s) that makes the observed data most likely. This is because the likelihood of the parameters given the data is defined to be equal to the probability of the data given the parameters (N.B. technically, they are proportional to each other, but this does not affect the principle).

If we were in the business of making predictions based on a set of solid assumptions, then we would be interested in probabilities - the probability of certain outcomes occurring or not occurring. However, in the case of *data analysis*, we have already observed all the data: once they have been observed they are fixed, there is no 'probabilistic' part to them anymore (the word data comes from the Latin word meaning 'given'). We are much more interested in the likelihood of the model parameters that underlie the fixed data.

A simple example of MLE

To re-iterate, the simple principle of maximum likelihood parameter estimation is: to find the parameter values that make the observed data most likely. Using a simple coin toss experiment, rather than assume that p is a certain value (0.5) we might wish to find the *maximum likelihood estimate* (MLE) of p , given a specific dataset.

Beyond parameter estimation, the likelihood framework allows us to carry out *tests* of parameter values. For example, we might want to ask whether or not the estimated p differs *significantly* from 0.5 or not. This test is essentially asking: is there evidence that the coin is biased? We will see how such tests can be performed when we introduce the concept of a *likelihood ratio test* below.

Say we toss a coin 100 times and observe 56 heads and 44 tails. Instead of *assuming* that p is 0.5, we want to find the MLE for p . Then we want to ask whether or not this value differs significantly from 0.50.

How do we do this? We find the value for p that makes the observed data most likely. As mentioned, the observed data are now fixed. They will be constants that are plugged into our binomial probability model :-

- $n = 100$ (total number of tosses)
- $h = 56$ (total number of heads)

Imagine that p was 0.5. Plugging this value into our probability model as follows :-

$$L(p = 0.5 | data) = \frac{100!}{56!44!} 0.5^{56} 0.5^{44} = 0.0389$$

But what if p was 0.52 instead?

$$L(p = 0.52 | data) = \frac{100!}{56!44!} 0.52^{56} 0.48^{44} = 0.0581$$

So from this we can conclude that p is more likely to be 0.52 than 0.5. We can tabulate the likelihood for different parameter values to find the maximum likelihood estimate of p :

p	L
0.48	0.0222
0.50	0.0389
0.52	0.0581
0.54	0.0739
0.56	0.0801
0.58	0.0738
0.60	0.0576
0.62	0.0378

Table 3.3.2: Shows the likelihood for different parameter values

If we graph these data across the full range of possible values for p we see the following *likelihood surface*, as illustrated in figure 3.3.2, overleaf.

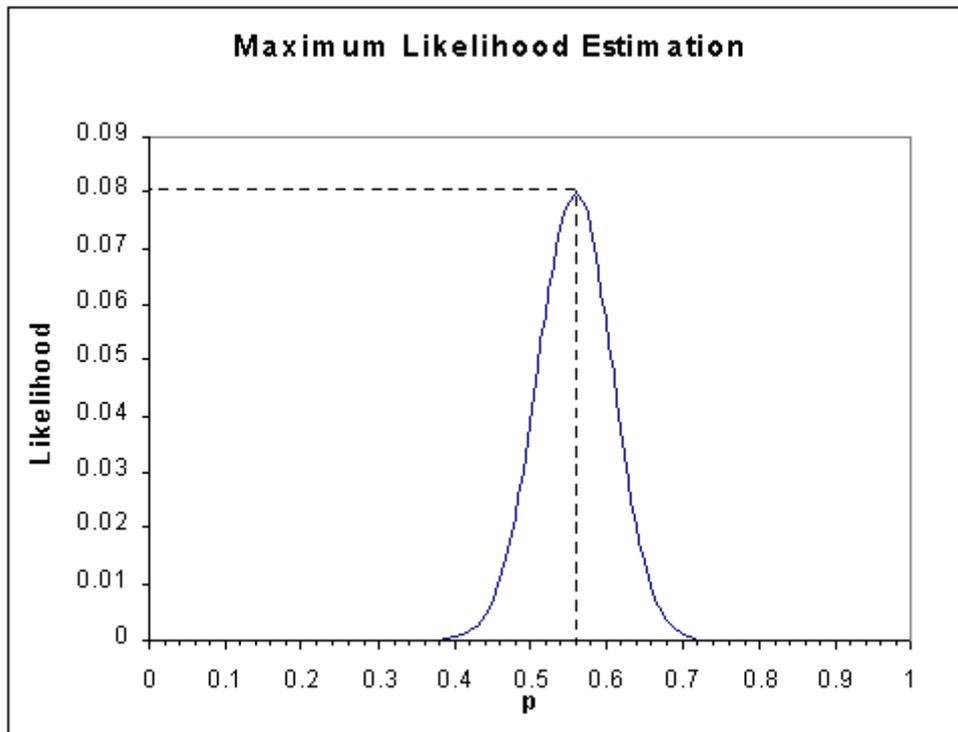


Figure 3.3.2: Likelihood surface, showing maximum likelihood for a number of parameter estimates.

We see that the maximum likelihood estimate for p seems to be around 0.56. In fact, it is exactly 0.56, and it is easy to see why this makes sense in this example. The best estimate for p from any one sample is clearly going to be the proportion of heads observed in that sample. (In a similar way, the best estimate for the population mean will always be the sample mean.)

For such a simple case, one might not expect to use such a complicated method.

However, if you use the simple frequency estimate $\hat{p} = \frac{56}{100}$, then you are using maximum likelihood estimation, even if it is unwittingly.

However, not all problems are this simple. This thesis goes on to examine more complicated models with a greater number of parameters, where it is often very difficult to make even reasonable guesses at the MLEs. The likelihood framework conceptually takes all of this in its stride however, and this is what makes it the work-horse of many modern statistical methods. It also has other good properties. For example it has a limiting Normal distribution, which provides a general theory for deriving the uncertainty associated with the MLE.

3.4 – Probable Maximum Precipitation

This section looks at the theory of Probable Maximum Precipitation (PMP), which is an alternative method of extreme rainfall estimation.

3.4.1 - The Theory of Probable Maximum Precipitation (PMP)

The world meteorological organization [WMO, 1986], defined PMP as:

“theoretically the greatest depth of precipitation for a given duration that is physically possible over a given size storm area at a particular geographical location at a certain time of year, with no allowance made for long term climatic trends”.

It is thus seen as a single deterministic number (governed by physical principles) that would never be exceeded.

The theory of Probable Maximum Precipitation (PMP) started as maximum possible precipitation. It was based around the concept that there are maximum physical limits for all of the elements which act together to produce rainfall [U.S. Department of Commerce, Washington D.C., 1960]. This method of peak rain fall estimation is frequently favoured by engineers who must design a structure to withstand a theoretical flood, referred to as the Probable Maximum Flood (PMF); one assumption being that the PMF is generated from PMP and that PMF can be calculated using a rainfall-run-off model which has been calibrated using a rain-gauge and stage recorder in the catchment. Another reason why this method might be favoured over statistical alternatives, is that the availability of rainfall records is relatively scarce, especially in less developed countries; even when rainfall data is available it is unusual to have complete records exceeding three or four decades. These two factors have made it very difficult to estimate rare or extreme rainfall events, which will ultimately generate large amounts of run-off and potentially floods.

Structures such as reservoirs must be able to safely pass these flows; as however unlikely it is believed to be, the potential loss of life should it occur is obviously unacceptable. For a long time, this method of flow estimation was seen as the safest way to minimize the risk of structural failure and hence protect the population which might be affected by such a failure. This is obviously an attractive proposition, a method which produces the greatest possible rainfall and hence run-off, which is perceived to remove all risk of failure.

3.4.2 - The Problem with PMP estimates

Unfortunately, however, there have been numerous occasions on which observed rainfalls have exceeded the PMP estimates. PMP estimates have also been produced that are believed to be unrealistically high by many experienced hydrologists.

These occurrences have led to a re-evaluation of PMP methodologies, of which there are many. Some of the more common techniques used for estimating PMP are:

- Storm Model Approach.
- Maximisation and transposition of actual storms
- The use of generalized data or maximised depth, duration and area data from storms; these are derived from thunderstorms or general storms.
- Use of empirical formulae determined from maximum depth duration and area data, or from theory;
- Statistical analyses of extreme rainfalls.

Many of these techniques have undergone numerous revisions over the years because they have not adequately described or estimated PMP at locations other than those chosen for the validation / calibration. Another reason why many of these techniques have been revised or modernised is the accessibility and power of modern computers. Computers have made it plausible to run more complicated PMP simulations as demonstrated by the Storm Model Approach and the Generalized Method of PMP estimation.

3.4.3. Storm Model Approach

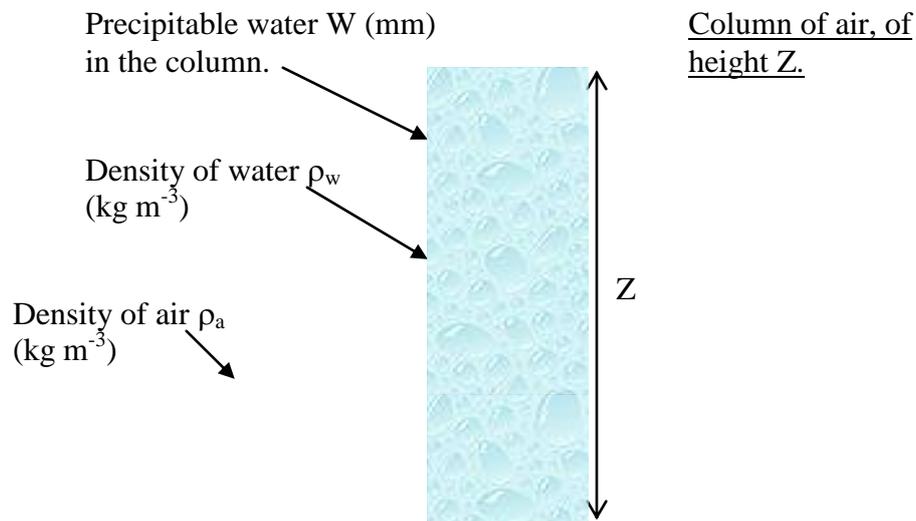


Figure 3.4.0: Schematic of precipitable water in a theoretical column of air.

The amount of precipitable water, W (mm), in a column of air of height Z is defined as:

$$W = \int_0^Z \rho_w dZ = - \int_0^{\rho_z} \frac{\rho_w}{\rho_a g} dp = \int_0^{\rho_z} \frac{q}{g} dp \quad \text{Equation 3.4.0}$$

Where:

q = specific humidity (or approximately the mixing ratio; in g kg^{-1}).

p = atmospheric pressure ($\text{hPa} \times 10^{-2}$).

g = acceleration due to gravity, 9.81 m s^{-2} .

These techniques were developed approximately thirty years ago, yet efforts are being made to build on this foundation using computer-based approaches. In addition to this, new sources of data have become available, particularly radar and satellite data which allow a much more detailed and accurate model to be built up.

3.4.4 Maximisation and Transposition of actual storms

PMP has historically been estimated in a variety of ways using techniques that maximise recorded storms. When the storms maximised are only those that occur on the catchment under consideration, the method is called the “in-situ maximisation” method. When storms that occur in adjoining and geographically similar regions to the catchment area (figures 3.4.1 and 3.4.2) are also considered, the method is called the “transposition and maximisation method” [WMO,1986:].

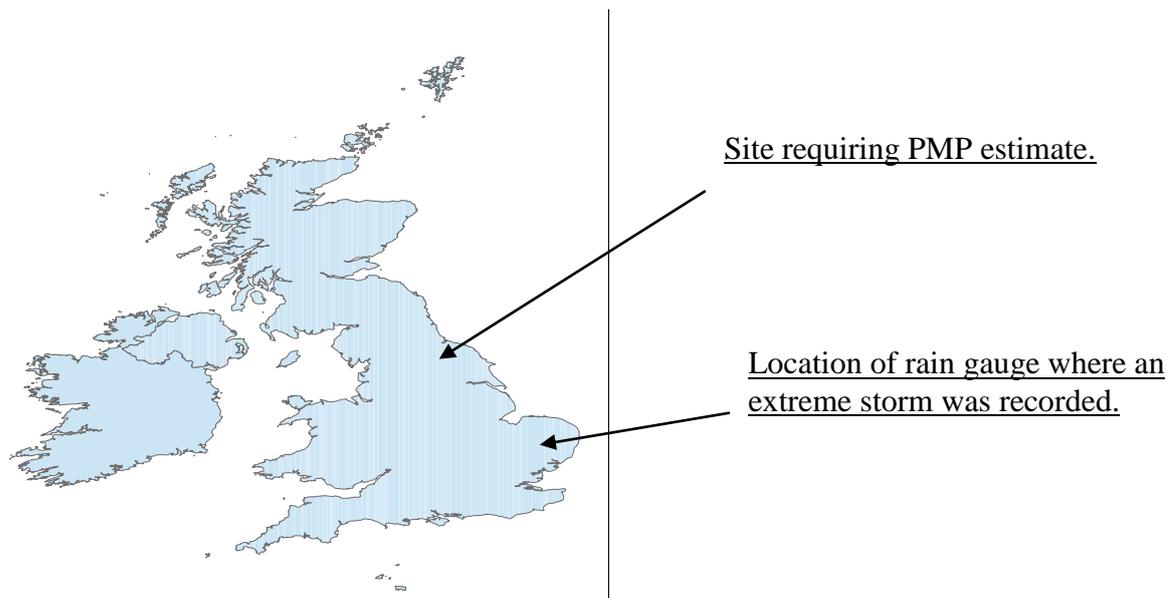


Figure 3.4.1: Map showing donor site and target site for an Extreme Event

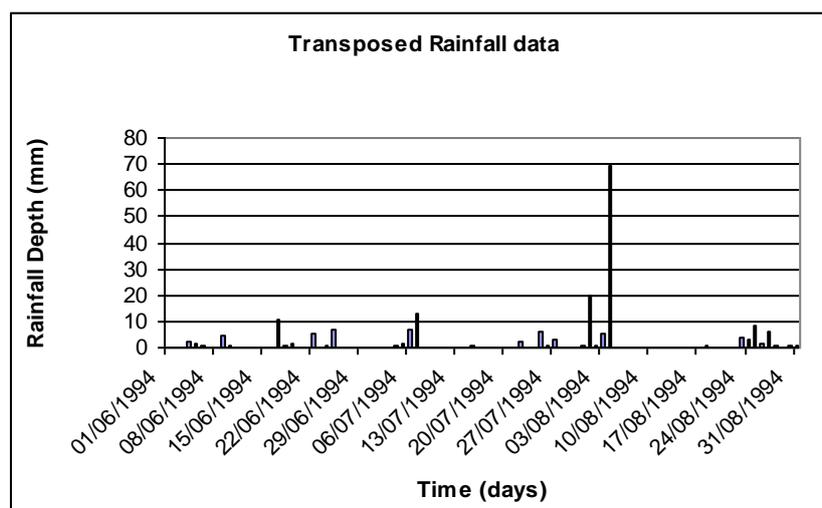


Figure 3.4.2: Transposed rainfall data from a donor catchment.

Observations of recorded storms are transposed over the region or catchment of interest. The rainfall is then maximised by using factors for orographic enhancement and other meteorological considerations, which are briefly described below.

Hart [1982] shows that the physical basis for storm maximization is based on a simple two-parameter model of the storm derived as follows. A storm is considered to consist of a convergent mass flow at low levels that rises and diverges in an upper outflow layer. The water vapour budget equation associated with the storm can be written as:

$$E - P = \int_0^{p_0} \left[\frac{\partial q}{\partial t} + \Delta \cdot (qV) \right] \frac{dp}{g} \quad \text{Equation 3.4.1}$$

Where: E is evaporation, P is precipitation, q is specific humidity, V is the horizontal wind vector, g is the acceleration due to gravity, p₀ is the surface pressure, and the vertical integration is carried out over the depth of the atmosphere. For major storms it is assumed that the evaporation term E, the rate of water vapour storage term (δq/δt), and the moisture gradient in the vicinity of the storm are negligible. With these assumptions, Equation 3.4.1 can be rewritten as:

$$P = - \int_0^{p_0} q \Delta \cdot V \frac{\partial p}{g} \quad \text{Equation 3.4.2}$$

That is, the rainfall is approximately equal to the vertically integrated product of the mass convergence and the specific humidity. If the model is further simplified to comprise an inflow layer Δp₁ and an outflow layer Δp₂ with uniform flows, D₁ and D₂ and specific humidities q₁ and q₂, the precipitation P reduces to:

$$P = - \left(\frac{q_1 \Delta p_1 D_1 - q_2 \Delta p_2 D_2}{g} \right) \quad \text{Equation 3.4.3}$$

From considerations of mass continuity, Δp₁D₁ = Δp₂D₂ and q₁ >> q₂, and hence the precipitation is approximated by:

$$P \approx - \frac{q_1 \Delta p_1 D_1}{g} \quad \text{Equation 3.4.4}$$

To calculate the maximised precipitation, the product of the moisture inflow and mass convergence needs to be maximised. The term q₁Δp₁ is the effective precipitable water (w_e) for a storm, and this can be maximised by using 24-hour persisting dew points to calculate the maximum effective precipitable water (w_{e,max}). The maximised precipitation is then calculated by adjusting the observed rainfall by a moisture

adjustment factor w_{emax}/w_e . However, as pointed out by Wiesner [1970], it is common practice to calculate the moisture adjustment factor from the actual precipitable water in a saturated atmospheric column and the maximised precipitable water w_{max} as given by the maximum 24-hour persisting dew points. The dew point uniquely defines the mixing ratio at cloud base and therefore the precipitable water in the saturated column. This indirect technique arises because there is usually no way of characterizing the extreme mass convergence, so the observed rainfall is taken as an implicit measure of this quantity. It is assumed that extreme precipitation storms have the highest efficiency. The maximised precipitation P_{max} is thus calculated from the precipitable water w derived from the observed dew point, the maximised precipitable water w_{max} , and the observed rainfall P (normally in the form of depth-duration-area (DDA) curves) as:

$$P_{\text{max}} = \left(\frac{w_{\text{max}}}{w} \right) P \quad \text{Equation 3.4.5}$$

The maximised precipitation is then calculated by adjusting independently the assumptions used in the simple two-parameter conceptual model that is used for PMP calculations. These assumptions are:

1. The precipitation is linearly related to the precipitable water (i.e., $P_2 = (w_2/w) * P$);
2. The precipitation efficiency of the storm does not change as the moisture available to the storm increases;
3. Terrain modulates the distribution of the precipitation but does not affect the synoptic-scale dynamics of the storm.

The relationship between the precipitable water and the precipitation (assumption 1) is particularly important since it is this relationship that underlies the foundations for both the moisture maximization and the storm transposition techniques currently employed in the GSAM (The Australian Bureau of Meteorology has developed three generalized methods that are applicable to the country (Australia): the generalized short duration method (GSDM), the generalized tropical storm method (GTSM), and the generalized south eastern Australia method (GSAM). The GSDM is applicable for small areas up to 1000 km² and for time periods up to 6 hours. The GTSM and the GSAM are used for larger areas of the order of 10⁴ km²).

The report of the National Research Council [1988] also concludes that the scientific foundations of the traditional PMP procedures, such as moisture maximization and storm transposition, require detailed study. The report points to numerical models as key tools for enhancing PMP procedures.

The following steps are used to evaluate the assumptions detailed above:

1. Use a numerical model of the atmosphere to simulate recent large storms.
2. Compare the model results with the observed rainfall and storm development.
3. Carry out sensitivity analyses to determine the maximum precipitation efficiency of the storms.
4. Develop a hypothetical “worst case storm” that would allow a comparison between the model-generated DDA curves and the DDA curves calculated using the maximization relationship of the current generalized technique.

There are a number of problems with this and similar methods:

1. The assumption that the record is long enough to have captured a truly extreme event approaching some theoretical upper-limit, not just a large storm.
2. The assumption that from a small number of rain gauges one happened to be located suitably to describe the storm when it reached its peak, see figure 3.4.3.
3. A large percentage of extreme rainfall events occur over a relatively short duration (hours or less) and the majority of rain gauges record a 24 hour total.

The location of the rain gauge may not be flawed but it is possible that it has not recorded the peak rainfall. Should the rainfall distribution be uneven, that is, it is more intense at one location than at another, and the point of maximum intensity has not occurred over the rain gauge, then the design of the structure is flawed.

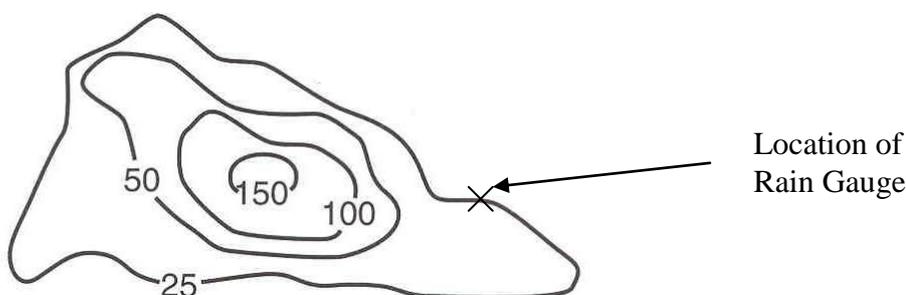


Figure 3.4.3: Accumulation from a convective rainfall storm in the UK

Looking at figure 3.4.3 it is easy to appreciate point 2 above. If the rain gauge is not in the ideal location, the recorded rainfall depth for the duration of the storm could vary from approximately 25mm to 150mm, assuming the rain gauge was located to capture the storm at all. This may not be a significant problem for a large catchment, but for a small, ‘flashy’ upland catchment, supplying a reservoir, this could have a significant impact.

More recent developments have used radar data to improve the descriptions of the spatial structure of the storms and to capture the maximum rainfall intensity within the storm.

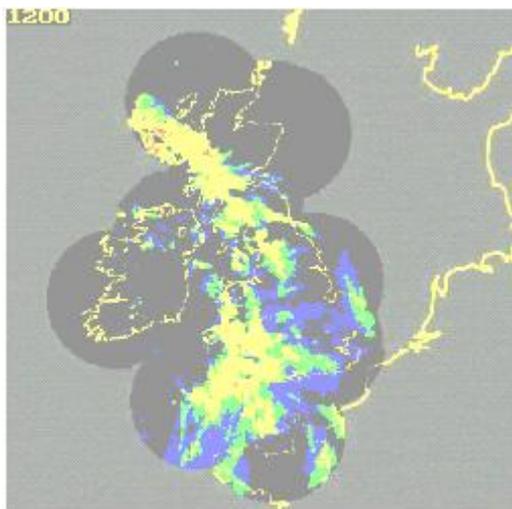


Figure 3.4.4: Radar image of a frontal weather system (November 2000, source: Hyrad)

3.4.5 - Generalized Method

More recently, a technique known as the “generalized method” has been developed to calculate PMP. This method uses rainfalls recorded over a large region and from a large database of storms. The storm database is generalized by separating out that portion of the rainfall attributable to regional meteorological conditions from that which may be considered to be due to site-specific (for example topography) characteristics.

Both the maximisation and transposition method and the generalized method involve the classification of storms by calculating the corresponding storm efficiency (E), which is defined as the ratio of maximum observed rainfall to the amount of precipitable water in the representative air column during the storm [NERC, 1975].

3.5 - Reservoir Flood Estimation, procedures and developments in other European countries; 10,000-year return period estimates

3.5.1 - Introduction

Having looked at the current flood estimation methods for the UK, an analysis of the methods applied by some European countries has been investigated. The aim of this chapter is to gain a greater understanding of the approaches used, for a comparison to those used in the UK.

Following some research, it was discovered that the Institute of Hydrology had carried out a review of reservoir flood estimation in a number of other countries. Where the information was available, this details the requirements and the approaches used in the Finland, Sweden and Norway, as discussed in “Reservoir flood estimation: another look” by Reed & Field (1992).

3.5.2 - Finland

New legislation came into force on 1st August 1984 regarding dam safety. This legislation applies to dams that are at least 3m in height or which pose a particular hazard.

The Dam Safety Act recognises four categories of dam:

Category P dams are those which in the case of an accident will endanger life or health, or cause serious damage to the environment or property.

Category O dams are those which in the event of an accident will cause will cause only minimal danger.

Category N dams are those which present an intermediate hazard.

Category T dams are temporary structures.

Table 3.5.2 shows the range of return periods associated with categories.

Category	Return period range
P	5,000 to 10,000 years
N	500 to 1,000 years
O	100 to 500 years

Table 3.5.2: Spillway design floods: Finnish practice.

Shorter return periods may be considered adequate for temporary dams.

The basic recommended method for estimating the design flood is the extrapolation of a Gumbel distribution fitted to the annual maximum series of gauged floods. Many Finnish rivers are heavily regulated by reservoirs and lakes. During most years, snowmelt floods occur regularly, with April and May characteristically providing the annual maximum floods. However, average flood growth-curves in Finland are no shallower than those in the UK (Gustard et al., 1989). Thus, the recommendation by Loukola et al. (1985) to base spillway design floods on simple extrapolation of peak flow records at the dam site is extraordinary. If followed, the guidance could lead to gross under- or over-estimation of design floods for a particular dam through over-reliance on statistically very short data series.

It is unclear how the peak flow estimate is converted to a hydrograph for the purpose of routing the design flood through the reservoir storage.

3.5.3 – Sweden

A distinctive feature of Sweden is the use of a unique 14-day design rainfall profile of unknown return period (possibly about 10,000-years). Corrections are made for geographical region, catchment area and altitude. The basis of the latter is unclear but other publications (e.g. Vedin and Eriksson, 1988) suggest that the adjustment derives from a ‘storm-centred’ rather than a ‘fixed’ areal reduction factor.

3.5.4 – Norway

On 1st January 1981, new regulations came into force in Norway regarding permanent dams more than 4m in height or which impound more than 500,000 m³. These regulations require that flood calculations be performed for both the design flood and the Probable Maximum Flood (PMF). The PMF sets the standard for dam safety, with the design flood setting the standard for normal spillway operations. A 1,000-year return period is specified for the design flood, which is determined by some type of frequency analysis of peak flows (see below). The PMF is calculated on the basis of Probable Maximum Precipitation (PMP) values and snow melt estimates, with allowance being made for reservoir routing.

Estimation of the 1,000-year flood:

The guidelines (Vassdragsdirektorat, 1986) recommend that several statistical distributions be considered when seeking an estimate of the 1,000-year flood.

If more than 50 years of annual maximum flood data are available, the mean annual flood (QBAR) is estimated from the observed series, while the growth factor (Q1000/QBAR) is taken from a two- or three-parameter G.E.V. distribution fitted to the observed series. If only 30 to 50 years are available, a two-parameter distribution is to be used. If fewer than 30 years of data are available, the Q1000/QBAR growth factor is based on regional analysis. If fewer than 10 years of data are available, QBAR is estimated by correlation with other series in the region or by catchment characteristic formulae.

In many cases, it is deemed appropriate to distinguish spring (largely snowmelt) and autumn (largely rainfall) floods. The spring floods yield a high QBAR and large hydrograph volume, but have only moderate growth rates, meaning that the curvature of the growth curve is moderate. In contrast, the autumn floods stem from shorter duration events of high intensity to which steeper growth curves apply.

A rainfall-runoff approach to estimating the 1,000-year flood is not generally recommended. This is because of the 'joint probability problem' of choosing appropriate initial catchment wetness and snowmelt/snow accumulations to combine with a 1000-year precipitation event to produce the required 1000-year flood.

Saelthun and Andersen (1986) describe what appears to be a fairly subjective method for converting statistically-derived estimated of the 1,000-year peak instantaneous and/or peak 1-day flow into a design hydrograph suitable for reservoir routing. They caution against the practice of nesting 1,000-year flows of different durations within a single design hydrograph.

3.5.4 – United States of America

Approaches to reservoir spillway design flood estimation are more varied across the USA than they are in the United Kingdom (Reed & Field, 1994). In part this arises from more diverse climatic and physiographic conditions; however, in part it may reflect the weaker institutionalism of reservoir flood estimation (Reed & Field, 1994). The interagency advisory committee on water data (1986) provides a useful summary

of US approaches to PMF and extreme flood estimation. The report recognises five approaches:

1. Extrapolation of flood frequency curves;
2. Combination of frequency distributions of casual factors (e.g. antecedent reservoir level and storm rainfall);
3. Regional approach to extrapolation (e.g. the station- year method);
4. Palaeoflood analysis (e.g. inferring historical flood levels by the position and dating of sediments); and,
5. Bayesian analysis (combining different sources of flood data, e.g. local, regional and historical).

3.6 – Summary

This chapter has described the techniques used by the Flood Estimation Handbook to produce and to measure extreme values.

This chapter has also considered the theory of Probable Maximum Precipitation (PMP), which is an alternative method of extreme rainfall estimation. Estimates of PMP are regarded as approximations which depend upon the amount and quality of the data available for applying the various methods. Further, as the WMO description of PMP states, there is no allowance for long term climatic trends.

This statement appears to have more and more significance in the light of research showing approximately a 0.5°C increase in global temperature over the past 30 years. More alarmingly, it is forecast that this increase will continue and that over the next 100-years the global mean temperature could increase by between 1 and 5.5°C (Intergovernmental Panel on Climate Change).

Perhaps most significant is the fact that when a very large estimate is produced it is often ignored because ‘experts’ believe it to be too large; however, PMP estimates have also been exceeded.

This chapter has also introduced current flood estimation methods for the UK and a number of European countries. It has introduced and explained some of the techniques described by and associated with the Flood Estimation handbook, which is in turn the accepted UK standard for flood estimation. In addition to this there has been a technical description of two distribution fitting techniques. The performance

of these distribution fitting techniques will be analysed in more detail in the next chapter; Chapter 4.

The remainder of this study will then focus on statistical methods of extreme value estimation for the following reasons:

- Statistical techniques associate probabilities with rainfall magnitude; and,
- Uncertainty is addressed in the form of confidence intervals.

Chapter 4

This page is intentionally blank.

4.0 – The Stationary Model

4.1 - Introduction

Using synthetic data, this chapter aims to gain a better understanding of:

1. Two distribution fitting techniques: L-Moments and Maximum Likelihood Estimates (MLE) (introduced in chapter 3.3); and,
2. The impact of spatial dependence upon pooling groups and therefore Regional Frequency Analysis.

Currently, the majority of rainfall and flood estimates are carried out using the assumption that the data is stationary. This chapter will continue to make this assumption.

Stationary data sets can be defined as having statistical properties that do not change over time; more precisely, the probability distributions of the process are time-invariant. The mean, variance and covariance of the process are defined as follows:

$$\text{Mean: } \mu(t) = E\{X_t\},$$

$$\text{Variance: } \sigma^2(t) = \text{Var}(X_t) = E\{(X_t - \mu(t))^2\}, \text{ and}$$

$$\text{Covariance: } \gamma(s, r) = \text{Cov}(X_s, X_r) = E\{(X_s - \mu(s))(X_r - \mu(r))\}.$$

Where: E denotes the expectation of a random process.

A simple summary of stationary data sets then, is that the mean, variance and covariance of the distribution do not change with time.

Following the approach of Matalas (1967) and Hosking and Wallis (1988), a spatially dependent, multi-variate model has been produced. The model generates spatially dependent annual maximum data, and has been used to demonstrate, and gain a better understanding of the effects of inter-site correlation upon pooling groups. The pooled annual maximum rainfall data (for each region within Great Britain as shown in Figure 1.2b) has then been fitted to, using the two techniques already described.

The standard approach to the estimation of extremes in hydrology (for example flood and rainfall data) is to use annual maximum series. The method of L-Moments was developed by Hosking (1990) for fitting extreme value (EV) distributions for flood frequency estimation and was adopted by the Flood Estimation Handbook (FEH). More recently, computational developments have led to the more widespread adoption

of Maximum Likelihood Estimation (MLE). However, L-Moments are still widely used by the hydrological community. This chapter shows that although there are good reasons for using L-Moments, there are also many advantages to using MLE.

To carry out the comparison of L-Moments and MLE, synthetic data has been used, firstly for single site analysis and then for multi-site analysis. The single site analysis has been achieved by using a random (normal) number generator to simulate (select) extreme events from a known distribution – this is the synthetic data. The two distribution fitting techniques have been compared by looking at the fitted quantile estimates and comparing them with the known, true values. The range of errors has been displayed in the form of confidence intervals from the generated data for each technique at chosen quantiles.

A similar comparison has taken place for regional frequency analysis. The data has been generated using a Multi-Variate Normal Random Number Generator. This method uses either:

1. Observed inter-site correlation from a known region; or,
2. User defined inter-site correlation.

It is possible therefore to demonstrate the effect of inter-site-dependence on:

1. Quantile confidence intervals.
2. The effective number of sites in the region, which is typically less than the actual number of sites in the pooling group.

This is shown in chapter 5.6 and 5.7.

Chapter 5.8 demonstrates an alternative method of homogeneity, which uses the likelihood value from the MLE fitting technique. The homogeneity test is a test of whether a sample distribution (individual site) belongs to a parent distribution (regional pooled group); this test is called the Likelihood Ratio Test (LRT). This method has been used to test predefined regions and also focused regional growth curve expansion methods such as the FORGE approach, which stands for FOCused Regional Growth curve Expansion.

4.2 - Comparison of two distribution fitting techniques: L-Moments and Maximum Likelihood Estimates (MLE)

Figure 4.2.1 shows why a comparison is required. The technique used for this test was:

1. Start by selecting a donor set of GEV parameters – for either a single site or pooling group of annual maxima rainfall data;
2. Generate a synthetic sample data set using a Monte Carlo technique randomly selected data from the known distribution – this was used to simulate a single site with 40 years of data; and,
3. Using the two techniques of L-Moments and MLE, fit a GEV distribution to the sample data.

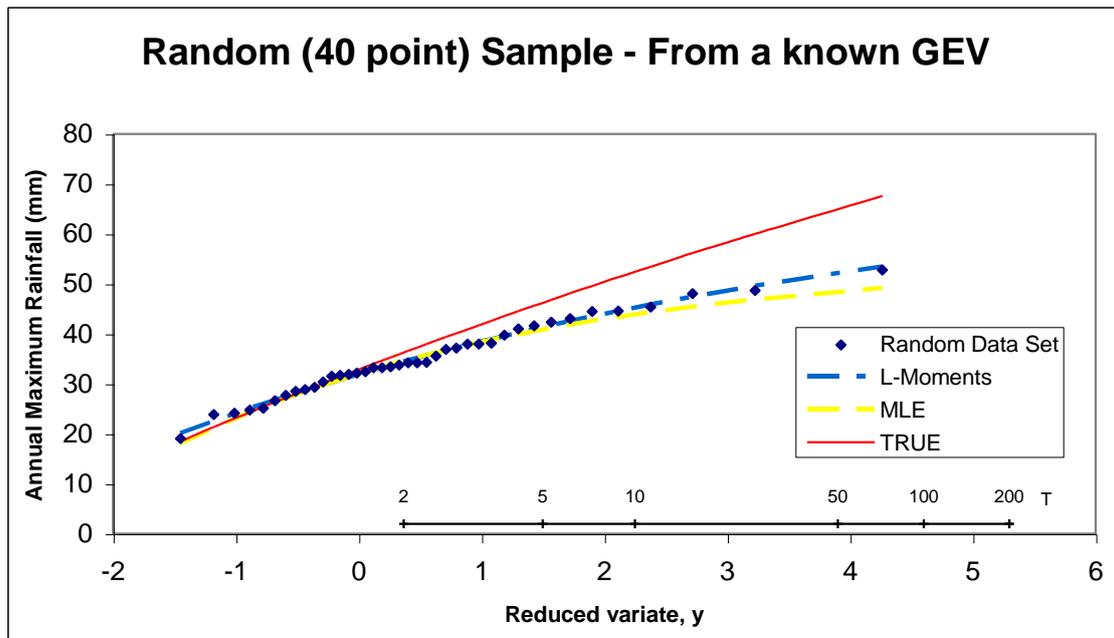


Figure 4.2.1: Comparison of L-Moments and MLE, for a short synthetic data set.

Figure 4.2.1 shows that, for this sample only, both techniques will underestimate at all quantile values when compared with the ‘true’ distribution. It can also be seen that the MLE technique, on this occasion, produces the greater error.

If this test is repeated multiple (10,000) times and the distribution curves are extrapolated to include the 100-year return period event, then the following results are obtained:

Simulated - 100 Year Return Period Events

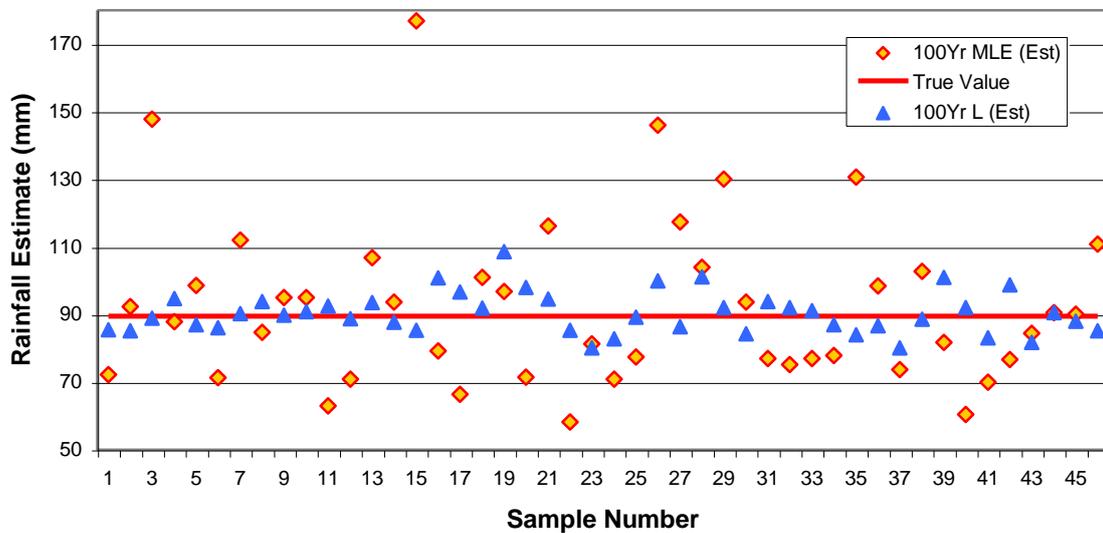


Figure 4.2.2: Comparison of L-Moments and MLE, for the 100-year Return Period Event.

Figure 4.2.2 shows a small sample from the 10,000 estimates produced for the 100-year rainfall estimate, using a single site. The size of the sample is chosen to assist the reader, as it displays a wide range of results. Primarily this figure shows that there is a larger spread of estimates (greater uncertainty) when using MLE than is produced by L-Moments.

4.3 - Method for Single site distribution fitting technique comparison:

The method adopted for comparison has been as follows:

- Starting with a time series of 24 hours (1 day fixed duration) annual maximum rainfall data, taken from a site in the UK, GEV distribution parameters were calculated using L-Moments and used as the control parameters for the following simulation.
- A random (normal) number generator was used to select points from the control distribution. This method was repeated to give 10,000 samples representing: 20, 40, 60 and 100 values (values in this case meaning the effective record length or the number of Annual Maxima at the synthetic site).

- Using the methods of L-Moments and MLE for parameter estimation, return period estimates were calculated and compared with the known true value. This comparison revealed the relative error for each estimate.
- All simulation work was carried out using a free software package and programming language called R, see: www.r-project.org/ for more details.
- The random number generator is available within R and called: rnorm.
- A routine was written to generate synthetic annual maxima, by randomly generating a vector of length n, and with values x, where: $0 \leq x \leq 1$. These random values are interpreted as Gringorten plotting positions, which when combined with known (control) GEV parameters, allows site specific synthetic extreme values to be generated.

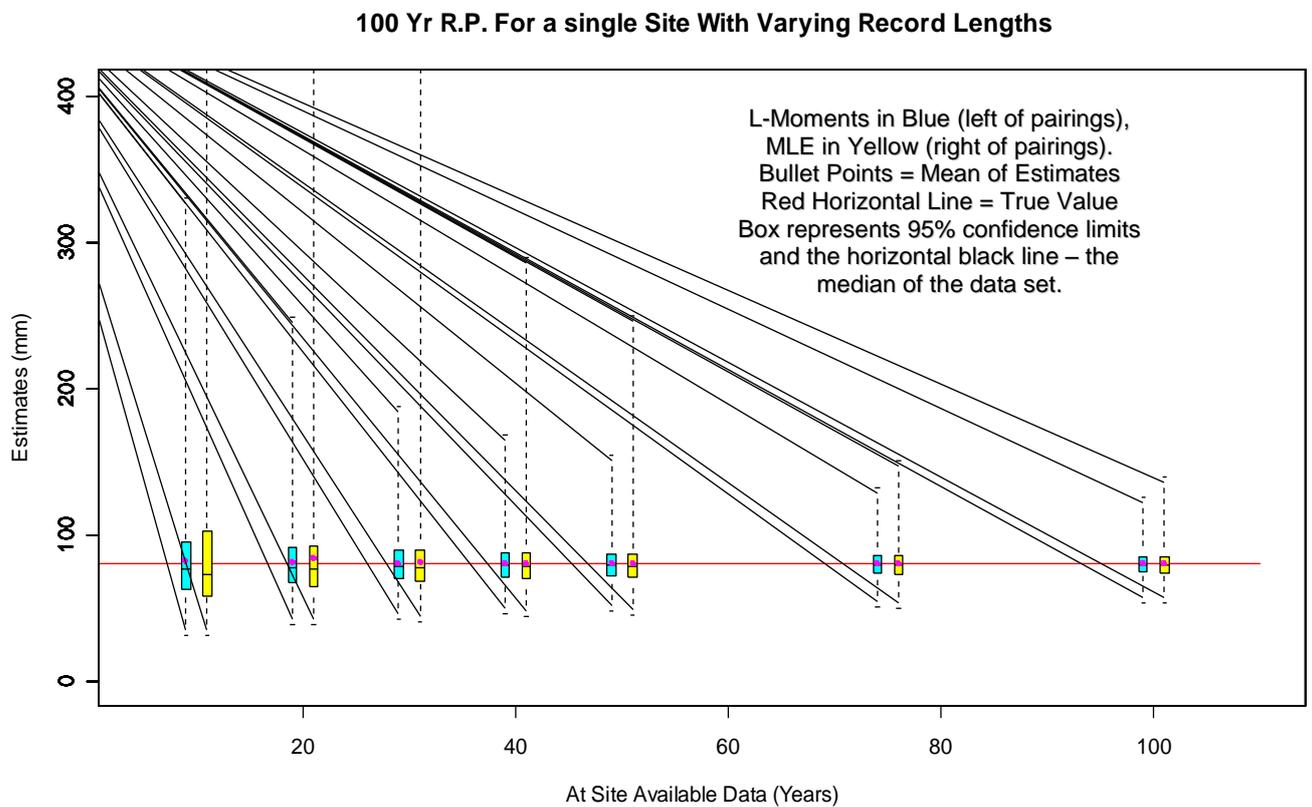


Figure 4.3.1: Summary of Estimates and Range of Errors produced by the two techniques.

(Some of the outliers for the 10, 20, and 30 year synthetic data sets have been excluded from view because they are too extreme and their inclusion would diminish the clarity of the graph.)

Summary of figure 4.3.1:

1. The range of errors (confidence interval and outliers) is reduced as the size of the data set is increased.
2. Figure 4.3.1 does not show the density of errors associated with the two techniques.
3. Clearly demonstrates the need to include confidence limits for return period estimates based upon the amount of data available and the resulting confidence, because, as already stated, figure 4.3.1 shows the confidence limits and outliers (most extreme estimates) produced when using each technique.

Although confidence intervals have been shown in figure 4.3.1 this has only been possible because from the 10,000 generated estimates the values were ranked and the 250th and 9750th values were used to show the 95% confidence interval, the outliers therefore account for the remaining 5%.

To achieve these results, routines have been created to generate and analyze the data as shown using the language 'R'.

R is a language and environment for statistical computing and graphical representation of the data. It is similar to the S language and environment that was developed at Bell Laboratories (formerly ATandT, now Lucent Technologies) by John Chambers and colleagues. There are some important differences between the two languages, but much code written for S runs unaltered under R [http://www.r-project.org/doc/R-FDA.pdf].

R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, and so on) and graphical techniques, and is highly extensible. The 'S' language is often the vehicle of choice for research in statistical methodology, and 'R' provides an Open Source route to participation in that activity [http://www.r-project.org/doc/R-FDA.pdf].

It is important to note that only MLE is capable of generating confidence intervals as a result of its use. L-Moments requires a form of re-sampling of the data, known as boot-strapping. One advantage of using MLE therefore is the instantaneous generation of confidence intervals. An example of this follows:

Focused Growth Curve Plot

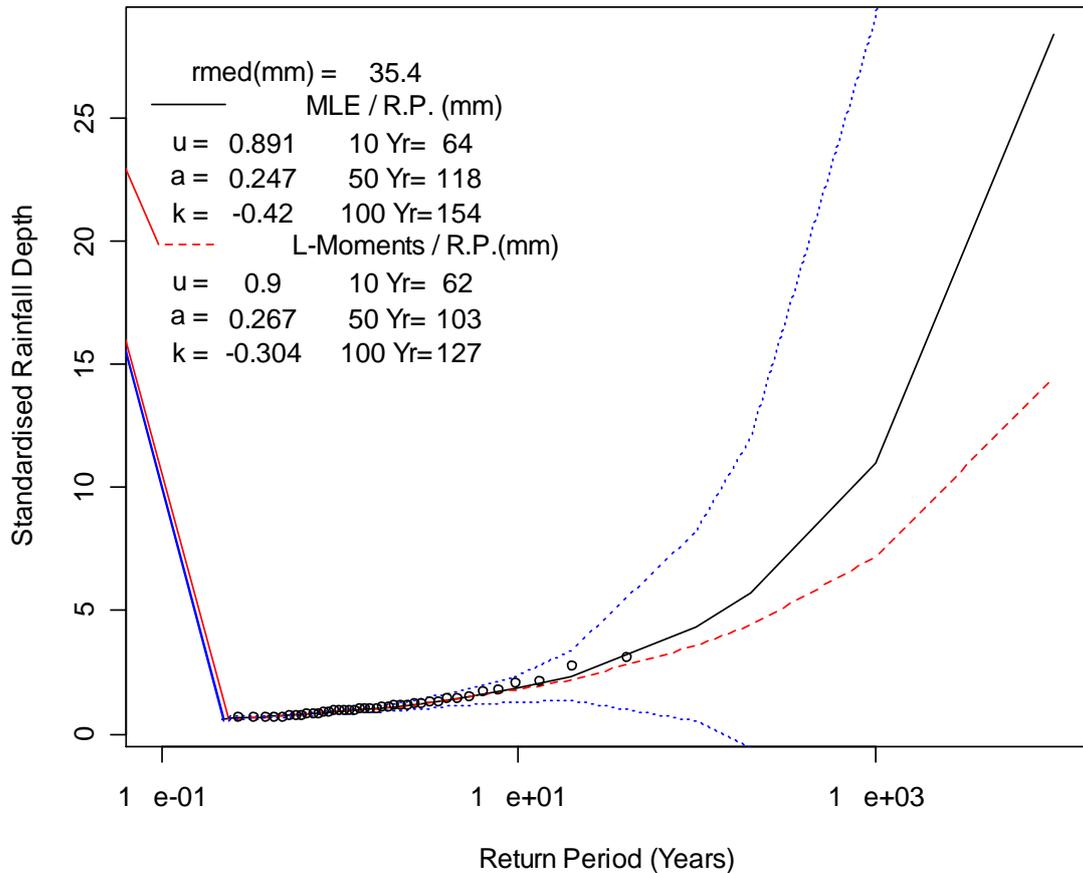


Figure 4.3.2 – MLE and L-Moment fit to annual maxima rainfall data; showing the confidence interval (faint, finely dashed lines, top and bottom) generated as a result of using MLE.

Figure 4.3.2 has been included for three reasons:

1. It shows the confidence interval generated as a result of using MLE
2. It shows that MLE does not always underestimate when compared with L-Moments, which may have been inferred in figure 4.2.1
3. It demonstrates one of the tools created during this thesis – using ‘R’ to aid with this investigation

4.4 – Summary

This chapter has demonstrated the strengths and weaknesses of two distribution fitting techniques. It has shown that the method of L-Moments appears to be more accurate for relatively short time series, at least for the distribution tested, but it also shows that any advantage demonstrated by L-Moments diminishes as the length of the time series increases.

Figure 4.3.2 shows that the confidence intervals for the estimates are very wide, especially beyond the 10 year return period event. Chapter 5 goes onto to look at accepted methods to reduce uncertainty and begins to quantify how accurate they are.

Chapter 5

This page is intentionally blank.

5.0 – Synthetic data generation for a region of N-sites with Spatial Inter-Site-Dependence

5.1 - Introduction

Following on from chapter 4, the next step was to look at methods of pooling data. When producing return period estimates, the recommended length of a time series is 5 times the required return period event. For example, the 10-year return period event requires a recommended minimum of 50 years of data and therefore for the 100-year event, ideally you would have a time series that is equivalent to 500 years of data. The only way to get long time series like these is to pool data from multiple sites in a region.

To gain a better understanding of the factors controlling uncertainty when pooling data, all of the analyses have been carried out using synthetic data. This approach allows the distribution parameters and inter-site-dependence (correlation) to be known; aiding understanding with regard to the effect of these on uncertainty.

Background

Rain gauges are typically operated by the Meteorological Office, water utilities and the Environment Agency. They are often located near to reservoirs, airfields, universities and other academic establishments; it is not unusual for a rain gauge to be located in a private garden, where that person has an interest in meteorology. Rain gauges are not located in a uniform manner and it is very often the case that rainfall data is required at a site without a gauge. When this happens it is standard practice to transfer data from one or more neighbouring sites using interpolation and or some scaling factor, perhaps based on elevation. Inter-site correlation (dependence) in this situation is clearly advantageous. This chapter however, is going to focus on the detrimental impact of inter-site-dependence.

Inter-site-dependence, or the correlation between sites, causes a duplication of data when it is pooled. Pooling groups are used in an attempt to augment a data set, for the purpose of generating greater accuracy for rare (extreme) event estimation. The level of correlation is important, as the size of the pooling group may have been selected to

achieve a station year total. The station year total is the product of the number of sites (N) multiplied by the number of years (n). For example, the desired station year total might be five times greater in length than the required return period estimate.

This chapter will also show that for Great Britain, the effective number of sites in a region or pooling group using the Station Year method, ranges from 74% – 93% of the total for 1 Day Annual Maxima (AM), and 61% - 88% of the total for 10 Day Annual Maxima. This means that a pooling group of 10 Sites, each with 40 years of data, does not equate to a time series of 400 station years in length, but to one of perhaps $74\% * 400 = 296$ station years for 1 Day AM or $61\% * 400 = 244$ station years for 10 Day AM. This is one of the main findings of this thesis.

5.2 – Synthetic Data Generation

Using the Generalised Extreme Value (GEV) distribution and the method that follows, synthetic data sets with varying inter-site-dependence within a region have been generated. These demonstrate the impact of varying inter-site correlation upon the confidence interval and later the effective number of sites in a pooling group.

5.3 – Method

The method described in this chapter is taken from Hosking and Wallis (1988).

The methodology outlines a procedure to allow the generation of simulated data for a pooling group, using primarily the cross-correlation of normalised Annual Maxima data between all of the sites in a specified region or group. Using the observed / generated matrix of site-to-site correlations (which are explained in more detail in this chapter), a multivariate normal random number generator is used to generate data samples of length t-years for each site; following some additional manipulation (explained in this chapter) this results in the synthetic annual maxima rainfall data.

What follows is a concise description of the methodology and the steps taken within the programme produced to carry out this analysis:

1. Read in the Annual Maxima for each site in the region (taken from real data, typically 40 years in length);
2. Read in site locations, in the form of grid references;
3. Calculate site-to-site separation, d_{ij} . Distance between sites 'i' and 'j';

4. Transform Annual Maxima data to a normal distribution using an empirical transformation;
 - a. Achieved using a pnorm routine in R. This is a predefined function, available from a library or toolkit within R.
5. Carry out pair-wise cross-correlation of normalised data between sites, which generates a correlation matrix;
6. Alternative to 5: Use a pre-defined ‘median correlation’ for the region and use equation: $\rho_{ij} = \exp(-\alpha d_{ij})$, to generate site-to-site correlation, using the site-separation matrix. The variable ‘ α ’ is optimised to give the required ρ_{med} ;
7. Carry-out a GEV distribution fit for each site (using observed data) and store parameters;
8. Using the pooled data, fit a regional (pooled) distribution and store GEV parameters; call this the regional GEV distribution or growth curve;
9. Using the ‘Multivariate Normal Random Number Generator’ (MVRNorm function in R) and the cross-correlation matrix, generate a data set of ‘n’ years (for example n = 40 years);
 - a. The MVRNorm function in R takes the cross-correlation matrix (which is also the covariance matrix because we’ve normalised the data) and generates as many new datasets as requested from the covariance matrix. Each new random dataset shares the same covariance/correlation between sites as seen in the original data.
10. Use the Regional GEV distribution to perform the Inverse Transformation of the data set, back to Annual Maxima Data – this is the artificial or synthetic data;
11. Combine or pool all of the data using the station-year pooling method, giving the number of sites multiplied by the length of the data sets, i.e. N sites * n years of synthetic data = Nn years of data with a known median regional correlation coefficient (ρ_{med});
12. Repeat step 9, 1,000 times to ensure a representative and realistic range of results. The number of samples (repetitions) was chosen based on consistency of results, where < 1% variation on summary statistics was achieved;
13. For each generated synthetic data set a GEV distribution is fitted and from this the rainfall return period (R.P.) estimates for the 50, 100, 1,000 and 10,000-year events are returned;

14. The range of results for each return period estimate is shown using a 'Box and Whisker' plot. The 'box' represents the median and the 95% confidence intervals; the whiskers show the most extreme values returned, these are called the 'outliers'; and,
15. For the sake of completeness, this test was repeated using a range of ρ_{med} values, as follows: $0 < \rho_{\text{med}} < 1$, $\rho_{\text{med}} =$ approximately 0, 0.1, 0.2, 0.3, 0.5, 0.75 and approximately 1;

5.4 - A summary of the methodology (with examples)

The aim of this section is to explain a part of the methodology, focusing on steps 2 to 6 of chapter 5.3. This part of the method is associated with correlation matrix. For normal use, this would be generated by performing a cross correlation of normalised annual maxima for the corresponding year for all of the sites in the region. This is the primary method used for this analysis. To explore and gain insight into the impact of inter-site dependence, it is also possible to generate a correlation matrix by using the equation:

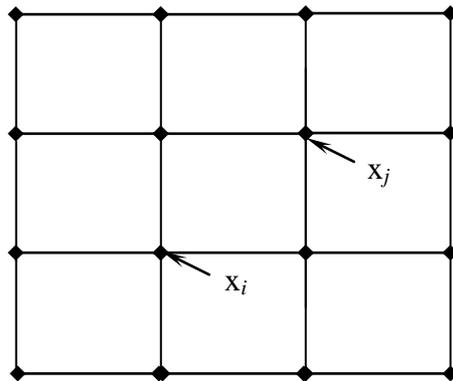
$$\rho_{ij} = \exp(-\alpha d_{ij}) \quad \text{Equation: 5.4.1}$$

Where:

- ρ_{ij} = the correlation (rho) between sites i and j .
- d_{ij} = the distance between sites i and j .

By varying the value of alpha (found by iterative analysis) the correlation between sites can be varied. Here, the median for the matrix has been optimised for academic interest, to generate a range of inter-site dependence values, ρ_{med} of $0 \leq \rho_{\text{med}} \leq 1$. This is explained in more detail within this chapter. What follows is an explanation of the methods used for the correlation matrix and also of Equation 5.4.1.

When looking at real data ρ_{med} is calculated straight from the correlation matrix for the regional pooling group. However, should we wish to generate a correlation matrix and the associated annual maxima data, this could be based upon manipulating the correlation between sites with real spatial locations (where 'i' and 'j' are specific to a region) or could be based on a uniform grid, as per figure 5.4.1 overleaf.



Where:
 Represents the location within the grid of site 'x'

Figure 5.4.1: Shows a uniform grid that could be used to demonstrate the location of sites within the synthetic region; sites x_i and x_j are shown for clarity, these could occupy any position within the grid or the region.

5.5 - The generation of a matrix of site-to-site separation, in kilometres:

Initial trials of this methodology were performed using 20 sites, each with 40 years of annual maxima data for the south west of England (SWE). The locations of these sites are shown in figure 5.5.1.

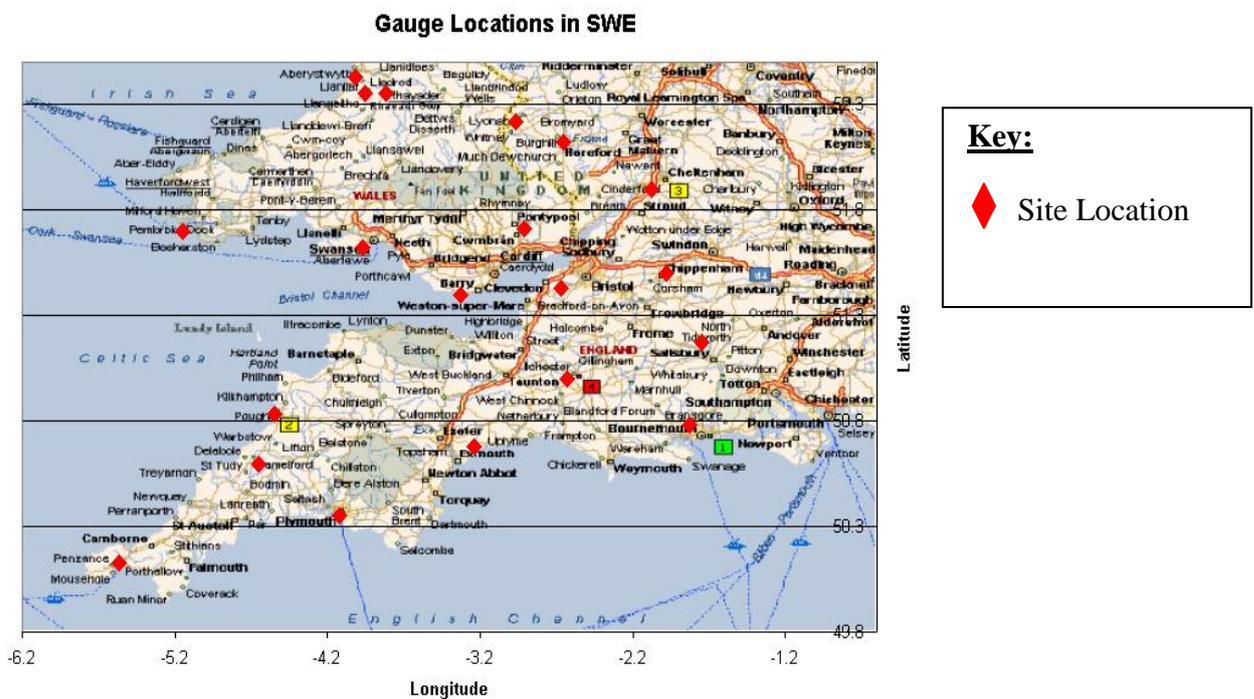


Figure 5.5.1: Locations of long-term daily rain gauges in the south west of England.

The grid reference for each of the sites was used to calculate the distance between sites and produce the following matrix:

	BOSCM	BUDEE	CHELT	CWMYS	DALEF	GOGER	HURNN	LNGAS	LYNEH	LYONS	PLYMO	PRESW	RHOOS	SIDMT	STANN	SWANS	TRAWS	TRENG	USKKK	YEOTN
BOSCM	0	200	84	194	244	210	44	70	40	144	190	123	113	118	214	162	201	294	101	65
BUDEE		0	209	177	106	182	192	148	195	190	61	196	107	94	27	97	175	106	151	136
CHELT			0	129	213	145	126	66	45	71	223	46	103	158	231	134	137	314	61	106
CWMYS				0	116	16	222	129	158	60	223	85	111	190	204	82	9	276	95	170
DALEF					0	112	254	175	221	161	167	179	131	176	129	82	109	178	155	191
GOGER						0	238	145	174	75	231	100	124	202	209	90	10	278	110	184
HURNN							0	93	81	179	169	160	126	100	200	176	228	275	128	62
LNGAS								0	49	90	157	77	46	93	167	93	135	251	35	47
LYNEH									0	105	197	83	94	127	213	139	165	295	69	71
LYONS										0	222	25	95	173	215	96	68	295	57	136
PLYMO											0	222	129	72	46	141	222	106	173	127
PRESW												0	94	166	221	108	93	302	50	124
RHOOS													0	81	129	51	114	212	45	65
SIDMT														0	100	117	192	176	117	56
STANN															0	124	201	83	174	149
SWANS																0	81	201	74	115
TRAWS																	0	272	101	174
TRENG																		0	257	230
USKKK																			0	80
YEOTN																				0

Figure 5.5.2: Matrix of site-to-site separation, d_{ij} , (km),

For the same sites, the annual maxima data was transformed to a normal distribution (step 4 of the method) and a pair-wise cross correlation performed to generate the correlation matrix, figure 5.5.3.

	BOSCM	BUDEE	CHELT	CWMYS	DALEF	GOGER	HURNN	LNGAS	LYNEH	LYONS	PLYMO	PRESW	RHOOS	SIDMT	STANN	SWANS	TRAWS	TRENG	USKKK	YEOTN
BOSCM	1	0.078	0.330	0.063	-0.031	-0.099	0.212	0.255	0.104	0.279	0.082	0.022	0.355	-0.035	0.074	0.246	0.119	0.148	0.022	0.305
BUDEE		1	0.144	0.243	-0.043	-0.001	0.013	0.147	0.123	0.276	0.421	0.092	0.036	0.332	0.599	0.018	-0.008	0.116	0.144	0.565
CHELT			1	0.015	0.157	0.070	-0.031	0.333	0.272	0.354	0.072	0.486	0.331	-0.007	0.280	0.170	0.199	0.041	0.382	0.268
CWMYS				1	-0.068	0.303	-0.007	-0.193	0.049	0.252	0.065	0.096	-0.173	0.007	0.175	0.021	0.749	0.162	-0.161	0.288
DALEF					1	0.199	0.135	-0.037	0.316	0.335	0.093	0.433	0.033	0.258	-0.124	0.149	0.182	0.146	0.263	-0.046
GOGER						1	0.044	-0.176	0.021	0.058	0.076	-0.155	-0.302	0.248	-0.070	0.143	0.432	0.030	-0.035	-0.115
HURNN							1	0.170	0.159	-0.201	0.027	-0.194	0.076	0.091	-0.259	-0.013	0.121	0.201	-0.071	0.133
LNGAS								1	0.594	0.054	0.117	0.273	0.575	0.222	0.267	0.193	-0.229	0.207	0.199	0.216
LYNEH									1	0.146	0.114	0.290	0.302	0.384	0.028	0.097	0.153	0.274	0.370	0.325
LYONS										1	0.134	0.406	0.289	0.229	0.267	0.206	0.295	0.151	0.159	0.277
PLYMO											1	0.007	0.036	0.211	0.234	-0.076	0.058	0.309	0.131	0.156
PRESW												1	0.320	-0.113	0.205	0.248	0.160	0.038	0.356	0.082
RHOOS													1	0.178	0.204	0.249	-0.105	0.111	0.188	0.222
SIDMT														1	0.142	0.029	0.152	0.177	0.066	0.273
STANN															1	0.141	-0.048	0.133	0.205	0.391
SWANS																1	0.047	0.029	0.324	-0.005
TRAWS																	1	0.192	-0.043	0.242
TRENG																		1	-0.028	0.186
USKKK																			1	0.091
YEOTN																				1

Figure 5.5.3: Matrix of site-to-site cross-correlation, ρ_{ij}

From figure 5.5.3, it is possible to define ρ_{med} , which is the median correlation for the region. This value defines the level of dependence for the region.

Alternatively, for research and validation purposes it is possible to optimise α (in equation 5.4.1 - $[\rho_{ij} = \exp(-\alpha d_{ij})]$) to give a ρ_{med} of $0 \leq \rho_{med} \leq 1$. For example if $\rho_{med} = 0$ were chosen, this would indicate that there is zero dependence between sites and for other values of ρ_{med} up to $\rho_{med} = 1$ for example, that greater levels of dependence exist within the region, and the corresponding effects on the accuracy of the station

year approach to pooling regional data; this has been shown in this chapter and chapter 5.7.1.

However, the generated Multivariate Normal data set (synthetic data), has been based on the at site cross-correlations for the south west of England. Because the ‘at site’ time series has been ‘normalised’ (step 4, Chapter 5.3), meaning the annual maximum values at each site have been transformed to a normal distribution, the correlation matrix for all of the sites is also the covariance matrix of the variables. The covariance matrix is used by the Multivariate Random Normal routine (available within the ‘R’ environment, called ‘mvrnorm’) to produce one or more samples from the multivariate normal distribution (this is explained in step 15 of the method). This study produced record lengths ranging from 10 to 40 years for each rain gauge site, of which there were 20 in the south west of England.

The generated data consist of the cumulative density function (c.d.f.) for each site in the region. The method used to perform the inverse-transformation of the c.d.f. for each site is to use this data as the Gringorten plotting positions (see Chapter 4.2.2). Using the GEV parameters for the pooling group in the chosen region, the inverse transformation was carried out; this produced the synthetic annual maxima data sets for the region.

By having 1,000 repetitions of this process for each site, it was found that this gave an adequate representation of the range of possible sample variations that could be encountered. This figure was decided upon when repeating the same experiment; the variability of the results was negligible, typically less than 1% for any of the stated values; for example, the mean, median and confidence limits, but not including the most extreme outliers.

The correlation vs. distance (spatial separation) relationships for each region can be seen in Appendix 1.

5.5.1 – Results and discussion

The graphs that follow in Figures 5.5.4 to 5.5.6 show a summary comparison of L-Moments and MLE for the fitting of distributions to the synthetic or artificial multi-site pooling groups. To aid in the comprehension of these results graphs have been produced showing $\rho_{\text{med}} = 0.1, 0.3, 0.5$ and 0.75 , with the estimates produced by the two methods for the 100, 1,000 and 10,000-year event.

The following graphs show the L-moment estimates in blue and MLE in yellow. For those readers viewing a black and white copy of this document, the pairings consist of L-Moment estimates on the left and MLE on the right. The box-plots show the 95% confidence interval and the horizontal black line in the box shows the median of the estimates. The pink bullet-point shows the corresponding mean of the estimates.

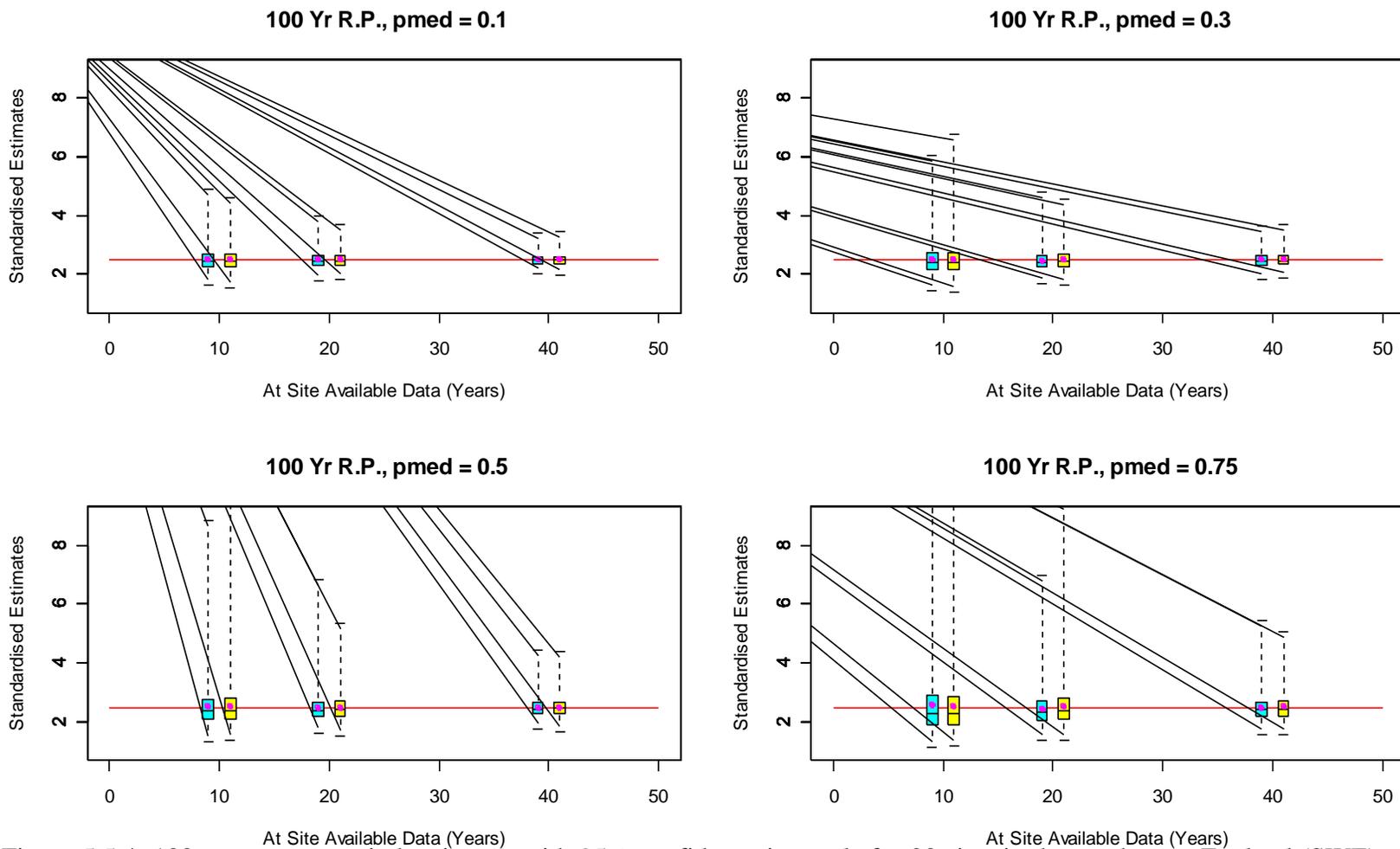


Figure 5.5.4: 100-year return period estimates with 95% confidence intervals for 20 sites in the south west England (SWE), each site having: 10, 20 or 40 Yrs of 1 day annual maxima data.

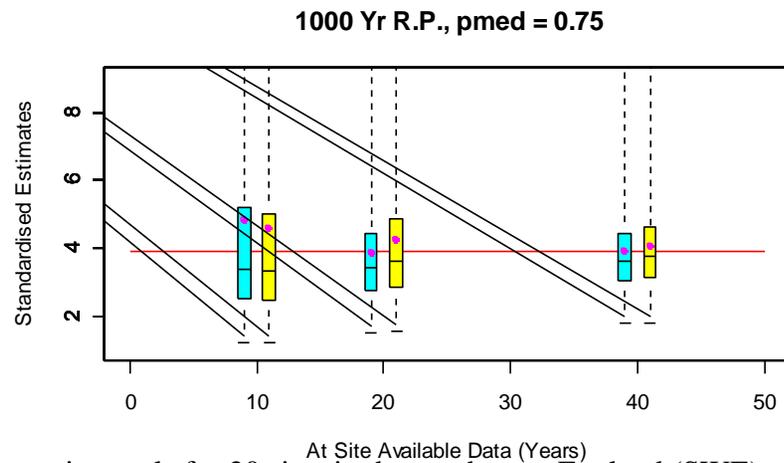
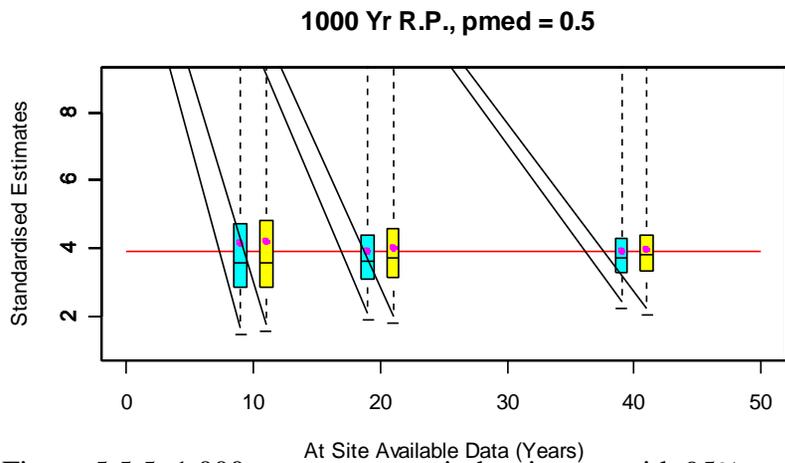
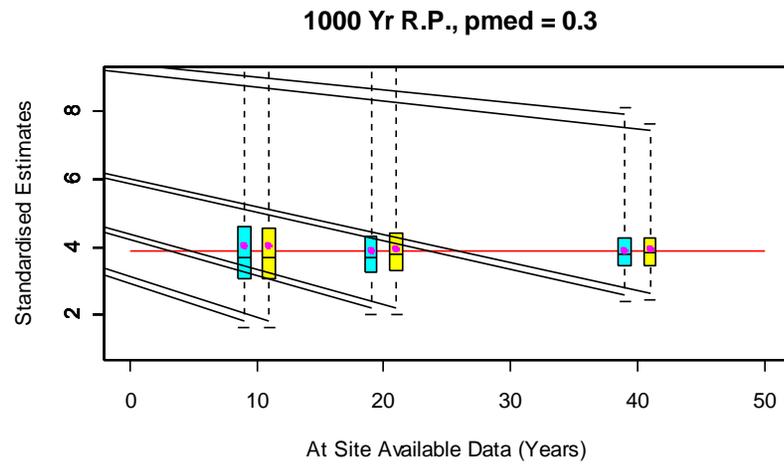
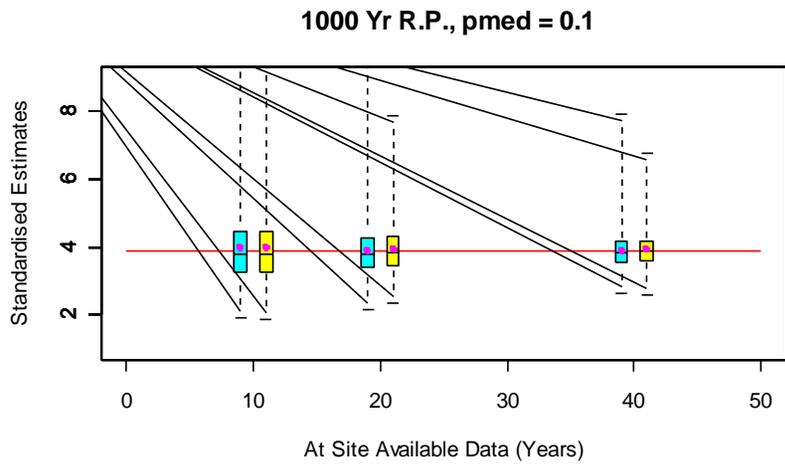


Figure 5.5.5: 1,000-year return period estimates with 95% confidence intervals for 20 sites in the south west England (SWE), each site having: 10, 20 or 40 Yrs of 1 day annual maxima data.

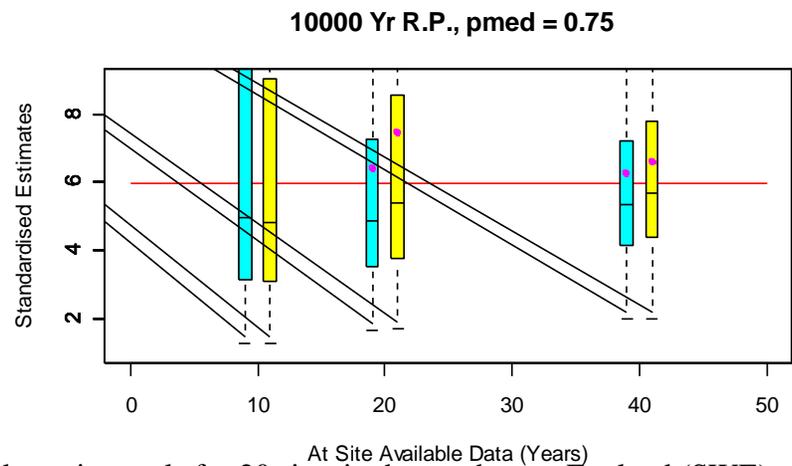
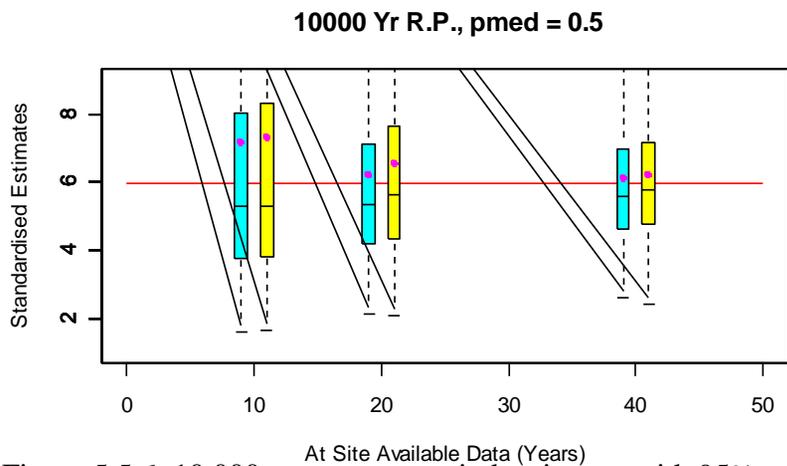
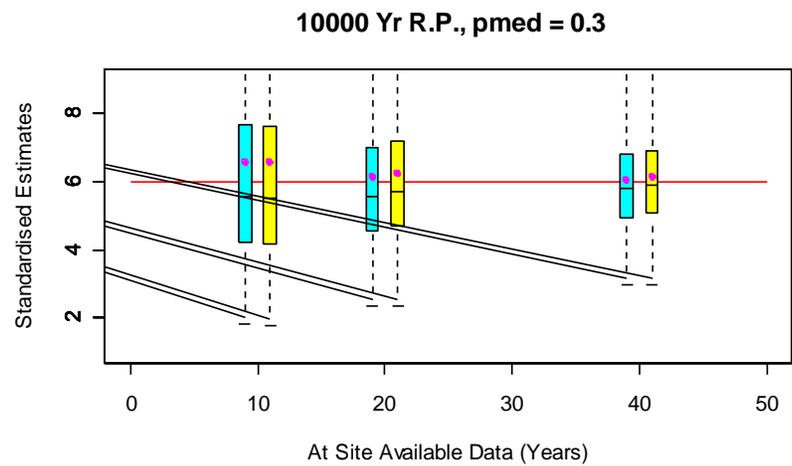
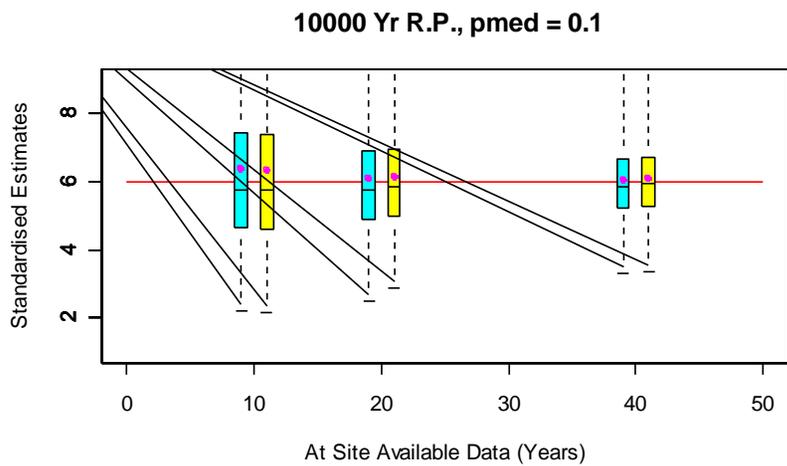


Figure 5.5.6: 10,000-year return period estimates with 95% confidence intervals for 20 sites in the south west England (SWE), each site having: 10, 20 or 40 Yrs of 1 day annual maxima data.

This page is intentionally blank.

Looking at Figures 5.5.4, 5.5.5 and 5.5.6, it is possible to start to visualise the distribution of the estimates. However, to simplify this further, two examples are shown below.

These are taken from Figure 5.5.6 and show:

1. A dashed line representing the true value of 5.97, plotted on the right-hand limit of this point – the histogram ‘bins’ increase with intervals of 1.
2. The number of estimates above or below this ‘True value’; and,
3. For this example, the distribution of estimates for each method.

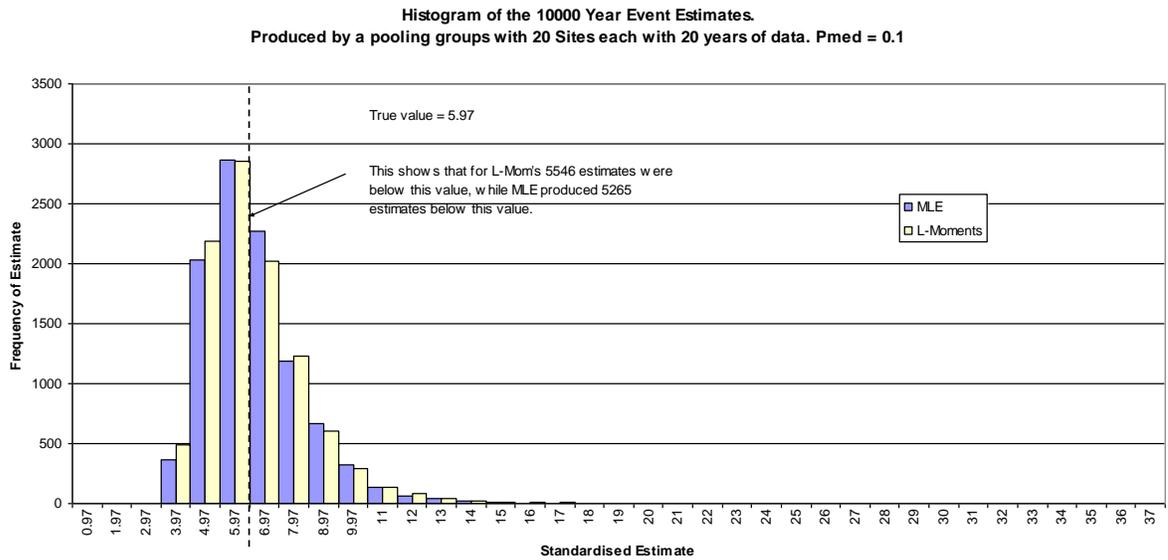


Figure 5.5.7: Histogram of Estimates produced by the synthetic data set where $\rho_{med} = 0.1$

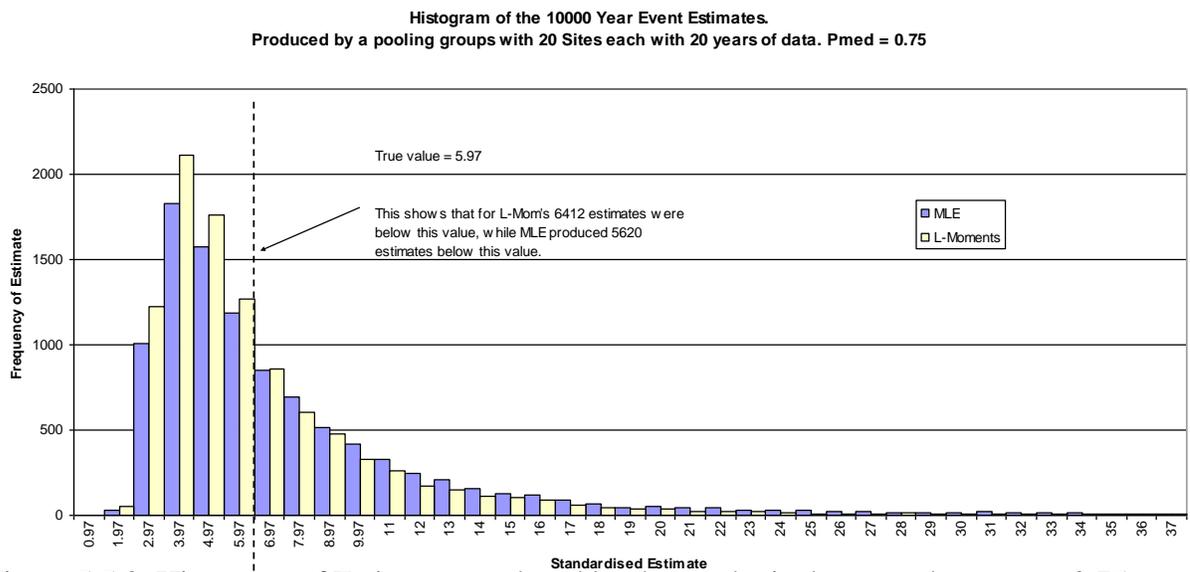


Figure 5.5.8: Histogram of Estimates produced by the synthetic data set where $\rho_{med} = 0.75$

So far, the figures in chapter 5.5 have shown that an increase in ρ_{med} results in decreased confidence in the quantile estimate as the 95% confidence intervals (upper and lower limits) move further away from the true (known) value. The decrease in confidence takes the form of increased standard deviation and a more skewed distribution of the estimates. The median values have a more pronounced negative bias (under-estimating) as the Inter-site-dependence increases (it is not known why); however, the mean has a more pronounced positive bias, increased by the dramatically over-estimating outliers that are also a consequence of the increased dependency. The 10,000-year return period estimates have been chosen to demonstrate these characteristics, purely because they are more pronounced and clearly more visible, but the characteristics remain the same at lower return periods. The median values give a more representative indication of the true value, where as the mean is more affected by bias of the estimates, typically producing an over estimation, due to the very large outliers.

Examining 40 years (1960 to 2000) of 1, 2, 5 and 10 Day Annual Maxima has shown that for multi-site (regional) pooling groups the regional ρ_{med} values are:

Regions	1 Day	2 Day	5 Day	10 Day
Southern Scotland	0.22	0.28	0.39	0.44
Northern Scotland	0.17	0.2	0.3	0.38
East Scotland	0.18	0.24	0.3	0.34
North East England	0.31	0.44	0.44	0.44
North West England	0.18	0.17	0.23	0.2
Central Eastern England	0.2	0.2	0.25	0.32
South West England	0.15	0.18	0.19	0.28
South East England	0.26	0.25	0.35	0.45

Table 5.5.9: Table of the observed regional ρ_{med} values for 1, 2, 5 and 10 day annual maxima.

In general, Table 5.5.9 shows that there is increased inter-site correlation as the duration increases from the 1 day to the 10 day duration annual maximum events. The degree of correlation corresponds to the chance that the maxima come from the same storm event.

N.B. Graphs of the distance versus correlation relationship can be seen in Appendix 1, these also show information relating to the Netmax concept (defined in chapter 5.6).

An alternative to comparing ρ_{med} for each region and varying duration is to compare ρ for a fixed separation, say 100km, giving ρ_{100} . The justification for this is that ρ_{med} will vary according to the size of the region and therefore may not give a clear comparison of inter-site-dependence in the various regions.

Regions	1 Day	2 Day	5 Day	10 Day
Southern Scotland	0.1	0.2	0.3	0.35
Northern Scotland	0.3	0.3	0.38	0.45
East Scotland	0.15	0.17	0.18	0.3
North East England	0.3	0.4	0.4	0.42
North West England	0.17	0.15	0.2	0.2
Central Eastern England	0.25	0.25	0.3	0.35
South West England	0.2	0.19	0.2	0.35
South East England	0.2	0.2	0.23	0.38

Table 5.5.10: Table of the regional ρ_{100} (approximate ρ - spatial correlation - at 100km separation) values for 1, 2, 5 and 10 day annual maxima.

With these results in mind the importance of confidence limits becomes more apparent. Confidence limits may be unpopular with designers and engineers who seek one definitive answer, but it is irresponsible and misleading to give one value, when a range of uncertainty exists. It should also be made clear that the range of uncertainty reduces with increased record lengths, but also that questions need to be asked about the homogeneity of pooled data sets.

5.6 - Netmax concept and $\ln(Ne)$, the effective number of sites

In this chapter, the aim is to show a working relationship between the Network Maximum values (Netmax) for a region and ρ , the correlation between sites within the regional network of rain gauges, which in turn is related to the effective number of rain gauges within a spatially defined region. This work follows on from the multi-variate, synthetic rainfall data generation.

Network Maximum values (Netmax)

The Network Maximum value is defined as the largest value for a given year across the network or region. For each annual maximum series in the pre-defined region or network, there is a maximum value for each year across the network.

When a distribution is fitted to these values, the resultant growth curve is called the Netmax growth curve.

The effective number of sites

The effective number of rain gauges is of interest because when sites are pooled together it is known that the accuracy or the improvement gained by pooling more and more sites is not proportional. For example, a regional pooling group of 20 sites with 20 years of data is not an equal replacement for one site within the region with a continuous record length of 400 years, the question being asked in this chapter is: what is the equivalent record length, or the effective number of sites in the pooling group?

Within this chapter, it has been shown (using synthetic data) that when $\rho_{\text{med}} = 1$ the median of the Network Maximum is equal to the true plotting position for the regional growth curve. When $\rho_{\text{med}} = 1$, there is total dependency between sites, meaning that the effective number of sites in the region is just 1. This can be further explained using the following example: when $\rho_{\text{med}} = 1$, it implies that just one rainfall event is being recorded; it is simply being recorded at different locations with varying magnitude. From

this, we realise we are seeing a range of estimates that fit one distribution with a median and mean equal to the true plotting position.

Conversely, this means that if $\rho_{\text{med}} = 0$, then there is no spatial dependency between sites and that no two rain-gauges record the same rainfall event, suggesting that the rainfall is highly localised or that the gauges are separated by a relatively large distance, compared to the size of the weather system; ρ_{med} therefore has an impact on the effective number of sites within a region and this has a significant impact on the amount of available data.

In reality, it is virtually impossible to achieve either $\rho_{\text{med}} = 0$ or $\rho_{\text{med}} = 1$ as random events, will occur simultaneously within a large enough sample; there will also be other factors, such as topography and separation between gauges which will prevent a ρ_{med} of zero or one. It is possible, however, for ρ to be negative, meaning that the opposite occurs at one site compared to another. An example of this is when it is raining in one location it frequently appears to be dry at the other location being examined. ρ and ρ_{med} must therefore lie between these two values, i.e. $-1 \leq \rho \leq 1$.

The examples shown overleaf for $\rho_{\text{med}} = 0$ and $\rho_{\text{med}} = 1$, are actually very close approximations, for example $\rho_{\text{med}} = 0.00072$ or $\rho_{\text{med}} = 0.999849$. The reason behind this are explained in the paragraph above.

With this in mind and the desire to find a relationship between ρ_{med} or ρ_{100} and the effective number of sites within a region, an empirical relationship between ρ_{ij} and the spatial separation d_{ij} , between gauges i and j was used, see equation 5.4.1. This allowed a correlation matrix to be generated with a ρ_{med} in mind, i.e.: 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75 and 1.

This method was proposed as a means for gaining insight into the tail of a distribution, that part which contains the rarest events. This technique was pioneered by Dales and Reed (1989). The assumption is that if the largest values are extracted from the annual maxima network for each year and plotted, there will be a constant separation between the Netmax Growth Curve and the Regional Growth Curve. The constant, N_e , represents the effective number of sites in the region. This assumption has also been tested and shown to be incorrect by this thesis and also by work carried out by the Cooperative Research Centre for Catchment Hydrology (1997).

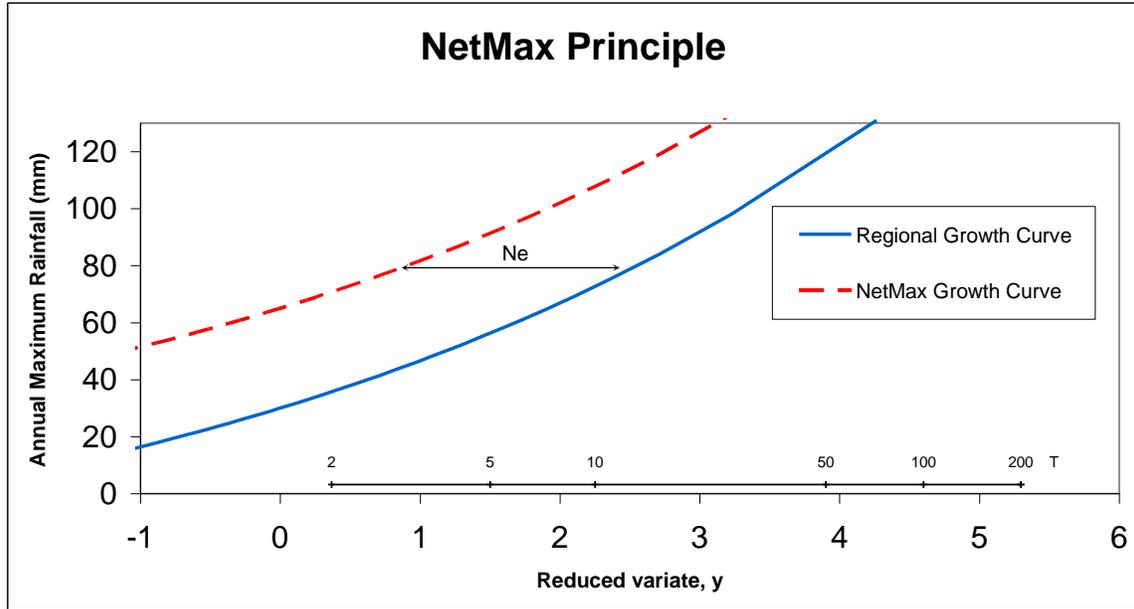


Figure 5.6.1 – Illustration of the Netmax principle, as explained by Dales and Reed.

Figure 5.6.1 shows that the separation between the Netmax growth curve and the regional growth curve to be constant, as explained by Dales and Reed (1989). However, this thesis finds that this separation is not a constant, but that it increases with return period. This has been demonstrated using synthetic data (Figure 5.6.3) and has been seen using observed data; however, the separation has also been shown to decrease when compared with an increasing return period, converging and then increasing again.

Using synthetic data, figures 5.6.2 and 5.6.3 show the effect of inter-site-correlation on the separation between the Netmax values and the regional growth curve from which they originate.

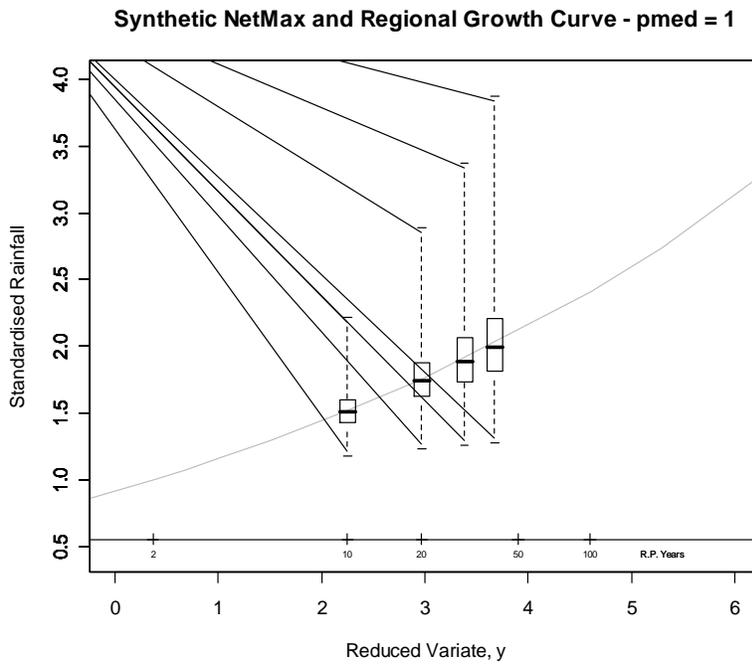


Figure 5.6.2 – Network Maximum values, illustration of total dependence between sites.

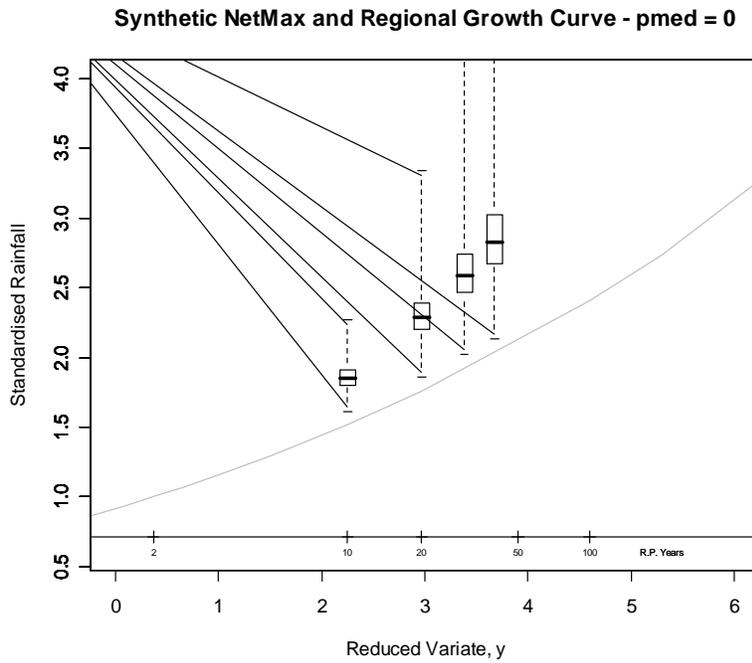


Figure 5.6.3 – Network Maximum values, illustration of zero spatial dependence between sites.

5.7 – ln(Ne) – Effective Number of sites for Observed 1, 2, 5 and 10 Day Annual Maxima Data. Regional Correlation / Dependence Data for regions within Great Britain

An analysis of the inter-site-dependence characteristics for each region in Great Britain has been carried out. The analysis was performed using pooled 1, 2, 5 and 10 day durations of annual maxima, for each region.

The regional network maximum is extracted from the data set and plotted along side the regional growth curve. The separation on the x-axis, using the reduced variate ‘y’, between corresponding rainfall depths of these two growth curves is calculated and using equation 5.7, a value for Ne, the effective number of sites, this is then calculated:

$$\frac{\ln(Ne)}{\ln(N)} = \text{Separation between corresponding data on the x axis.} \quad \text{Equation 5.7}$$

Where N is the total number of sites in the pooling group.

This study has shown with synthetic data and observed regional data, that Ne is not a constant. This is significant result because Dales and Reed, who introduced this technique, assume that $\ln(Ne)/\ln(N)$ is a constant; this assumption is used by the Flood Estimation Handbook. A complete set of results for each region within Great Britain can be found in Appendix 1.

5.7.1 – Results for the Effective Number of sites for regions within Great Britain

Table 5.5.9 shows the correlation between sites, expressed in the form of the median regional correlation descriptor. Appendix 1 shows the effective number of sites in each region. Table 5.7.1.1 summarises the effective number of sites in each region by taking the mean of $\ln(Ne)/\ln(N)$ at F10, F20 and F30. The notation: F10, F20, F30 and F40 have been used to represent the Gringorten plotting positions for the 10, 20, 30 and 40 year return periods.

Regions	1 Day	2 Day	5 Day	10 Day
Southern Scotland	0.75	0.76	0.70	0.63
Northern Scotland	0.93	0.85	0.80	0.71
East Scotland	0.83	0.81	0.66	0.61
North East England	0.87	0.71	0.72	0.70
North West England	0.90	0.91	0.86	0.88
Central Eastern England	0.92	0.91	0.85	0.79
South West England	0.84	0.90	0.91	0.75
South East England	0.74	0.71	0.71	0.61

Table 5.7.1.1: Table of the regional $\ln(N_e)/\ln(N)$ values for 1, 2, 5 and 10 day annual maxima.

The results in Table 5.7.1.1 can be interpreted as a percentage of the total number of sites, for example the 1 Day $\ln(N_e)/\ln(N)$ result for Southern Scotland is 0.75, this means that the effective number of sites = 75% or 18.75 sites from the regional total of 25 sites; effectively reducing the station year total from 1,000 station years to 750 station years.

Plotting ρ_{med} and ρ_{100} (tables: 5.5.9 and 5.5.10) against the effective number of sites in each region, then fitting a linear regression to each, an equation (equation 5.7.1) relating ρ_{med} to the effective number of sites in the pooling group was created. Where ρ_{med} is the median of the correlation values between the sites in the region and ρ_{100} is correlation value which corresponds to a 100km separation between sites within the regional being assessed.

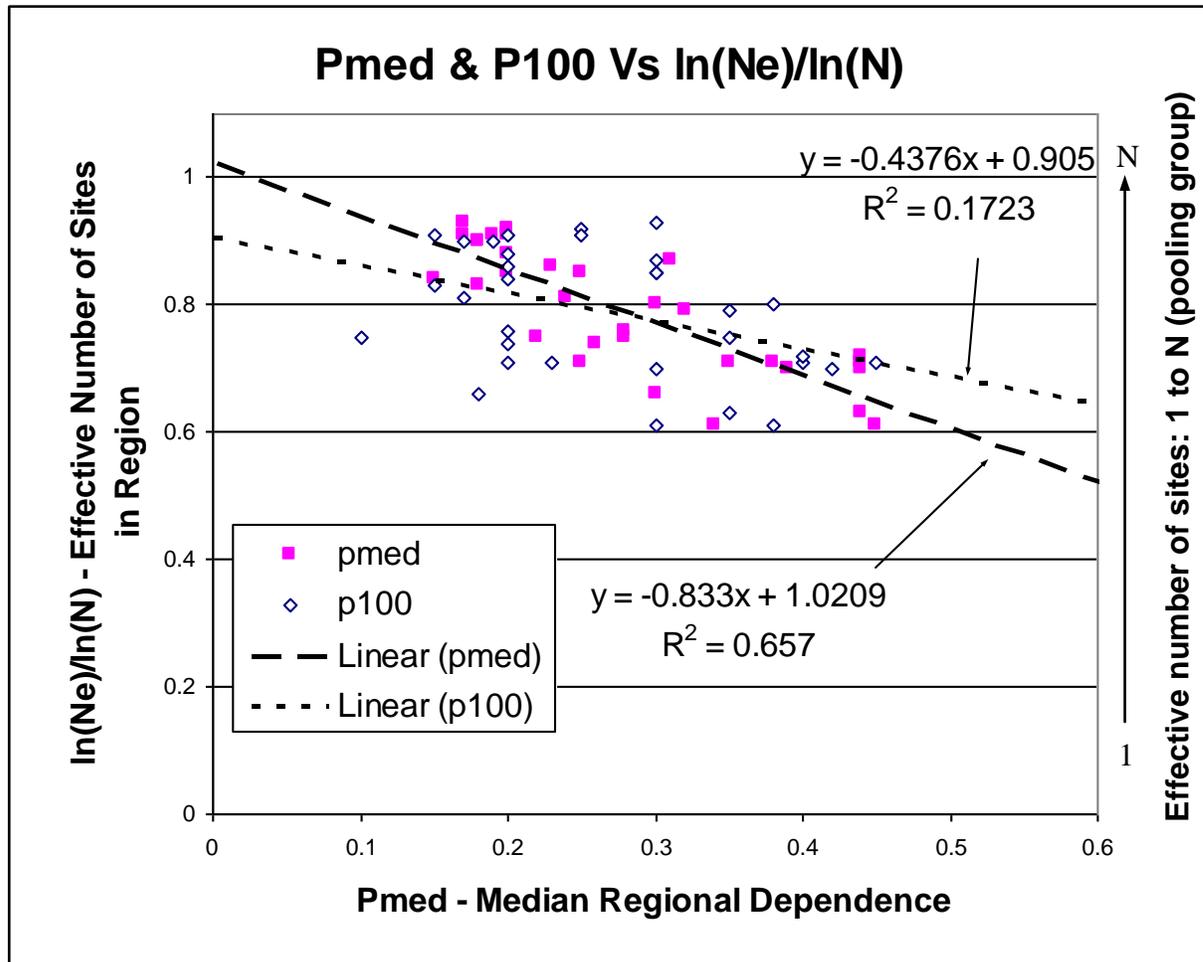


Figure 5.7.1.2: Plot of ρ_{med} and ρ_{100} vs. $\ln(Ne)/\ln(N)$ values for each region, for durations of : 1, 2, 5 and 10 day annual maxima.

From previous explanations on the effect of inter-site-dependence (see Chapter 5.6) it was explained that as the inter-site-dependence reduced to zero, the effective number of sites increased to 100% of the pooling group. It was also explained that as the inter-site-dependence (ρ) approaches 1, the effective number of sites reduces to 1.

Inspection of the R^2 values and the x and y intercepts for the trend lines in figure 5.7.1.2, shows that the ρ_{med} values best describe the inter-site correlation for the region when compared with the $\ln(Ne)/\ln(N)$ values. The x and y intercepts are also not far away from the expected values (y-intercepts: 0.9 for ρ_{100} and 1.02 for ρ_{med}) they could be forced, but figure 5.7.1.2 shows the fit with the largest R^2 values.

Figure 5.7.1.2 clearly shows that there is a relationship between the inter-site correlation and the effective number of sites in the region. In fact, using the regression equation, the relationship can be approximated as:

$$\text{Ln}(\text{Ne})/\text{Ln}(\text{N}) = -0.833 \rho_{\text{med}} + 1.0209 \quad \text{Equation 5.7.1}$$

If the observed ρ_{med} values from Table 5.7.9 are entered into Equation 5.7.1, then it is possible to compare the observed or calculated Ne values with those estimated by Equation 5.7.1; this shows a typical range of errors of approximately +/- 15%, the exceptions are highlighted, see Table 5.7.1.3:

Region	1 Day		2 Day		5 Day		10 Day	
	Estimated Ne/N	% Error	Estimated Ne/N	% Error	Estimated Ne/N	% Error	Estimated Ne/N	% Error
SS	0.83764	12%	0.78766	4%	0.69603	-1%	0.65438	4%
NS	0.87929	-5%	0.8543	1%	0.771	-4%	0.70436	-1%
ES	0.87096	5%	0.82098	1%	0.771	17%	0.73768	21%
NEE	0.76267	-12%	0.65438	-8%	0.65438	-9%	0.65438	-7%
NEW	0.87096	-3%	0.87929	-3%	0.82931	-4%	0.8543	-3%
CEE	0.8543	-7%	0.8543	-6%	0.81265	-4%	0.75434	-5%
SWE	0.89595	7%	0.87096	-3%	0.86263	-5%	0.78766	5%
SEE	0.80432	9%	0.81265	14%	0.72935	3%	0.64605	6%

Table 5.7.1.3: Estimates and Range of percentage errors from using Equation 5.7.1 with observed correlation data from Table 5.7.9.

5.8 – Homogeneity Testing

5.8.1 – Introduction

This section (Chapter 5.8) aims to introduce an existing homogeneity testing technique and to demonstrate the potential of a new technique, which is dependent upon the use of Maximum Likelihood Estimation (MLE) and the Likelihood ratio test.

The hypothesis of homogeneity is that the at-site frequency distributions are the same except for a site-specific scale factor [Hosking and Wallis, 1997].

5.8.2 – Existing techniques

If a region is homogenous in the sense that the data for each site within the region represents a random realisation of the same underlying physical process, then the regional L-Moments can be used to fit a probability distribution [Hosking 1990; Hosking and Wallis 1990; Wallis 1989]. Hosking and Wallis [1991] constructed a method to evaluate goodness of fit. This measure is based upon the difference between L-Kurtosis of the fitted distribution and the regional average L-Kurtosis of the sample data. Assessment of goodness of fit is based on L-Kurtosis, the fourth L-moment, because the first three L-Moments are used to estimate the three parameters of the distribution.

Homogeneity is a basic requirement when pooling data and is often assumed in regional frequency analysis. This chapter sets out to test for homogeneity within the predefined regions, which have been used up to this point. Homogeneity testing using L-Moments is achieved by making comparisons of the L-Moment ratios. The graphical comparison checks for similarity / homogeneity in the distribution properties, typically L-CV (coefficient of variation). The measure of similarity is often achieved through a ‘goodness’ of fit measure, such the R^2 value obtained following a linear regression.

5.8.3 – Possible future technique

The homogeneity test that will be demonstrated here is the Likelihood Ratio Test (LRT). This is testing at a significance level, for example 5%, the probability of the sample (at

site distribution) originating from the population distribution (pooling group). The Null Hypothesis is that the pooling group is homogeneous. A significant result therefore is where the test statistic returns a probability of less than a chosen significance level, for example 5%. Then it is a significant result and the test has indicated a high probability of Heterogeneity.

It is important to note that the 5% significance level refers to the associated error, or probability of the result being incorrect. To explain this further, a p-value of 5% means that 5 times out of 100 an incorrect result will be returned, meaning the data will show that the region is heterogeneous when it is in fact homogeneous.

5.8.4 – Method description

This chapter (5.8.4) contains an example and a more complete description of the method used to carry out homogeneity testing using MLE and the Likelihood Ratio Test.

In order to carry out the likelihood ratio test, the log likelihood for two different models are required:

Model 1 - All sites are assumed to be homogeneous, so that the same GEV parameters u , a and k fit all. In order to find these values, the data are pooled into a single sample, and the GEV parameters are calculated. The log-likelihood value returned is the target value for this model, which will now be referred to as LL1.

Model 2 - All the sites are allowed to have their own values of (u,a,k) . This time the GEV parameters are calculated for each site individually. Log-likelihood values are calculated for each site. The sum of these values is calculated and referred to as LL2.

N.B. LL2 must be greater than or equal to LL1. The test statistic is $2(LL2-LL1)$, (i.e. twice the difference $LL2-LL1$). This should result in a positive number. Under the null hypothesis, which is that the pooling group is homogenous, then this should come from a Chi-Square Test Statistic distribution on ν degrees of freedom, where ν is (total number of parameters in model 2 - total number of parameters in model 1).

So, a significant result, i.e. a test statistic which is significantly large when you look at the Chi Square Test Statistic tables for ν degrees of freedom, is evidence against the null hypothesis, in favour of the alternative, i.e. non-homogeneity.

5.8.4.1 - Example

Below is an example of the output from the routine written in R to test for homogeneity. This test has been applied to all of the regions used up to this point; a complete set of results can be seen in Appendix 2.

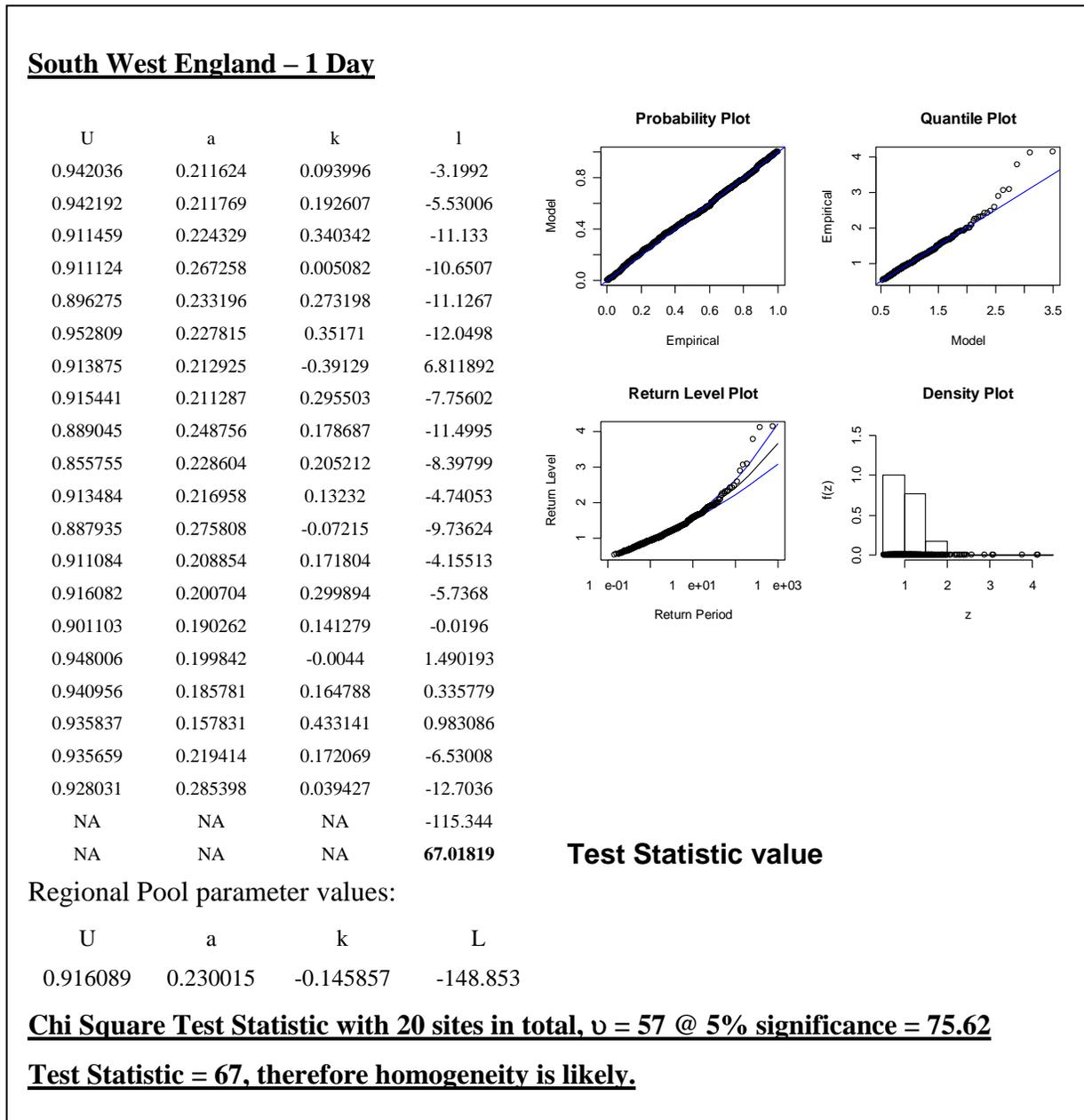


Figure 5.8.2.1 – Homogeneity test for SWE using 1 Day Annual Maxima

5.8.5 – Discussion

The main motivation for the trial of this method was the desire that this technique might enable the formation of very large pooling groups (regions). The inspiration was the method of catchment selection used by the FEH; where catchment descriptors are used to select comparable sites for the purpose of data augmentation, not being restricted to neighbouring catchments.

The aim for the homogeneity test (described in this chapter) was to select sites or perhaps more accurately to reject sites based upon a comparison of their fitted distribution with the existing yet expanding pooling group. One example of why this method might be more powerful than a straight forward radial expansion of the pooling group (focused on the site of interest) is: should the focal site be located close to but clearly on one side of the Pennines for example, then a radial expansion might not be suitable, given the influence of the Pennines on the rainfall characteristics of the available rain gauges. It is also clearly preferable to choose sites that are further away yet still statistically homogeneous as this would lead to reduced inter-site-dependence (see Chapters 5.6 and 5.7 for information on inter-site correlation) and therefore the advantage of the additional site is two fold:

1. Increasing the number of station years in the pooling group; and,
2. Reducing the inter-site-dependence.

Chapter 1.1 provided the background and introduced the problems associated with extreme value rainfall estimates; primarily this is a lack of long duration rainfall data – especially at the point of interest. This chapter has demonstrated a method using statistical analysis to form significantly larger data sets.

Although some results have been included, it was decided not to pursue this technique of site selection for pooling groups. The reason is that this statistical test uses a significance level in the hypothesis testing each time a site is added to the pooling group. Therefore, each time a site is added there is, for example, a 5% chance of error. What has not been investigated is the compounding effect of repeating this test and the associated errors. It has been included here though, because it is hoped that someone with sufficient statistical expertise might be able to develop this technique further.

A summary of results from using an experimental method of focused pooling group expansion, based upon homogeneity testing of 179 rain gauges in Great Britain, follows. Having written a programme that will expand upon a single site, the focal site, by adding sites according to (or in the order of) their proximity to the focal site and testing for homogeneity each time a site is added until a pooling group which is deemed to be non-homogeneous is produced. The following summarised results were obtained for each site in Great Britain:

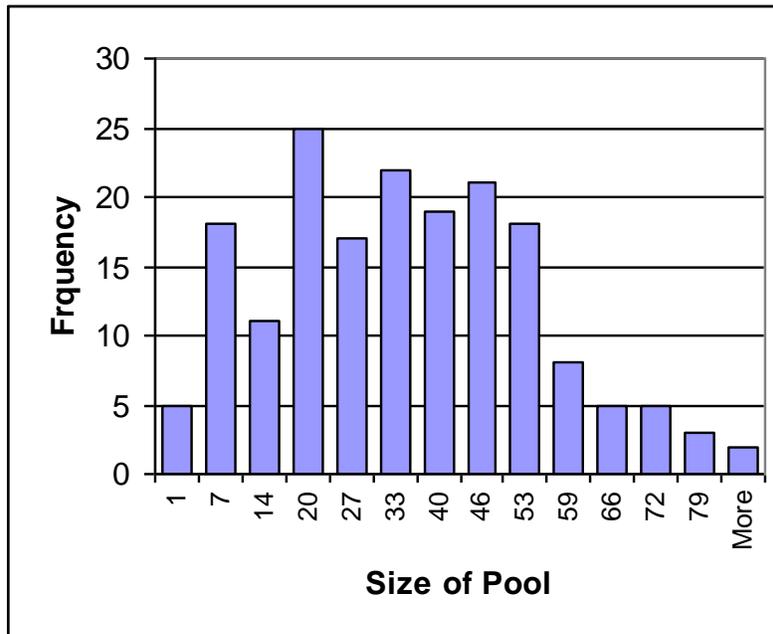


Figure 5.8.2.2 - Histogram of number of sites per pool.

Starting from any one of 179 rain gauges in Great Britain (selected by the user, randomly if desired), additional sites were added to the target site. This Pool (x-axis) was expanded upon and until a non-homogeneous Pool is produced. Figure 5.8.2.2 shows the number of rain gauges in the homogeneous Pool and the number of Pooling groups (y-axis) for the corresponding Pooling group size.

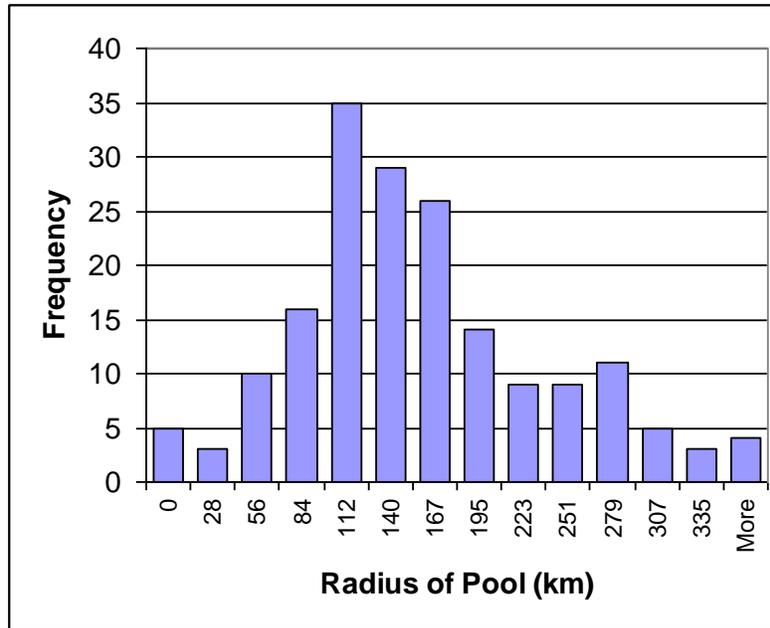


Figure 5.8.2.3 - Histogram of maximum radial separation of sites per pool, up to but not including the site that caused heterogeneity.

Taking the average number of sites per pooling group, 31.4, and assuming that each site has 40 years of data (the average for the sites used in this thesis), then using this method the average pooling group contains 1256 station years of rainfall data.

The largest pooling group contained 85 sites which produced 3400 station years of rainfall data (figure 5.8.2.4).

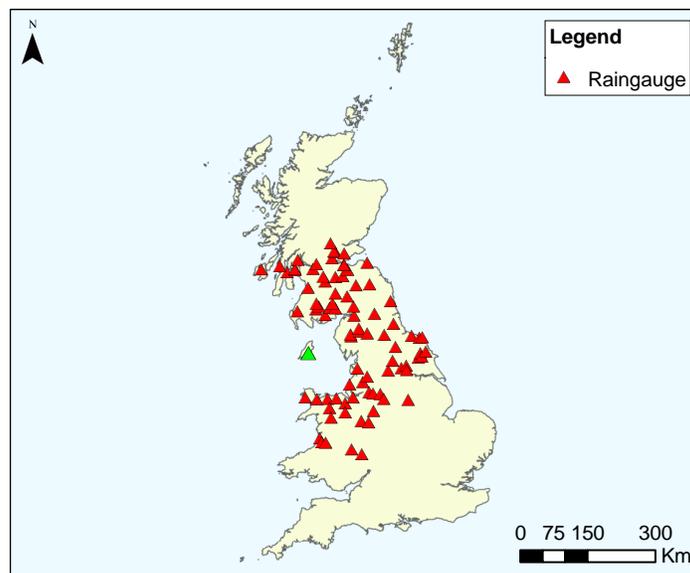


Figure 5.8.2.4 - Shows the 85 sites which produced the largest homogeneous pool.

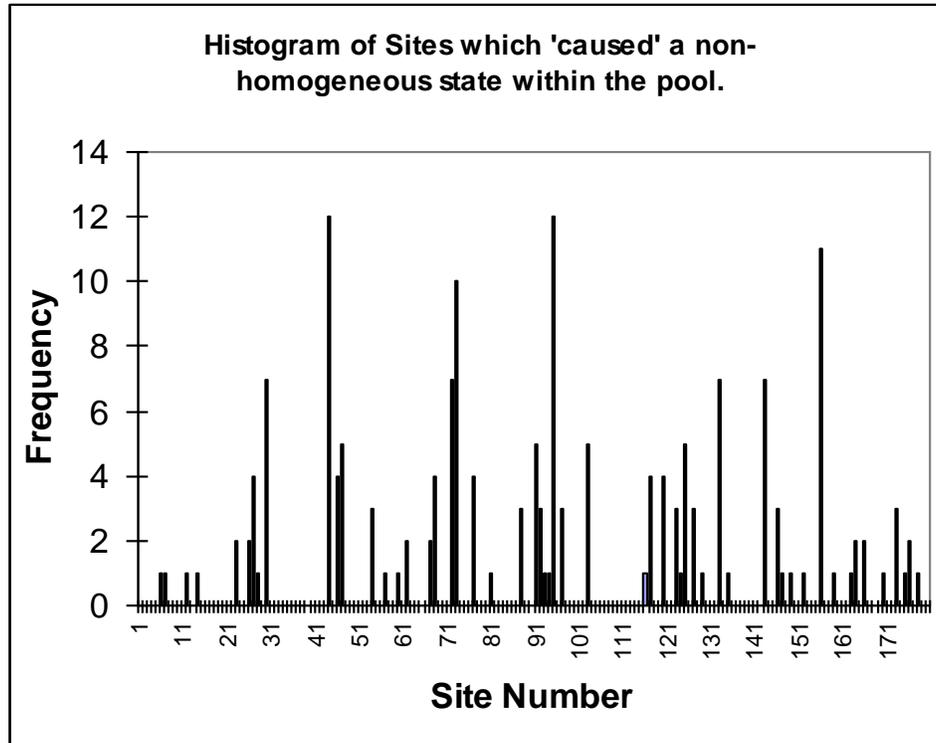


Figure 5.8.2.5 - Shows the 54 sites which produced / caused a non-homogeneous pool.

If we choose those sites with a frequency >6, we are left with the following sites:

Site	Frequency	Site Name	Region
30	7	Lyonshall	SWE
72	7	Llanuwchllyn	NEW
133	7	Balmoral	ES
143	7	Frandy	ES
73	10	Loggerheads	NEW
156	11	Benmore Younger Botanic Garden	SS
44	12	Elmdon	CEE
95	12	Cockle Par	NEE

Figure 5.8.2.5 – Table of sites shown to have repetitively caused heterogeneity.

It was hypothesised that the location of the sites in Figure 5.8.2.5 could be directly attributable to, for example:

1. Geographical/ topographical uniqueness
2. Anomaly with equipment
3. Inappropriate human interference
4. Proximity to coast

However, only Benmore Botanic Gardens appears close to the coast, the others can be classed as inland.

Of the 179 sites used as the focal site, 125 of those sites were not responsible for creating a non-homogeneous pool.

Taking those sites as a 'region' and attempting one pooled analysis of the region for homogeneity, the group of sites was found to be homogeneous with a test statistic = 376.55, the corresponding Chi-Square value @ 5% significance = 414.79

125 sites each with 40 years of data = 5000 station years.

5.9 – Summary

This chapter has tried to explain some of the controlling factors, and to understand the limitations associated with the available data when analysed.

This chapter started by introducing the method that has been used to generate the synthetic, multi-site rainfall data. It then went on to explain the Netmax concept and the effective number of sites in a pooling group.

A method for homogeneity testing was proposed and explored, the objective of which was to optimise pooling groups for extreme value estimates, therefore reducing uncertainty. This method requires further analysis. Hydrologically, however, the potential of this method is apparent and very attractive.

This page is intentionally blank.

Chapter 6

This page is intentionally blank.

Chapter 6 – Using a non-stationary model to represent and test for a trend in Annual Maxima rainfall

6.1 – Introduction

This chapter introduces and describes the GEV distribution, its parameters and the annual maxima growth curve with which they are associated. It also describes how a trend can be simulated and tested for, first by using synthetic rainfall data, and then by testing for trends in observed or recorded annual maxima rainfall.

A non-stationary model is one that is changing with time. The models presented in this chapter assume ‘climate change’ in relation to extreme rainfall can be modelled, and more importantly detected, using the parameters of an Extreme Value distribution.

Non-stationary data sets can be defined as having statistical properties that do change over time; more precisely, the probability distributions of the process are time-variant. A simple summary of a non-stationary data set is that the mean, variance and covariance of the distribution can all change with time, either individually or proportionally.

An assumption or even a requirement of most distribution fitting techniques is that the data be from a stationary distribution, meaning with no climate change. One significant advantage of the Maximum Likelihood (MLE) technique over Linear Moments (L-Moments) is that MLE is able to fit to a non-stationary data set and detect whether the data set is non-stationary by using a likelihood ratio test (LRT). All of these properties will be demonstrated in this chapter using synthetic data. However, at the end of this chapter, the tools that have been developed and proven will be applied to 179 observed annual maximum time series from Great Britain.

6.2 – Non-stationary GEV distribution

The distributions that will be used are the Generalized Extreme Value (GEV) distributions, of which there are 3 special cases:

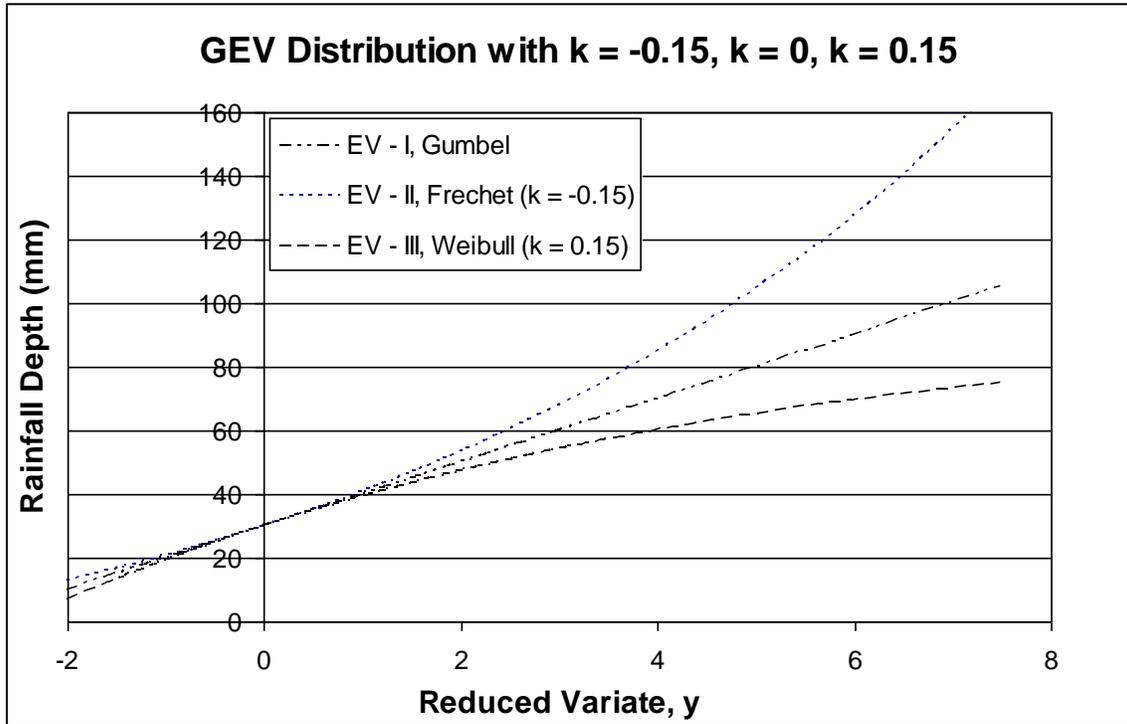


Figure 6.2.1 – GEV distribution with shape parameters $k=-0.15, k=0, k=0.15$

If X is a random variable with $GEV(\mu, \alpha, k)$ distribution, then:

Equation 6.2.1:
$$\Pr(X \leq x) = F(x) = \exp \left\{ - \left[1 - k \left(\frac{x - \mu}{\alpha} \right) \right]^{\frac{1}{k}} \right\}$$

If: $k < 0, -\infty < x \leq \mu + \frac{\alpha}{k}$ (EV-II, Frechet Distribution)

If: $k > 0, \mu + \frac{\alpha}{k} \leq x \leq \infty$ (EV-III, Weibull Distribution)

and a special case of the GEV, the Gumbel distribution, equation 6.2.2.

Gumbel:

Equation 6.2.2:
$$F(x) = \exp \left\{ - \exp \left[- \left(\frac{x - \mu}{\alpha} \right) \right] \right\}$$

If: $k = 0$ (EV-I)

Where: μ = Location parameter;
 α = Scale parameter; and,
 k = The Shape parameter.

The parameters μ and α correspond to the mean and standard deviation, the other parameter, k , gives an indication of how skewed the distribution is.

The GEV distribution is justified because under very broad conditions, for any sequence of independent, identically, distributed (iid) random variables X_1, X_2, \dots, X_n , the GEV distribution is the limiting distribution of $\max \{ X_1, X_2, \dots, X_n \}$ as $n \rightarrow \infty$, after appropriate normalisation (As explained by Coles, 2001).

In practise, this means that in many realistic situations, the GEV distribution is a very good approximation for the distribution of maxima obtained over fixed time intervals.

As has already been stated this distribution has three forms:

1. Gumbel ($k=0$)
2. Frechet ($k<0$)
3. Weibull ($k>0$)

These have been shown graphically in Figure 6.2.1.

6.2.1 – GEV parameters

Chapter 6.1 introduced the GEV distribution and demonstrated the changes upon the distribution caused by varying k , the shape parameter. This chapter demonstrates in greater detail the significance of all three parameters upon the potential estimates returned by the distribution and will allow the reader to visualise the impact of a trend (year on year increase) in one or more of the GEV parameters.

Chapter 6.2 has already explained how the parameters μ and α correspond to the mean and standard deviation; the other parameter, k , gives an indication of how skewed the distribution is. A number of graphs follow that demonstrate much more clearly the effects of a change in one or more of these parameters.

The Location parameter, μ , controls the y-intercept and the mean positioning in the y-axis. This is illustrated in Figure 6.2.1.1.

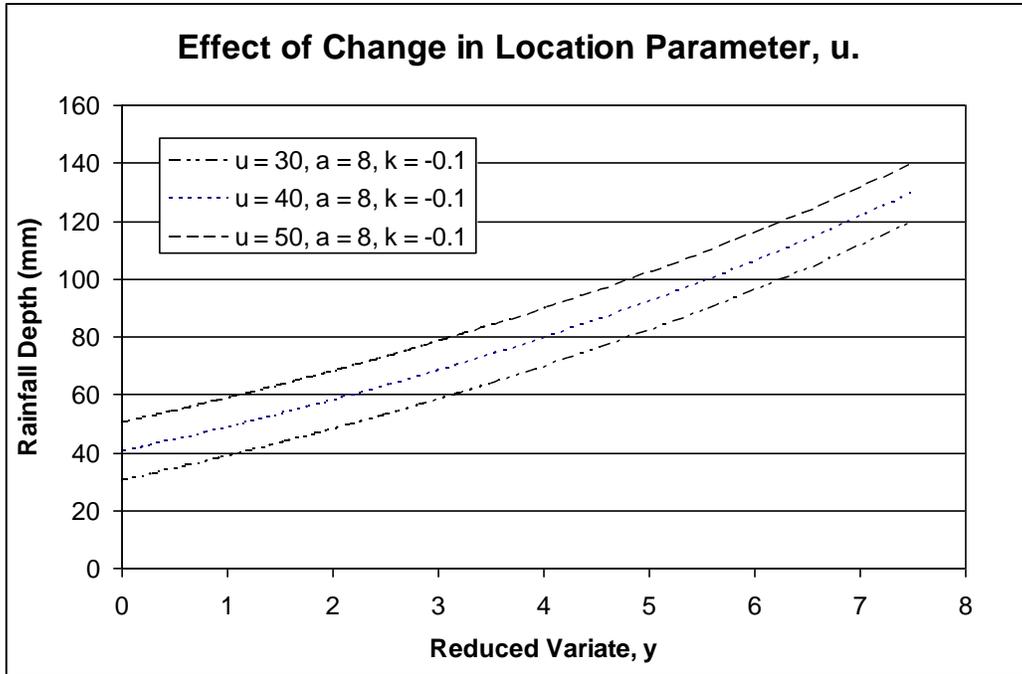


Figure 6.2.1.1 – Graph showing the effect of varying μ , the Location parameter.

The Scale parameter, α , controls the overall gradient, as illustrated in figure 6.2.1.2

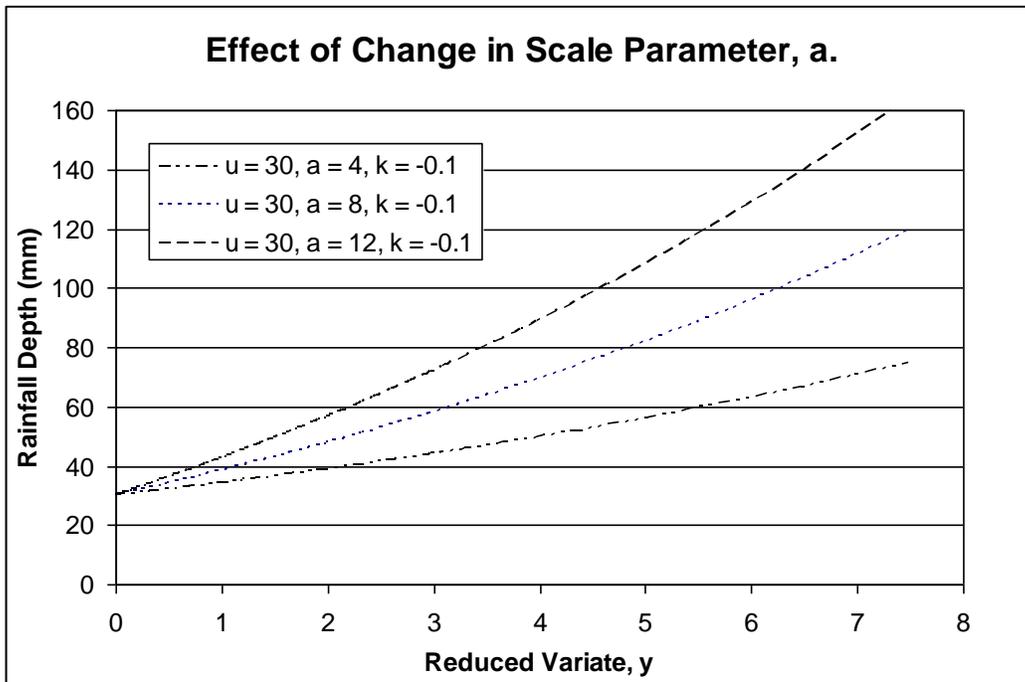


Figure 6.2.1.2 – Graph showing the effect of varying α , the Scale parameter.

The Shape parameter, k , controls the asymptotic curvature. This is shown in figure 6.2.1.3 below.

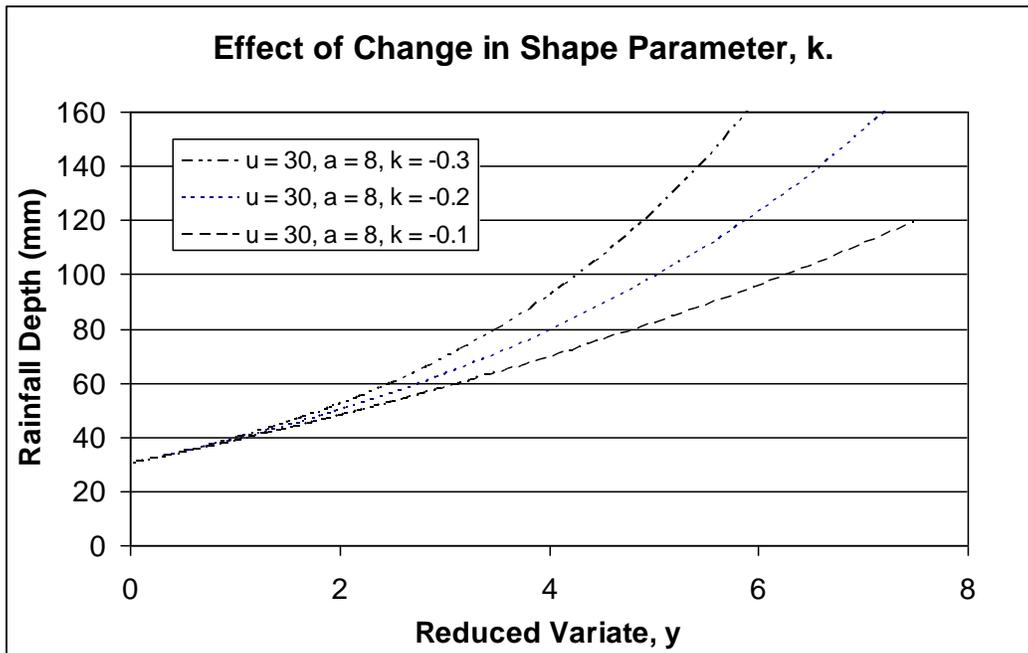


Figure 6.2.1.3 – Graph showing the effect of varying k , the Shape parameter.

6.3 - Generation of a trend

Trends have been introduced into synthetic data sets using a constant linear increase for each time interval, which in this case will be annual. There has been no additional complication in the form of seasonal cycles as it is intended to generate and to make use of the annual maxima only – the largest event recorded in a year. The data has been generated using a ‘random normal number generator’ within the R-package as per the method in chapter 5.3. The difference will be that the parameter values which are used will not be stationary and will therefore increase with each time step.

The time series for a parameter can be represented as:

Equation 6.3.1
$$\mu(t) = B_0 + B_1 t$$

Where: $\mu(t)$ = the parameter value at time interval t .

B_0 = the starting value of the parameter;

B_1 is the incremental increase; and,
 $t =$ is the time interval.

The incremental increase, B_1 , which can be detected, has been found using a sensitivity analysis of the synthetic data in the trend detection model. Some examples of the amounts by which a parameter would need to change if detection were to be possible using this method, are shown in figures: 6.4.3.1 to 6.4.3.5.

6.4 – Trend detection

6.4.1 - Method

The method of Maximum Likelihood (ML) produces parameter estimates that maximise the sample likelihood. These parameter estimates are known as the Maximum Likelihood Estimates (MLE). Having chosen a probability distribution model, in this case the GEV distribution, then the method of ML optimises or maximises the parameters to give the most likely estimates. It is possible, however, to introduce additional parameters, or more accurately replace one parameter for multiple parameters. This is what has been done with the introduction of trends and this is what ML estimation allows for in the detection of trends.

It is not enough, however, to simply fit additional parameters to a distribution. A form of hypothesis testing is required to distinguish between the Null hypothesis, which is that the time-series is stationary and therefore fitted to by a simpler model (one which does not contain a time varying parameter) and the alternative hypothesis, which is that the time-series is non-stationary (includes a time varying component) and therefore fitted to by the more complicated model.

The hypothesis test takes the form of the Likelihood Ratio Test (LRT). This is a statistical test of the goodness-of-fit between two models. As already described, a relatively more complex model is compared to a simpler model to see if it fits a particular data set significantly better. Adding additional parameters will always result in a higher

likelihood score. The likelihood ratio test begins with a comparison of the likelihood values for the two models:

$$LR = 2 * (\ln[L2] - \ln[L1]) \quad \text{Equation: 6.4.1.1}$$

Where: LR = the likelihood ratio

L1 = Likelihood for test 1 – stationary model (not changing with time)

L2 = Likelihood for test 2 – non-stationary model (time varying)

The likelihood value L, is calculated during the Maximum Likelihood parameter estimation optimisation.

The LRT statistic approximately follows a chi-square distribution. To determine if the difference in likelihood scores among the two models is statistically significant, the number of degrees of freedom needs to be considered. The number of degrees of freedom is equal to the difference in the number of parameters for the two models being compared. If the LRT statistic is greater than the Chi-square distribution statistic for the number of degrees of freedom and also at the chosen significance level, then the Null hypothesis has been shown to be incorrect, for a chosen significance level. In this case the result would indicate that the time-series is not stationary, or is non-stationary and contains a trend.

N.B. - As has already been stated, the LRT is assigned a significance level. This gives an indication of how likely it is that a stationary data set could be interpreted as a non-stationary data set due to the chance occurrence of a trend. A simple example of this would be: at the 5% significance level, it would be expected, that on generating a random data set 100 times, that a trend be detected 5 times out of 100 when one does not exist (referred to as a Type I error). Caution must therefore be exercised when looking at multiple data sets for this reason.

To counter this problem when dealing with synthetic data (containing a known trend), a measure of the ‘power’ of the test can be used to better explain the results from the significance test. The power and power curve show the following relationship:

$$\text{Equation 6.4.1.2} \quad \text{Power} = \frac{\text{Number of Significant Sites}}{\text{Total Number of Sites}}$$

The definition ‘significant site’ is used to denote a site where a trend has been detected.

It is interesting to note that the significance level and the point at which the power curve is interpreted to yield a sufficiently powerful result can be chosen by the ‘user’; they are arbitrary values that must be chosen with the test, and the impact of the decision in mind.

6.4.2 – Trend Detection using a two parameter, Gumbel ($k=0$) distribution

Initially, it was decided to simplify the problem of trend detection in the UK by using the Gumbel Distribution. This decision was made for two reasons:

1. The shape parameter for one-day annual maxima rainfall at individual sites varies from $-0.5 < k < 0.5$. However, a significant proportion (approximately 70%) of single site k values fall into the range of $0 > k > -0.2$, as shown by Figure 6.4.2.1. Further, regional pooling groups tend to have a shape parameter in the region of $-0.15 < k > 0$, it was therefore believed to be a reasonable simplification to make.
2. It was not known how well the fitting method would handle the complexity of fitting multiple parameters with a covariate and whether it would be possible to interpret the fitted parameters, other than to say that a trend had been detected.

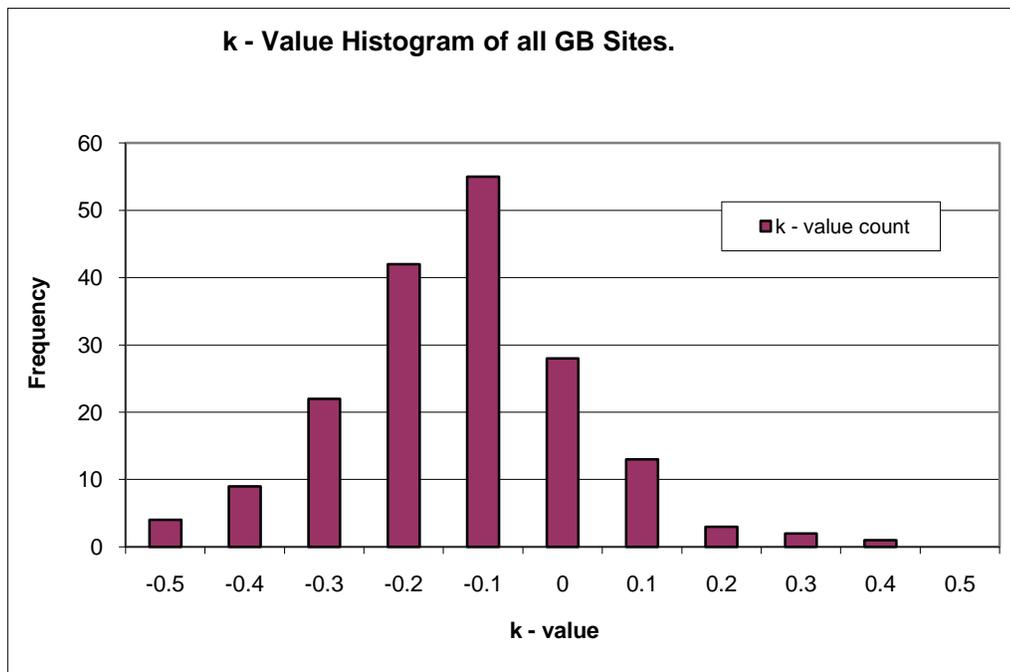


Figure 6.4.2.1 – Histogram of k values for 179 sites in Great Britain

When looking at extremes, especially in excess of (rarer than) the 100-year return period event, single site time series would not be used – due to the increased uncertainty. Regional pooling group ‘k’ values (for Great Britain) range from approximately -0.15 to approximately 0, and are shown in table 6.4.2.2, below.

Regional pooling group	Shape parameter, k
South East England	-0.023
South West England	-0.146
Central Eastern England	0.0328
North West England	-0.09
North East England	-0.0215
Northern Scotland	0.033
Eastern Scotland	0.069
Southern Scotland	-0.098

Table 6.4.2.2 – Shape parameter k, for regional pooling groups in Great Britain

Except for the South West of England (SWE), all of the other regions could be approximated to Gumbel, $k=0$, for the purpose of this analysis. In addition, if the level of uncertainty for all quantile estimates is included, then once again it could be argued that the Gumbel distribution can adequately describe the SWE. See figure 6.4.2.3, below.

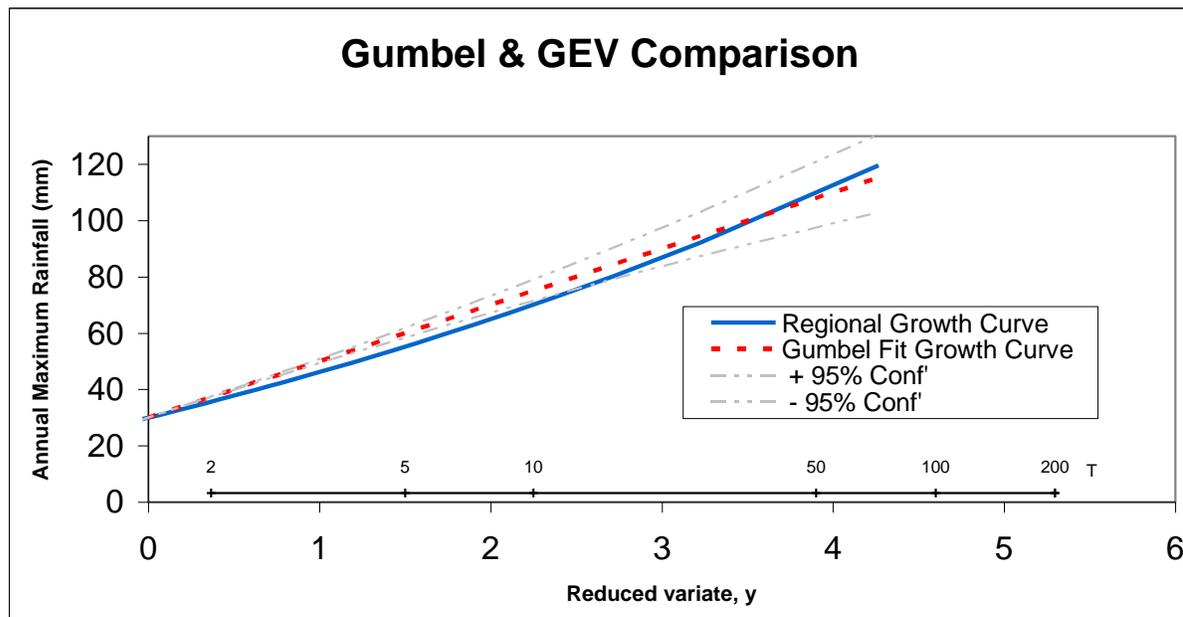


Figure 6.4.2.3 – Comparison of Gumbel and GEV fit to the SWE pooling group.

Figure 6.4.2.3 shows a regional growth curve where $k = -0.15$.

Having demonstrated that it is not unreasonable to use the Gumbel distribution, primarily for reasons of simplification, the results of the sensitivity testing of this technique and the parameters follow a short explanation of the sensitivity testing procedure.

6.4.3 - Sensitivity testing of parameters

The aim of this chapter is to test for the level of trend required in each parameter, and combinations of these, to enable detection. The test used to detect the trend has been defined in chapter 6.4.1. The ‘power of detection’, as defined by equation 6.4.1.2, has been plotted for each time series tested. This has been repeated for a range of trends in the available parameters.

Starting values for the parameters were chosen by looking at observed standardised parameters within the UK. Taking the regional mean of a parameter as the starting point, for example B_0 , then using 100th ($B_0/100$) of the observed parameter range as a starting point for the incremental increase, B_1 , it was possible to then carry out a sensitivity analysis (varying the parameter until a range of results were obtained) and allowing the model to indicate the ‘power’ associated with the trend in one or more parameters. The model has been structured to generate 1000 samples of length, $t = 40, 60, 80, \dots, 200$. This allows the reader to see the impact of both the ‘magnitude’ of the trend and the ‘length’ of the data set or time series, allowing a greater understanding of the trend to be gained. A representative sample of the results follows:

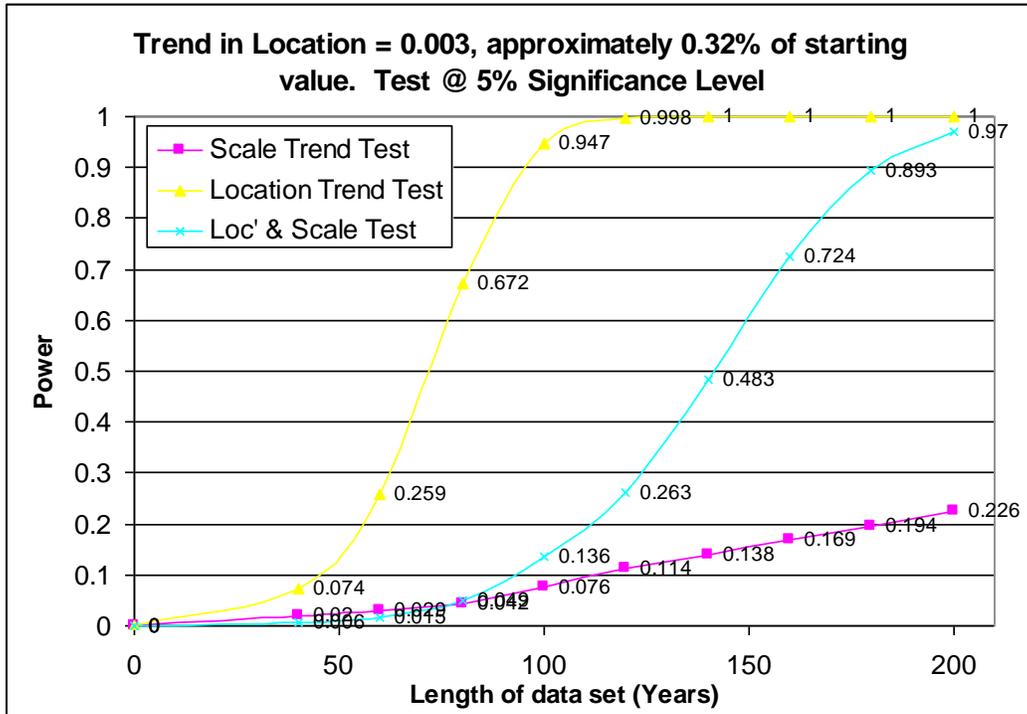


Figure 6.4.3.1 – Power curve for Trend Detection in the Location parameter within a Gumbel Distribution.

Figure 6.4.3.1 demonstrates the power of detection for a trend of 0.32%/yr in the location parameter. It is important to realise that a 0.32% annual increase in the starting value of the location parameter equates to a 32% increase at $t=100$ and a 64% increase when $t=200$.

Figure 6.4.3.1 also demonstrates that:

1. The correct parameter (the one containing a trend) was identified by the test; and,
2. The single parameter test is more powerful than the multi-parameter test when the trend exists in one parameter only.

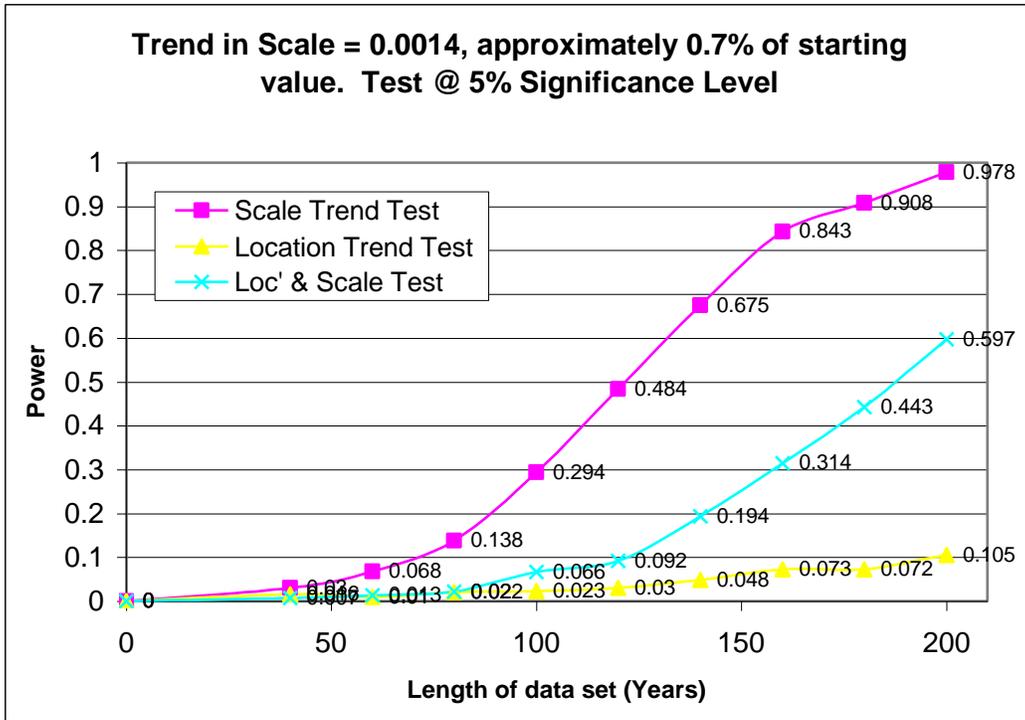


Figure 6.4.3.2 – Power curve for Trend Detection in the Scale Parameter within a Gumbel Distribution

Figure 6.4.3.2 shows a result at the lower end of the acceptable range of the sensitivity analysis, for a trend in scale parameter. This demonstrates the power of detection for a trend of 0.7% in this parameter. A 0.7% annual increase in the starting value of the scale parameter equates to a 70% increase at $t=100$ and a 140% increase when $t=200$.

Figure 6.4.3.2 demonstrates that:

1. The correct parameter (the one containing a trend) was identified by the test; and,
2. Once again the single parameter test is more powerful than the multi-parameter test when the trend exists in one parameter only.

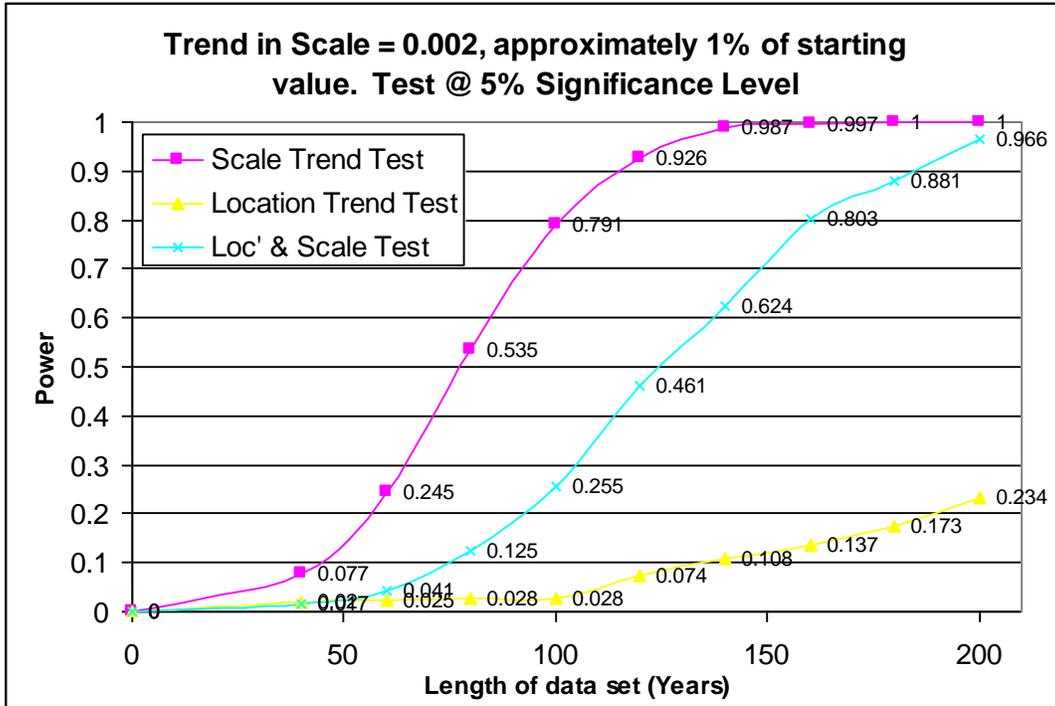


Figure 6.4.3.3 – Power curve for Trend Detection in the Scale Parameter within a Gumbel Distribution

Figure 6.4.3.3 shows a larger trend in scale parameter; 1% instead of 0.7%. This demonstrates the increased power of detection. A 1% annual increase in the starting value of the scale parameter equates to a 100% increase at $t=100$ and a 200% increase when $t=200$.

Figure 6.4.3.3 demonstrates that:

1. The correct parameter (the one containing a trend) was identified by the test; and,
2. Once again the single parameter test is more powerful than the multi-parameter test when the trend exists in one parameter only.

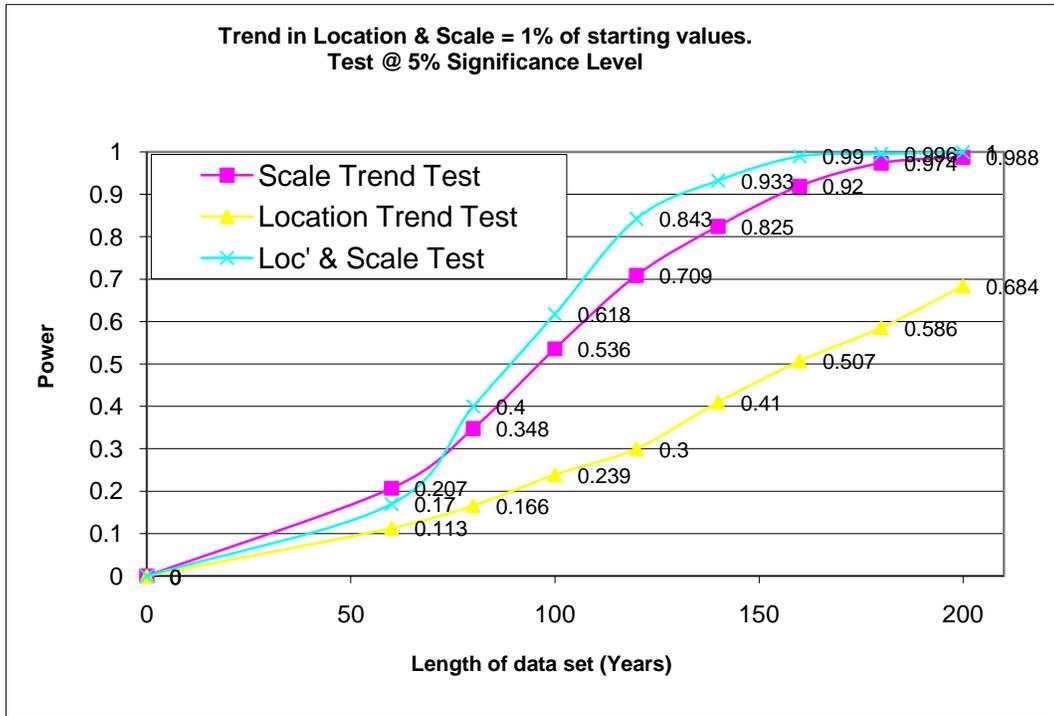


Figure 6.4.3.4 – Power curve for Trend Detection in the Location and Scale Parameters simultaneously, within a Gumbel Distribution.

Figure 6.4.3.4 shows that once again the test correctly identifies the source of the trend. Looking at the results (as illustrated by the figures within this chapter), the reader will start to appreciate the annual percentage increase in each parameter that is required for the likelihood ratio test to have sufficient power of detection.

Parameter	Annual % change	Length of dataset (Yrs)	Total % increase ¹	Power
Location	0.32	80	25.6	0.672
		100	32	0.947
Scale	1	80	80	0.535
		100	100	0.791
Both	1	80	80	0.4
		100	100	0.618

Table Notes:
1 – This is the total increase in the selected parameter.

Figure 6.4.3.5 – Summary of Power Curves – Gumbel Distribution only.

These results (figure 6.4.3.5) show the relationship between the percentage increase of the applied trend, and the time period required by this method to successfully detect the trend using synthetic data. The location parameter appears to be the most amenable parameter for early detection of change in extreme rainfall (for the distribution tested).

6.4.4 – Trend Detection using a 3 parameter GEV ($k \neq 0$) distribution

Having demonstrated that the proposed trend detection method works for the Gumbel distribution (two parameters, $k=0$), it was decided to increase the complexity of the trend detection test to include the k (shape) parameter, therefore using all three parameters of the GEV distribution.

It is also important to note that when this technique is applied to observed data sets (instead of synthetic data), it will be to single site data, meaning the available record lengths will be approximately 40 years. As section 6.4.2 demonstrated, variation in the k parameter for single site analysis is such that it is difficult to defend the use of the Gumbel based trend detection test if a full GEV trend detection test exists.

The graphs/results that follow (overleaf) are similar in appearance to those in chapter 6.4.3. The only significant difference is the addition of the k parameter.

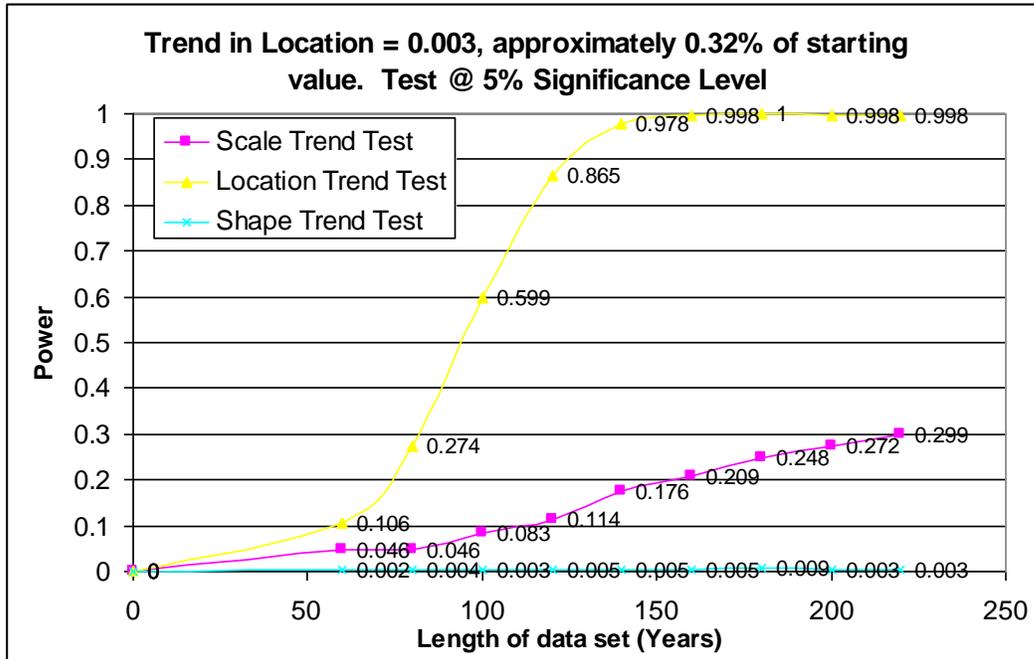


Figure 6.4.4.1 - Power curve for Trend Detection in Location Parameter using the full GEV Distribution

Figure 6.4.4.1 demonstrates the power of detection for a trend of 0.32% in the location parameter. As stated previously, a 0.32% annual increase in the starting value of the location parameter equates to a 32% increase at $t=100$ and a 64% increase when $t=200$. Figure 6.4.4.1 also demonstrates that the correct parameter (the one containing a trend) was identified by the test.

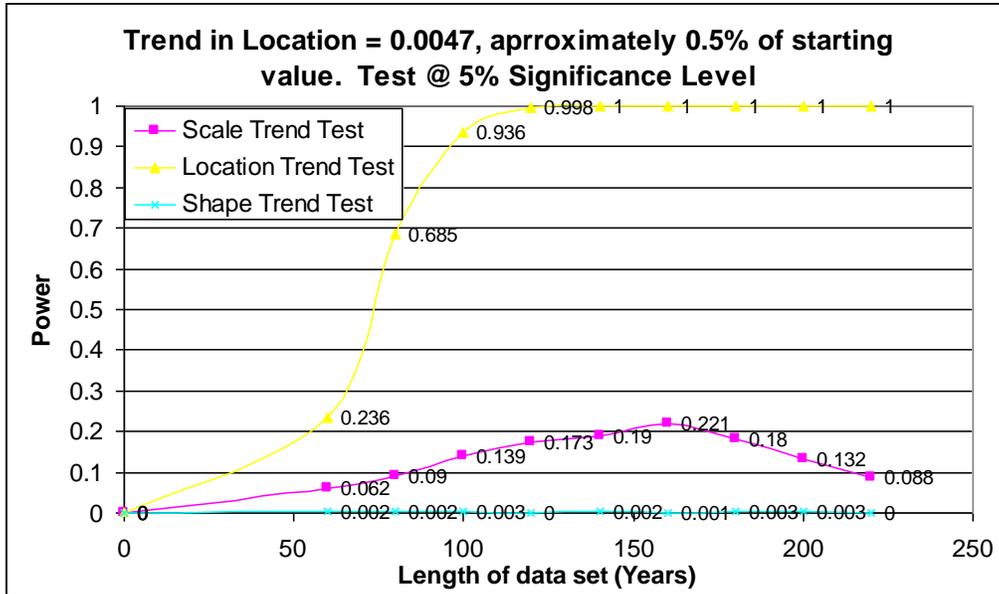


Figure 6.4.4.2 - Power curve for Trend Detection in Location Parameter using the full GEV Distribution.

Figure 6.4.4.2 demonstrates the power of detection for a trend of 0.5% in the location parameter. As stated previously, a 0.5% annual increase in the starting value of the location parameter equates to a 50% increase at $t=100$ and a 100% increase when $t=200$.

Figure 6.4.4.2 also demonstrates:

1. That the correct parameter (the one containing a trend) was identified by the test;
2. That the power of detection has increased inline with the increased annual trend; and,
3. Where the test has incorrectly detected a trend (Scale parameter, albeit with a significantly lower power), with additional data/stronger trend, the power is seen to reduce after a period of time.

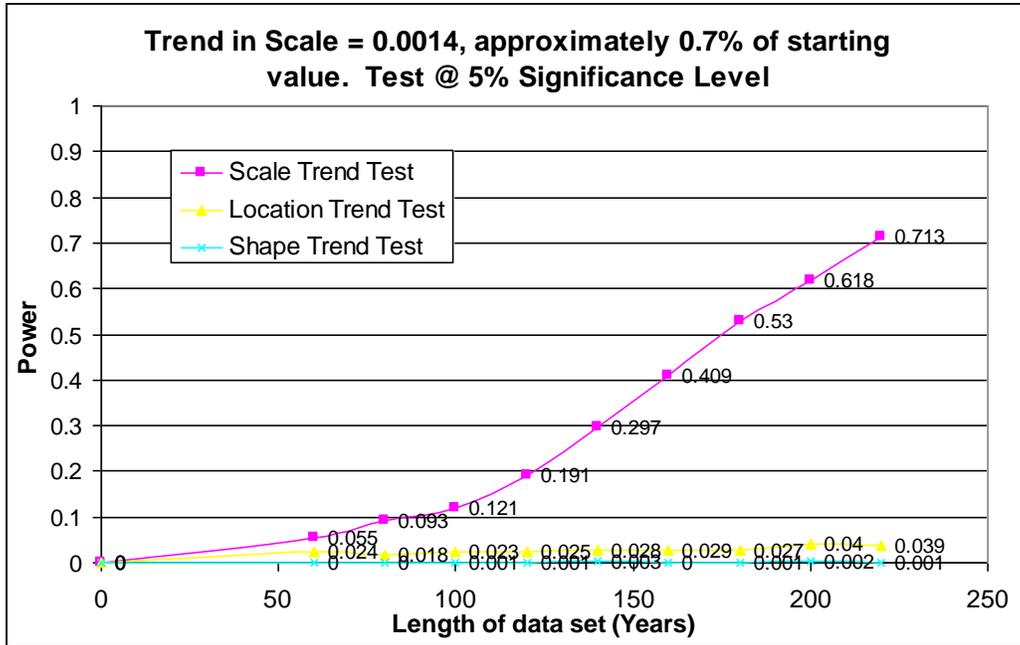


Figure 6.4.4.3 - Power curve for Trend Detection in Scale Parameter using the full GEV Distribution

Figure 6.4.4.3 demonstrates the power of detection for a trend of 0.7% in the scale parameter. As stated previously, a 0.7% annual increase in the starting value of the scale parameter equates to a 70% increase at $t=100$ and a 140% increase when $t=200$.

Figure 6.4.4.3 also demonstrates that the correct parameter (the one containing a trend) was identified by the test.

Comparing figure 6.4.4.3 with figure 6.4.3.2, it appears that the Gumbel trend detection test is more powerful on this occasion.

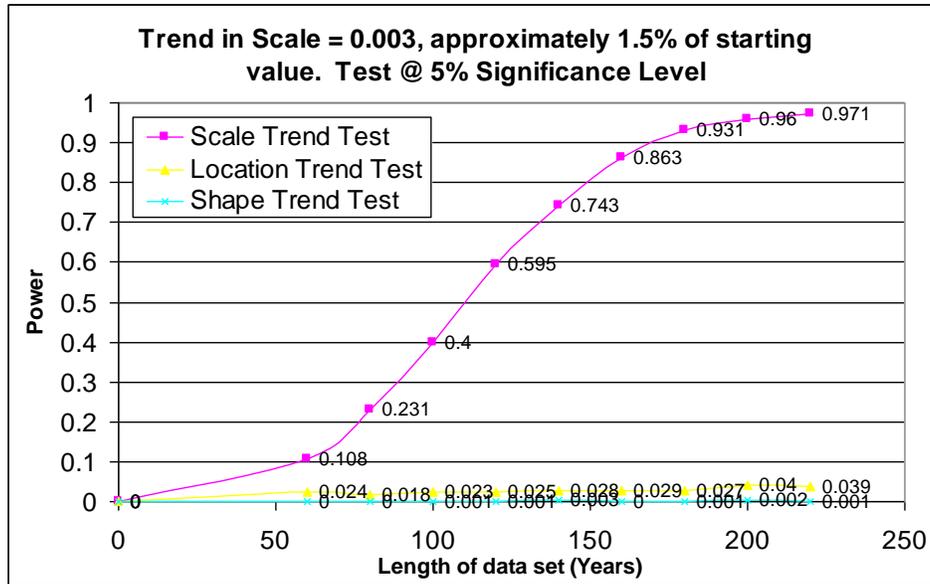


Figure 6.4.4.4 - Power curve for Trend Detection in the Scale Parameter, using the full GEV Distribution

Figure 6.4.4.4 demonstrates the power of detection for a trend of 1.5% in the scale parameter. A 1.5% annual increase in the starting value of the scale parameter equates to a 150% increase at $t=100$ and a 300% increase when $t=200$.

Figures 6.4.4.3 and 6.4.4.4 show that a doubling of the scale parameter (0.7% to 1.5%) is required to achieve a similar power of detection to those observed in the preceding tests. It is important to note that the test continues to correctly identify the parameter containing a trend.

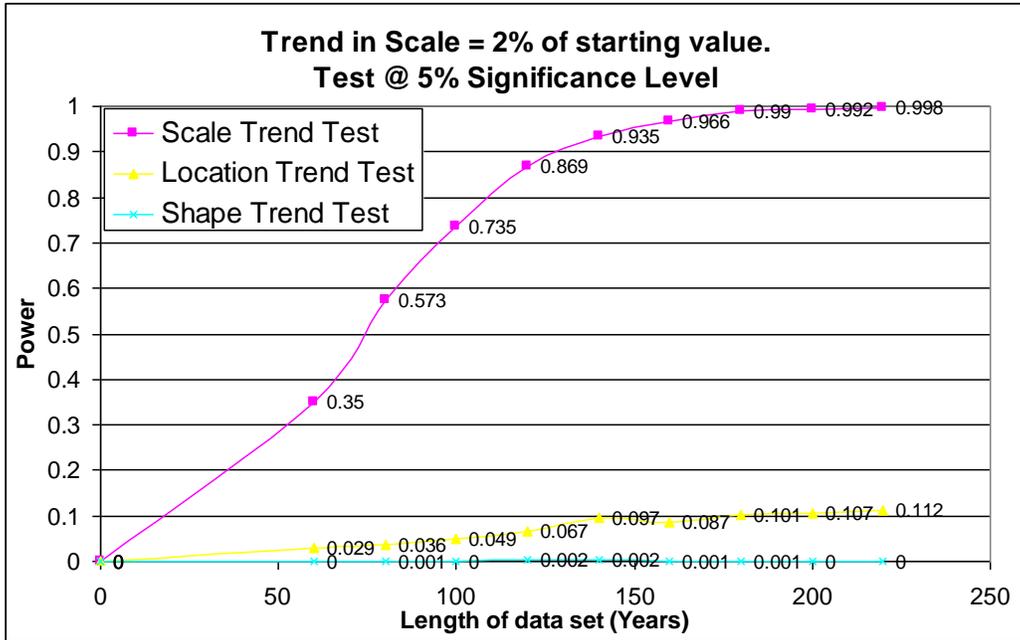


Figure 6.4.4.5 - Power curve for Trend Detection in Scale Parameter using the full GEV Distribution

Figure 6.4.4.5 demonstrates the power of detection for a trend of 2% in the Scale parameter. A 2% annual increase in the starting value of the scale parameter equates to a 200% increase at $t=100$ and a 400% increase when $t=200$.

Figures 6.4.4.6 to 6.4.4.10 show the varying performance of the power of detection test associated with trends in the shape (k) parameters; where the trend ranges between 1.9% and 6.67%.

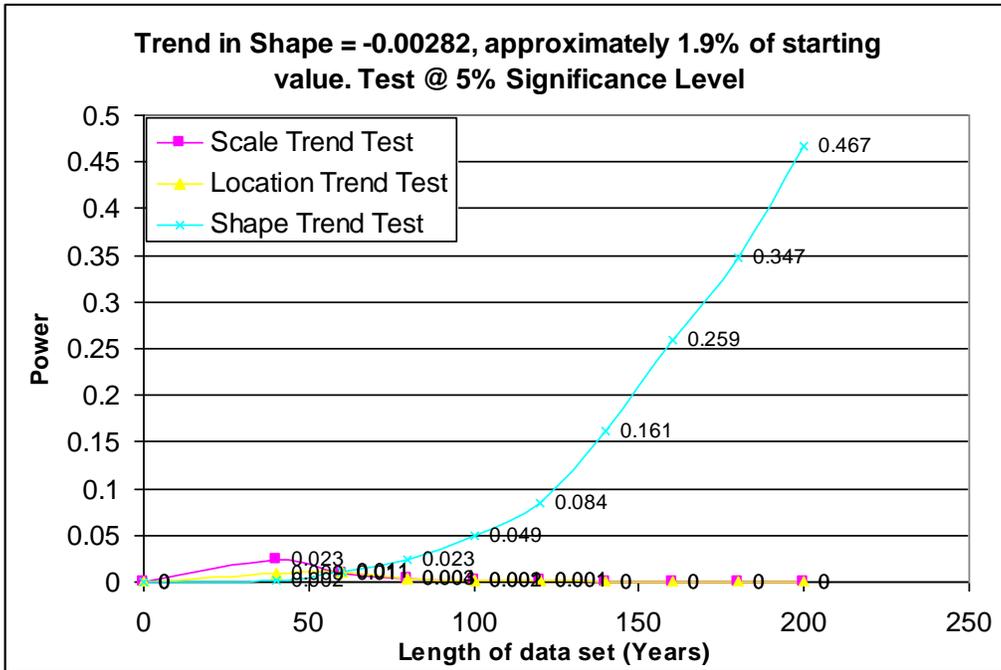


Figure 6.4.4.6 - Power curve for Trend Detection in Shape Parameter using the full GEV Distribution

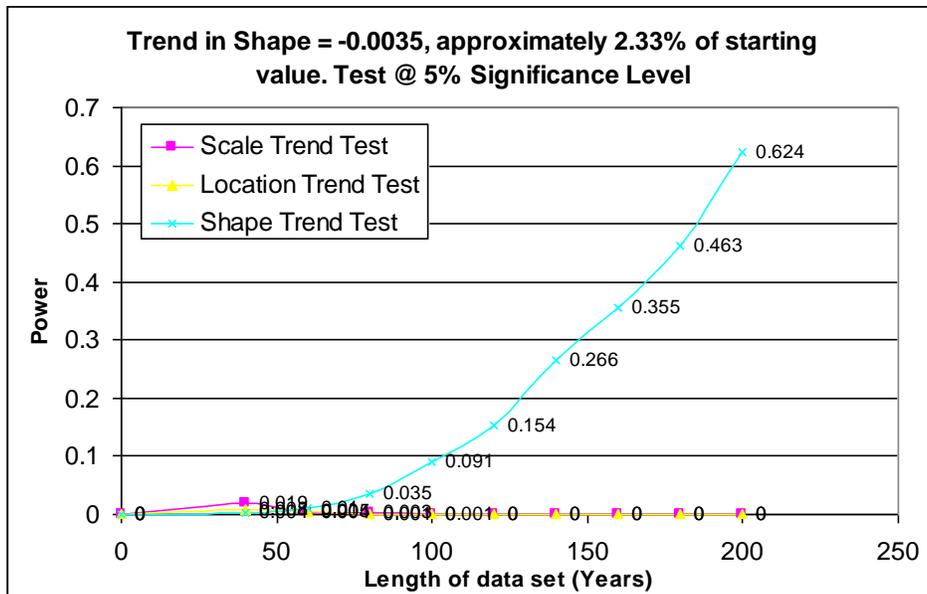


Figure 6.4.4.7 - Power curve for Trend Detection in Shape Parameter using the full GEV Distribution

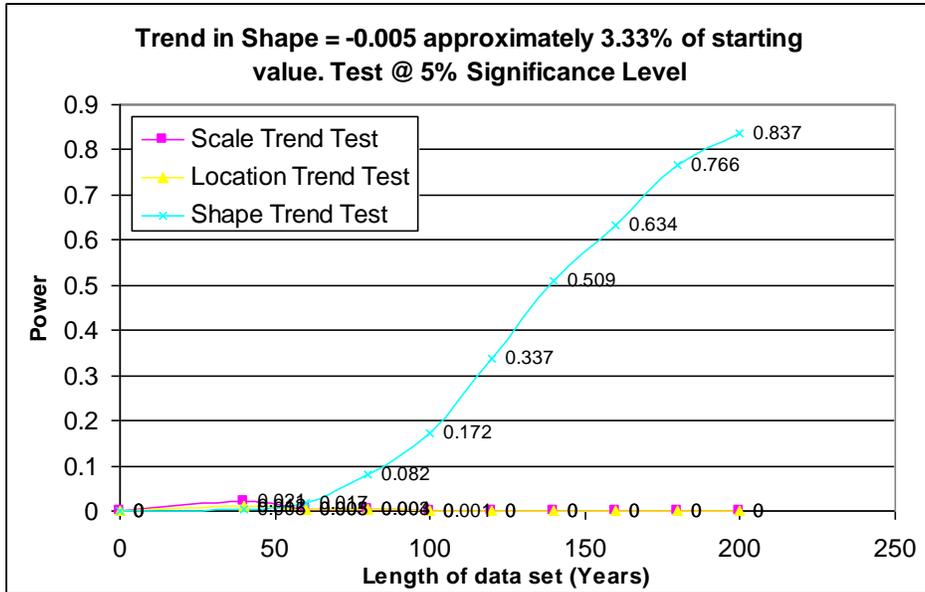


Figure 6.4.4.8 - Power curve for Trend Detection in Shape Parameter using the full GEV Distribution

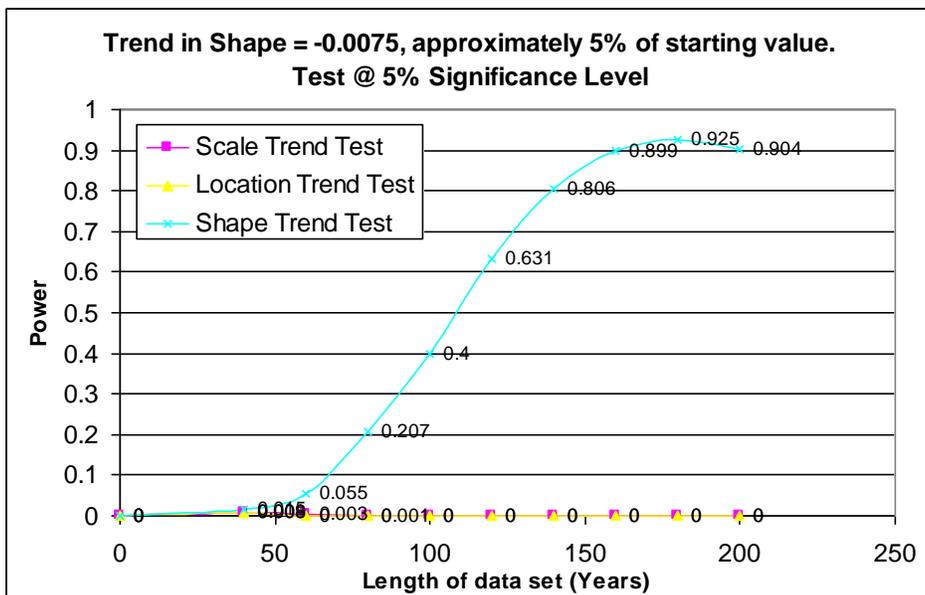


Figure 6.4.4.9 - Power curve for Trend Detection in Shape Parameter using the full GEV Distribution

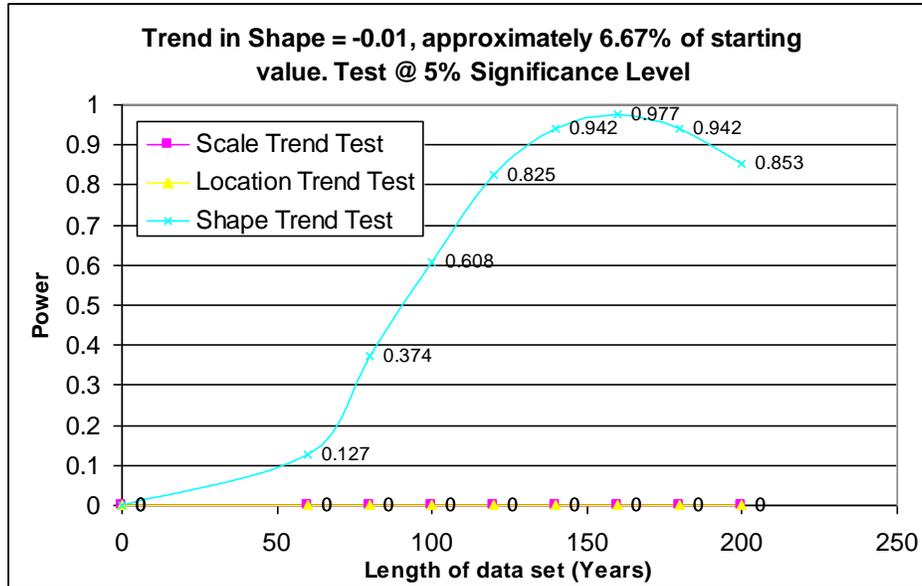


Figure 6.4.4.10 - Power curve for Trend Detection in Shape Parameter using the full GEV Distribution

By increasing the trend in the shape (k) parameter (Figures 6.4.4.6 to 6.4.4.10), the characteristic power curve becomes visible.

6.4.5 - Equivalence of changes in GEV parameters on quantile estimates

Taking the 100, 1,000 and 10,000-year return period events, and applying an increase of 10%, 25% and 50% to each parameter individually, the effects of these changes has been assessed. The results of this analysis are shown in the figures below:

Unless otherwise stated $u = 30$, $a = 8$, $k = -0.1$

Percentage Increase	100-year Estimate with varying:		
	u	a	k
0%	76.73	76.73	76.73
10%	79.73	77.31	77.90
25%	84.23	78.19	79.74
50%	91.73	79.65	83.00

Figure 6.4.5.1 – Table showing the effect of a percentage increase, in any one of the three GEV parameters, upon the 100-year return period event.

Percentage Increase	1,000-year Estimate with varying:		
	u	a	k
0%	109.61	109.61	109.61
10%	112.61	110.61	112.75
25%	117.11	112.10	117.76
50%	124.61	114.59	126.97

Figure 6.4.5.2 – Table showing the effect of a percentage increase, in any one of the three GEV parameters, upon the 1,000-year return period event.

Percentage Increase	10,000-year Estimate with varying:		
	u	a	k
0%	150.95	150.95	150.95
10%	153.95	152.46	157.58
25%	158.45	154.73	168.38
50%	165.95	158.51	188.99

Figure 6.4.5.3 – Table showing the effect of a percentage increase, in any one of the three GEV parameters, upon the 10,000-year return period event.

Figures 6.4.5.1 to 6.4.5.3 show a maximum increase in any one parameter of 50%. This may seem significant, but it is important to remember that the Scale parameter may be caused to increase by a greater amount than this due to the recording of a significantly large event. For example, at Manston in South East England, Figure 6.4.5.4 overleaf:

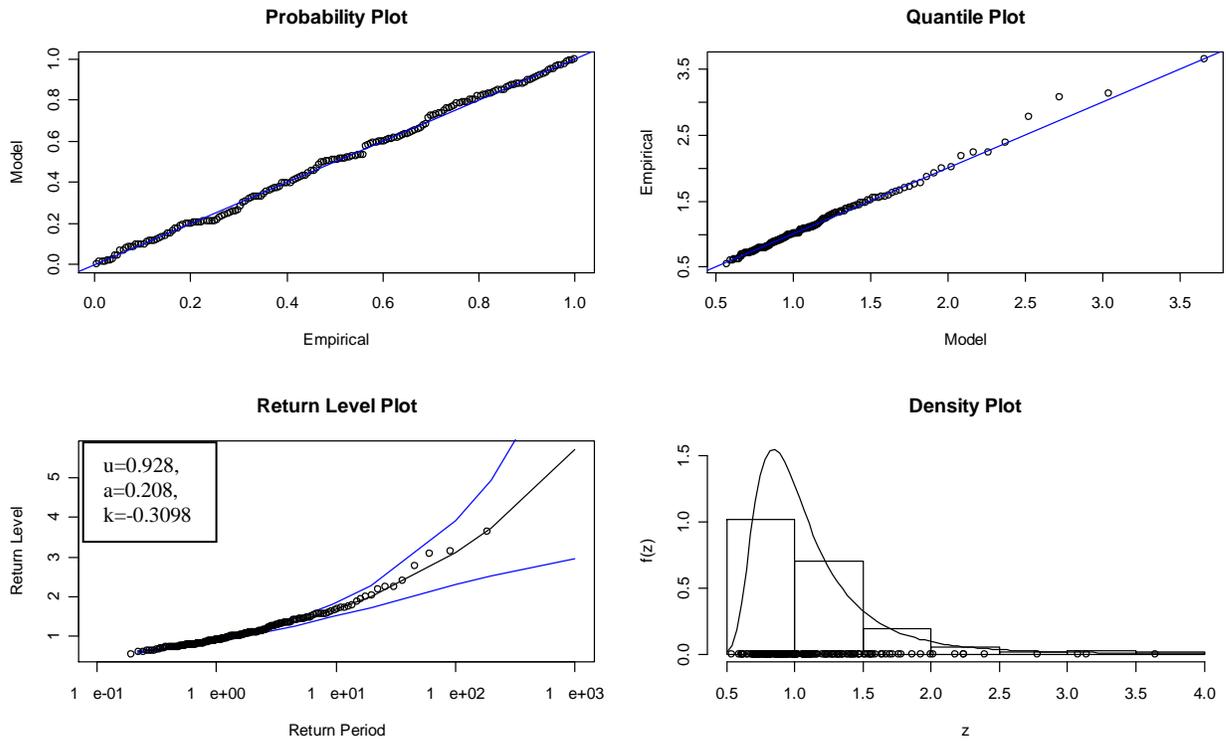


Figure 6.4.5.4 – Graphs relating to the fit of the Manston Data set, with the largest event removed from the record.

Comparing Figure 6.4.5.4 and Figure 6.4.5.5 (without and with the largest event on record) there is a noticeable increase in the k parameter from $k=-0.3098$ to $k=-0.4846$, an increase of 56.4% (in the k parameter), due to the inclusion of the largest event. This example has been included to demonstrate that such increases in a parameter, whilst large, can be observed by varying the available data. With this in mind the percentage increases required for trend detection do not seem unreasonable.

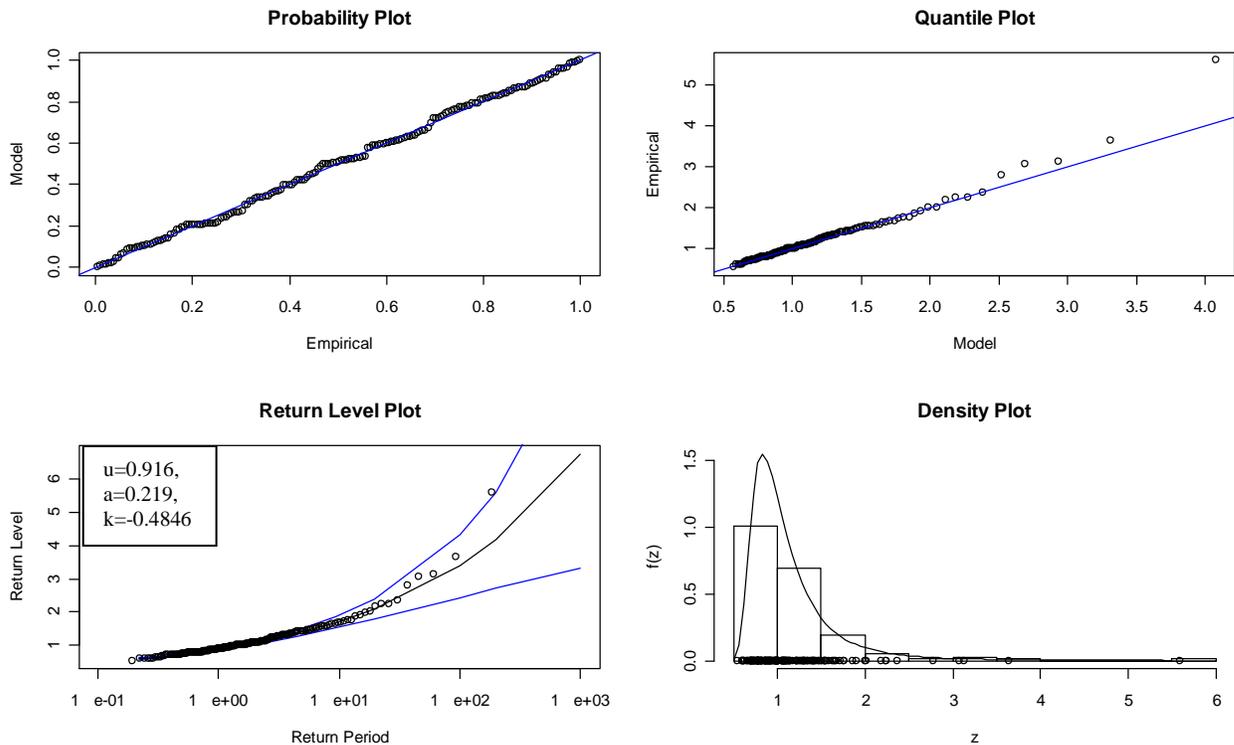


Figure 6.4.5.5 – Graphs relating to the fit of the Manston data set, with the largest event included in the record.

6.5 – Trend detection in observed annual maxima data

This section applies the trend detection test to the 179 rain gauges that have been used. When looking at observed data, decision makers might ask to be shown more detail than a test which is limited to a p value of say 0.05, believing that additional information (for other significance levels) would be of value. For this reason, it might be preferable to display the p-values for each site. It is important to allow some form of spatial awareness, because, the nature of the test is to show a significant result, 5% of the time (at the 5% significance level) even when one does not exist (type 1 error). Hence, any clustering of significant results reinforces their importance. Perhaps a better response is to inform the user of this data and allow them to choose a value with which they are comfortable. For example, the significance level of 5% excludes all test results which are greater than this value. If however, a decision maker is concerned that a number of sites

in his or her region come back with a p value between say 0.05 and 0.1, then the user may believe these results should not be ignored. For this reason, when looking at the observed annual maxima data sets, it was decided to display all of the results that were returned with a $p \leq 20\%$; this thesis does not wish not imply that these results be regarded as significant, but if they are located within a cluster of sites deemed significant at 5%, then this additional information adds weight to the primary test. These results have been plotted on maps, as well as displaying the point values; an inverse distance weighting was calculated and plotted to interpolate the point values.

Clustering of significant results has been interpreted as meaning that there is more likely to be a trend present, than for those sites where only one or two sites showed evidence of a trend. However, caution must be exercised as it is of course possible for the error sites to be located within or near to the actual cluster of sites which contain a trend.

Figures 6.5.1, 6.5.2 and 6.5.3 have been plotted (overleaf) and show the p-values (graphically) for each site, as well as a spatially correlated interpretation of the single site data.

The results that follow are for the location parameter only as this was found to contain the greatest number of trends at the 1, 5 and 10 day duration of annual maximum rainfall in Great Britain.

Interestingly, M. Ekstrom et al (2004) found that the HadRM3H projects (following the IPCC SRES scenario A2 for 2070–2100) showed a 30% increase in rainfall intensity, for longer duration events (5–10 days), and that event magnitudes at given return periods show large increases in Scotland (up to c30%), with greater relative change at higher return periods (25–50 years).

1 Day Annual Maxima Trend Detection Map of England, Scotland & Wales

N.B. Trend detection shown in Location Parameter only, however, this parameter showed the greatest number of sites with a trend.

Period of record used: 1961 – 2000.

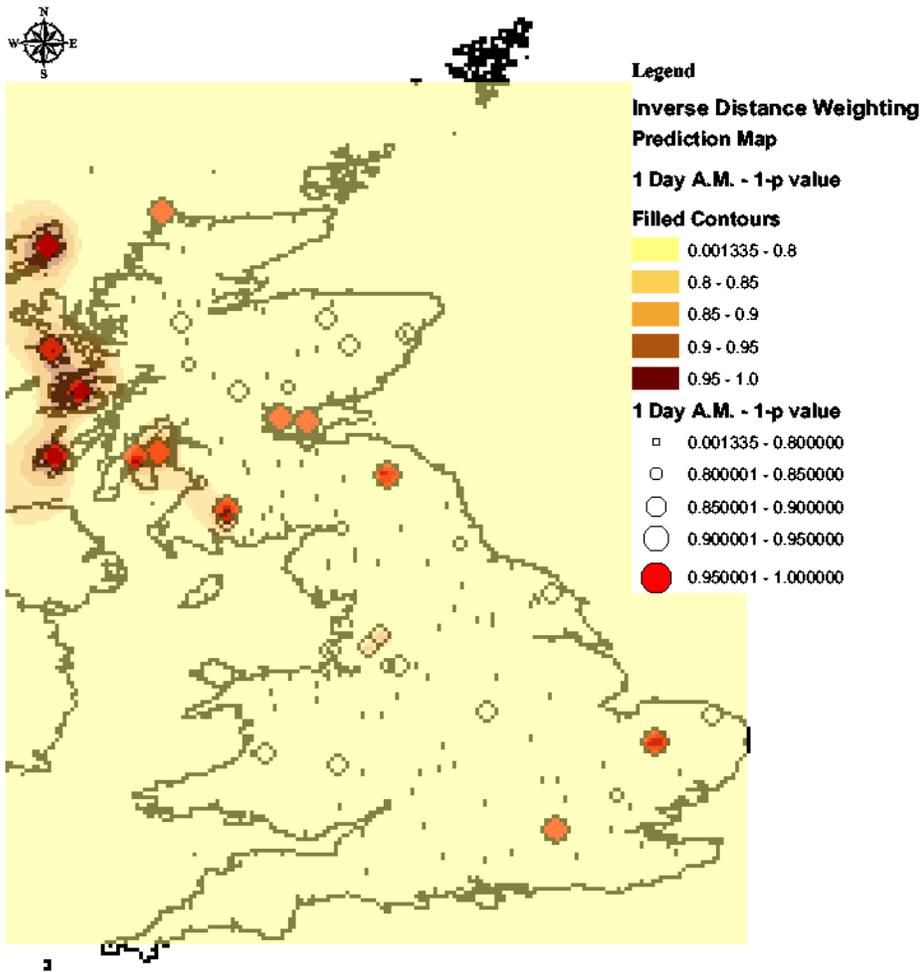


Figure 6.5.1 – Detected trends in 1 Day Annual Maximum Rainfall Data

5 Day Annual Maxima Trend Detection Map of England, Scotland & Wales

N.B. Trend detection shown in Location Parameter only, however, this parameter showed the greatest number of sites with a trend.

Period of record used: 1961 – 2000.

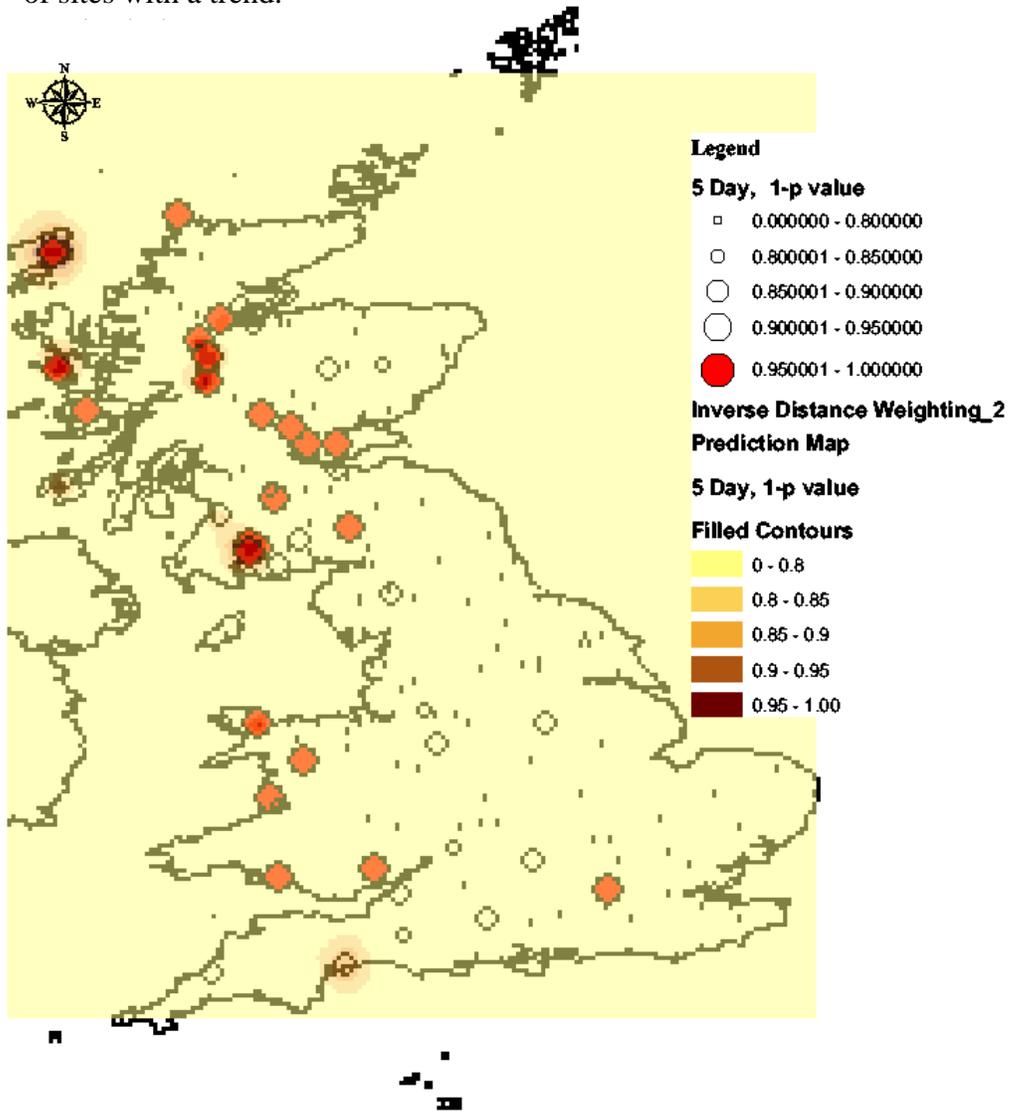


Figure 6.5.2 – Detected trends in 5 Day Annual Maximum Rainfall Data

10 Day Annual Maxima Trend Detection Map of England, Scotland & Wales

N.B.

Trend detection shown in Location Parameter only however, this parameter showed the greatest number of sites with a trend.

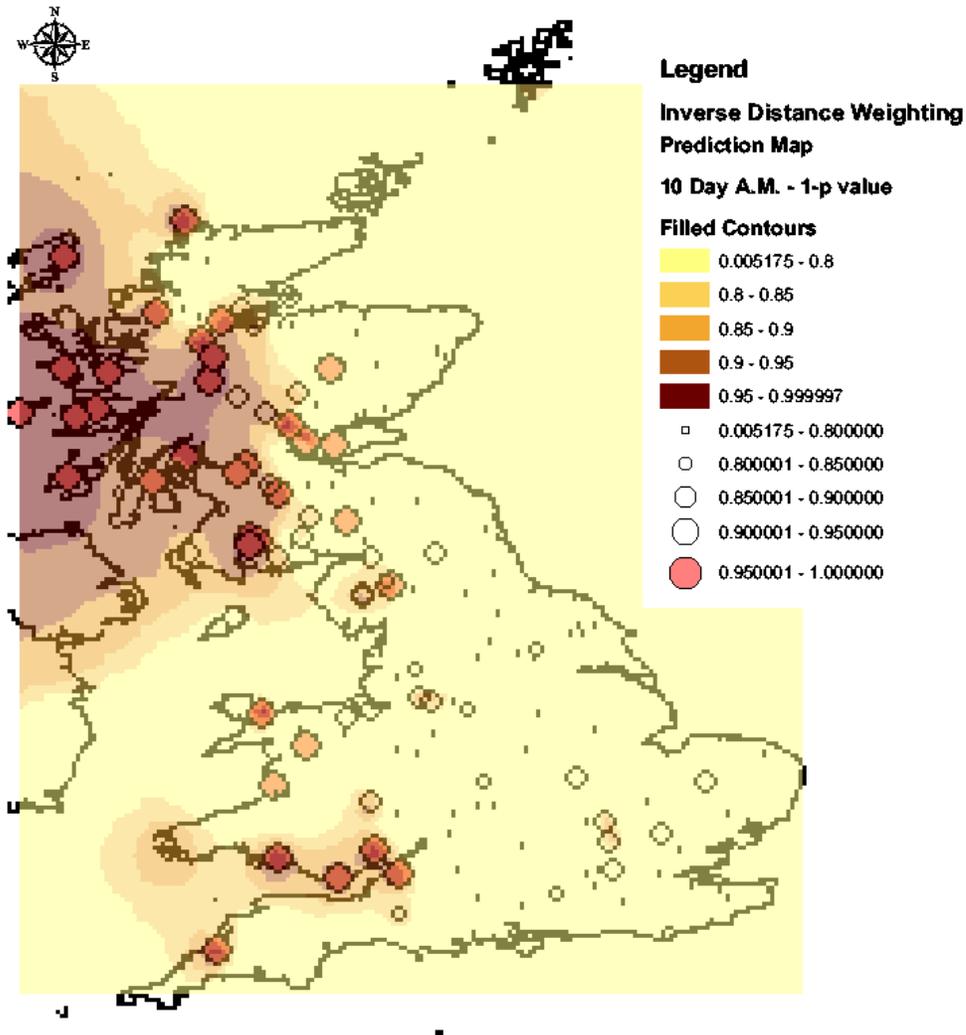


Figure 6.5.3 – Detected trends in 10 Day Annual Maximum Rainfall Data

6.6 – Summary

This chapter has introduced and described the GEV distribution, its parameters and the annual maxima growth curve with which they are associated. It has also described how a trend can be simulated and tested for using synthetic rainfall data. Having demonstrated the ability of this technique to accurately identify a trend, and having also shown the strength of the trend required for detection, this technique was applied to observed data. Interestingly and perhaps surprisingly, trends were detected. The greatest signal strength was observed at longer durations (10 day annual maxima) over western Scotland.

Reviewing the figures in this chapter leads to an important question: at what power can the statement ‘a trend has been detected’, be justified? This power test can only be carried out at a regional or national level. However, it is important to remember when looking at the power curve that 5% (or the chosen significance level) could be errors and therefore not contain a trend (Type 1 error).

In all of the synthetic trend scenarios observed in this chapter, when in error, none of the incorrectly identified trends exceeded a power of 0.3 or 30%. Following more analysis, it might be sufficient to say that the minimum threshold for detection, should be set just above this maximum observed error. Ultimately, as with the significance level, it must be the analyst that makes the final decision once presented with all of the available facts.

Having seen the shape of the power curves and knowing that the majority of rainfall data sets in the UK are typically 40 – 50 years in length, with a few exceptions, this thesis recommends a power of 0.51; or 51% of sites containing a trend.

This page is intentionally blank.

Chapter 7

This page is intentionally blank.

Chapter 7 – Summary of Thesis

7.1 – Summary of Chapters

The initial two chapters introduced the aims of this study and the theory behind dam safety practice.

Chapter 3 introduced flood estimation methods currently used within the UK and Europe. It further described some of the techniques associated with the Flood Estimation Handbook, which is the accepted UK standard for flood estimation. This chapter also included a technical description of two distribution fitting techniques.

Chapter 3 also looked at PMP, with estimates of PMP being considered to be approximations that depend upon the amount and quality of the data available for applying the various methods. Further, as the WMO description of PMP states, there is no allowance for long term climatic trends. This information appears to have increasing significance in consideration of research showing that over the past 30 years there has been an increase in global temperature of approximately a 0.5°C. Of greater concern, it is forecast that this trend will continue and that an increase in temperature of between 1.5 and 4.5°C over the next 100-years could occur.

Chapter 4 then considered two distribution fitting techniques and demonstrated the strengths and weaknesses of each. It showed that the method of L-Moments appears to be more accurate for relatively short time series, but it also shows that any advantage demonstrated by L-Moments diminishes as the length of the time series increases.

Chapter 5 started by introducing the method that has been used to generate the synthetic, multi-site rainfall data and then continued to explain the Netmax concept and the effective number of sites in a pooling group. To optimise pooling groups for extreme value analysis and therefore reduce uncertainty, a method for homogeneity testing was

proposed and explored. From a hydrological viewpoint, the potential of this method is apparent and very attractive. However, this method requires further analysis.

Chapter 6 described the GEV distribution, including its parameters and the associated annual maxima growth curves. This chapter also described the use of synthetic rainfall data to simulate and test a trend. It was demonstrated that this technique can accurately identify a trend, with the technique then being applied to observed data.

7.2 – Main Outcomes

The four main outcomes of this thesis have been:

1. Gaining an increased understanding of spatial dependence and the impact this has on the effective number of sites (amount of data) in a pooling group, with analysis showing that this varies with event rarity (return period).
2. One of the underlying assumptions for the Netmax concept, used by the FORGEX method, has been shown to be wrong. This is particularly true for rare events where the return period is greater than approximately 100 years.
3. Developing a method for trend detection in the parameters of the GEV distribution.
4. Having tested this on synthetic data, it was then applied to observed data for Great Britain and trends were found.

Each of these points will now be discussed in more detail.

1. Spatial dependence and trends within hydrological and meteorological data have a significant effect upon reliable estimation of the 10,000-year (or other extreme) event. Spatial dependence has been shown to have significant impact on the amount of data in a pooling group but has also been shown to vary with return period (event rarity). This has a significant impact on the effective number of sites in a pooling group and also highlighted that the FORGEX method was in need of review (this will be explained in point 2). During the analysis of spatial dependence, this thesis also found that the effective number of sites in a region or pooling group using the Station Year method has been shown to range from 74% – 93% of the total for 1 Day Annual

Maxima (AM), and 61% - 88% of the total for 10 Day Annual Maxima rainfall data. This means that a pooling group of 10 Sites, each with 40 years of data, does not equate to a time series of 400 station years in length, but to one of perhaps $74\% * 400 = 296$ station years for 1 Day AM or $61\% * 400 = 244$ station years for 10 Day AM. This information is provided for example only. Full regional analysis for Great Britain is reported upon in chapter 5.7.

2. As discussed in chapter 5.6, the effect of spatial dependence (inter-site correlation) has been shown to vary with return period; spatial correlation reduces with increasing return period. It was this realisation that led to the questioning of the Netmax concept. This discovery is in agreement with earlier research in 1997 by CRCCH (Cooperative Research Centre for Catchment Hydrology), which demonstrated that there was not a constant separation between the Netmax and regional growth curve. This research was carried out for rainfall data in the area of Victoria in Australia.
3. In addition to the work described above, a method for trend detection in annual maxima rainfall has been tested and shown to be effective (in chapter 6). This was initially demonstrated using synthetic data with known trends. This method introduced a constant trend in one or many of the GEV parameters, the test was then applied and results have been produced showing the 'power of detection'. With this knowledge, it is possible to assess how strong a trend, and or how long a time-series is required to detect a trend.
4. Following the investigation described above, observed (recorded) rainfall data sets were analysed, using 1, 5 and 10 day duration annual maximum rainfall data for 179 rain gauges in Great Britain (from 1960 to 2000). Interestingly, trends were detected with the greatest signal strength being observed at longer durations (10 day annual maxima) over western Scotland.

7.3 Summary of methods developed to achieve the main outcomes

Research was carried out in to the accepted methods for extreme value rainfall estimates in Great Britain and parts of Europe. Reservoir design and the Reservoirs Act (1975) have also been investigated to gain insight into how these extreme values are used and why they are needed.

In parallel with a literature review, work started on a comparison of statistical methods that are widely used by hydrologists, and these were compared with a method that is widely used by statisticians for extreme value (rainfall) estimates, namely L-Moments, and MLE (Maximum Likelihood Estimates). Whilst studying these methods, it became apparent that the majority of rainfall and flood estimates are produced using techniques that assume the data (rainfall time-series) are stationary, where stationary data sets are defined as having statistical properties that do not change over time. With a growing consensus among the scientific community that the climate is changing (non-stationary), it would appear that the time has come to stop using these inappropriate techniques.

This thesis has shown, in a hydrological context, that an alternative in the form of MLE should be considered, for the following reasons:

- MLE can be adapted (without compromise) to fit to a non-stationary time series;
- MLE is able to plot confidence intervals without re-sampling the available data; and,
- MLE can test for trends in non-stationary data sets.

This thesis has shown that a great deal of uncertainty exists when extreme rainfall estimates are produced, however this is very rarely reported or made use of. One example where this is important is rainfall runoff modelling using extreme estimates.

By using MLE, this valuable information (uncertainty) could be relayed to the user, and be considered in design calculations.

7.4 – Future work

7.4.1 – Introduction

Whilst we have achieved the aims and objectives both set out and that evolved during the thesis, it is also apparent that some of this work could and should be either expanded upon or challenged as part of some future studies. The one area which is in need of questioning is the proposal for a homogeneity test which allows the formation of large pooling groups. As already explained, I have concerns over the suitability of this proposal. The concerns relate to the principle of the test, that there is a chance 5% (or chosen level of uncertainty) chance of error. Normally, this test would be applied once so the level of uncertainty is known. The iterative nature of the proposed test however, raises concerns about the suitability and accuracy of this test after multiple each additional gauge is added. Statistical analysis of this repetitive method and the potential compound errors that might be associated with it need to be investigated.

7.4.2 – Areas to be considered for future work

1. Further development of the test for detecting trends in a GEV distribution.
2. Greater analysis of the finding that spatial dependence is a function of return period.
3. The homogeneity test for defining large pooling groups.

7.4.3 – Future work explained

1. It is proposed that methods are explored which aim to increase the power of trend detection in extremes. Further, this thesis has looked at linear trends only; further work needs to be done on time varying trend detection. The problem here is that each data point must have an associated time-index; for this reason it is not possible to adopt a station-year method of pooling and fitting. However, it is possible to fit to multiple sites simultaneously and test for similarities. For example each site could be fitted too using

the hypothesis that one value of k could satisfactorily fit the distribution at each site in the pooling group. The likelihood ratio test would then return a test statistic for or against this hypothesis. This test could then be repeated for each GEV parameter. More importantly when testing for a trend, the same principle could be applied to the covariate of one or more parameters – the covariate being the time varying trend. The test would then be based on the hypothesis that the trend in say the location parameter could be fitted using one value for multiple sites, whilst allowing the other parameters to be site dependent and therefore independent of each other. This test would be especially useful where clustering has been observed in the single site trend detection test.

Fitting to the time varying distributions does return a time varying set of parameters, meaning that estimation of the changes in user defined return period event can be calculated. For example, the 100-year, 24 hour rainfall event using a stationary fitting technique may be 50mm. The non-stationary fitting technique allows an estimate to be generated of the equivalent event in 100-years time (from the calculations being carried out). In 100-years the same event might equate to a rainfall depth of 65mm; assuming that the trend is linear. Work should be carried out testing the accuracy of the covariate parameter (trend element) and its ability to successfully extrapolate to future events.

2. Following the discovery that spatial dependence varies with return period, an assessment of the impact upon techniques that are currently in use by hydrologists within the UK needs to be carried out. This specifically impacts the Netmax element of the FORGEX method for growth curve extension and therefore extreme value estimation. Initially, it would appear that this finding will result in reduced rainfall totals for a given return period. So, the current technique is effectively providing a conservative estimate, meaning greater than the actual value. This work has not been pursued during this thesis as it did not form part of the initial objectives.

3. Another area for future research is the proposed homogeneity test for pooling rainfall sites, chapter 5.8. This method requires careful analysis of the implications surrounding the test; the repetitive process of testing and adding another site carries the potential risk

of accumulating errors. This was not explored further during this thesis due to time constraints, but is worthy of further analysis due to the potential gains.

This page is intentionally blank.

References

Bass K. T., 1975, Selection of design flood - the engineer's dilemma. Inspection, Operation and Improvement of Existing Dams. *Proceedings of BNCOLD Symposium*, University of Newcastle-upon-Tyne, paper 4.1.

Buishand, T.A. and Brandsma, T. 2001. Multi-site simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbour resampling. *Water Resources Research*, 37, 2761-2776.

Binnie & Partners, 1986: Modes of dam failure and flooding and flood damage following dam failure. Contract report to the Department of the Environment, (Contract No PECD 7/7/184).

Brunetti M., et al., 2000, Precipitation intensity trends in northern Italy, *International Journal of Climatology*, 20, 1017-1031.

The Centre for Ecology & Hydrology, 1999, Flood Estimation Handbook, Volume 2
Rainfall frequency estimation

Centre for Ecology & Hydrology, Calver, P., Flood Frequency Quantification for Ungauged Sites using Continuous Simulation: a UK approach

Collier, C.G. and Hardaker, P.J., 1997, Estimating probable maximum precipitation using a storm model approach, *Journal of Hydrology*, 183, 277-306.

Cooperative studies section, hydrologic services division, 1960. Generalised estimates of probable maximum precipitation for the U.S. Technical paper 38, Weather Bureau, U.S. Department of Commerce, Washington D.C.

Crichton, D., 2002, UK and Global Insurance Responses to Flood Hazard, *Water International*, 27, 119-131.

Dales, M.Y. and Reed, D.W., 1989, Regional flood and storm hazard assessment. *Report No. 102, Institute of Hydrology*, Wallingford, UK, 159pp.

Ekström, M., Fowler, H.J., Kilsby, C.G. and Jones, P.D. 2005. New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies. *Journal of Hydrology*, 300, 234-251.

Environment Agency, Flood and Coastal Risk Management Conference 2010, Telford, UK, 29 June - 1 July 2010. (Unpublished)

Fowler, H.J. and Kilsby, C.G., 2003a, A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000, *International Journal of Climatology*, 23: 1313–1334.

Fowler, H.J. and Kilsby, C.G., 2003b, Implications of changes in seasonal and annual extreme rainfall. *Geophysical Research Letters*, 30 pp. 1720, doi:10.1029/2003GL017327.

Fowler H.J., Ekström M., Kilsby C.G. and Jones P.D., [New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate.](#) *Journal of Hydrology* 2005, 300(1-4), 212-233.

Fowler, H.J., Ekström, M., Blenkinsop, S. and Smith, A.P. 2007. Estimating change in extreme European precipitation using a multimodel ensemble. *Journal of Geophysical Research*, **112**, D18104.

Frei, C., and C. Schär, 2001, Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *Journal of Climatology*, 14, 1568-1584

Groisman, P.Y., Karl, T.R., Easterling, D.R. et al. 1999. Changes in the probability of heavy precipitation: Important indicators of climate change. *Climatic Change*, 42, 243-283.

Gruner, E., 1963: Dam disasters. Proc. ICE, Paper No. 6648.

Gustard, A., Roald, L., Demuth, S., Lumadjeng, H. and Gross, R., 1989, Flow Regimes from Experimental and Network Data (FRIEND) 2 Vols. *Institute of Hydrology*, Wallingford, Oxon, UK

Hart, T.L., 1982, Survey of probable maximum precipitation using the synoptic method of storm transposition and maximization. Workshop on Spillway Design, *AWRC Conf. Ser., Vol. 6*, Aust. Gov. Print. Serv., Canberra, Australia.

Haylock, M.R., Cawley, G.C., Harpham, C., Wilby, R.L. and Goodess, C.M. 2006. Downscaling heavy precipitation over the UK: a comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology*, **26**, 1397-1415.

Hegerl, G.C., Zwiers, F.W., Stott, P.A. and Kharin, V.V. 2004. Detectability of anthropogenic changes in annual temperature and precipitation extremes. *Journal of Climate*, **17**, 3683-3700.

Hosking, J. R. M., and Clarke, R. T., 1990, “Rainfall-Runoff Relations Derived from the Probability Theory of Storage,” *Water Resources Research*, 26, 1455–1463.

Hosking, J. R. M., and Wallis, J. R., 1988, "The Effect of Inter-Site Dependence on Regional Flood Frequency Analysis," *Water Resources Research*, 24, p588–600.

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, p105–124.

Hughes, A., et al, 2000, *Risk Management for UK Reservoirs (C542)*, CIRIA

Institute of Hydrology's 1999, *Flood Estimation Handbook (Volumes: 1 and 2)*

Iwashima, T., and Yamamoto, R., 1993, 'A statistical analysis of the extreme events: Long-term trend of heavy daily precipitation.' *J. Met. Soc. Japan*, 71, 637-640.

Jones, P.D. and P.A. Reid, 2001, Assessing future changes in extreme precipitation over Britain using Regional Climate Model integrations, *International Journal of Climatology*, 21, 1337–1356

Karl, T.R. and R.W. Knight, 1998, Secular trends of precipitation amount, frequency, and intensity in the United States., *Bulletin of the American Meteorological Society* 79 (2), 231-241.

Kjeldsen, Thomas R.; Jones and David A., 2006, Prediction uncertainty in a median-based index flood method using L moments. *Water Resources Research*, 42, W07414. 10.1029/2005WR004069

Leisch, F., 2003, *The R Project for Statistical Computing*. In: www.r-project.org/ - for information on the R Software package and programming language [Accessed Dec 2007]

Loukola, E., Kuusisto, E. and Reiter, P., 1985, The Finnish approach to dam safety, *Water Power & Dam Construction*, 37, 22-24.

Maraun, D., Osborn, T.J. and Gillett, N.P. 2008. UK daily precipitation intensity: improved early data, error estimates and update from 2000 to 2006. *International Journal of Climatology*, doi: 10.1002/joc.1672

Matalas, N.C., 1967, Mathematical assessment of synthetic hydrology, *Water Resour Res* 3:937 - 945

Milly, P.C.D., Betancourt, J., Falkenmark, M., et al.. 2008. Stationarity is dead: Whither water management? *Science*, 319, 573-574.

Moffat, A. I. B., 1982, Dam deterioration - a British perspective. *Proceedings of BNCOLD Conference*, Keele, pp 103-115. (This paper was not bound in the original volume.)

Natural Environment Research Council (NERC), 1975, *Flood Studies Report (FSR)*, Vol.11, London.

Osborn, T.J. and Hulme, M. 1997. Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *Journal of Climate*, **10**, 1855-1908.

Osborn, T.J., Hulme, M., Jones, P.D., and Basnett, T.A., 2000, Observed trends in the daily intensity of United Kingdom precipitation, *International Journal of Climatology* **20**, 347-364

Palmer, T.N., and Räisänen, J., 2002, Quantifying the risk of extreme seasonal precipitation in a changing climate, *Nature*, **415**, 512-514

Purcell, S., 2007, *Maximum Likelihood Estimation (MLE)*, In: http://statgen.iop.kcl.ac.uk/bgim/mle/sslike_3.html, [Accessed June 2007]

Reed, D.W., Faulkner, D.S. and Stewart, E.J., 1999, The FORGEX method of rainfall growth estimation, II: Description. *Hydrol. Earth System Sci.* **3**, 197-203.

Reed, D. W. and Field, E. K., 1992, Reservoir flood estimation: another look. Wallingford, *Institute of Hydrology*, 93pp. (IH Report no.114)

Reed D.W. and Stewart, E.J., 1994, Inter-site and inter-duration dependence in rainfall extremes. Chapter 8 in: *Statistics for the Environment 2: Water related Issues*, Wiley, Chichester, UK, 125-143.

Reservoirs Safety Provisions Act 1930

Reservoirs Act 1975

Sælthun, N. R. and Andersen, J. H., 1986, New Procedures for Flood Estimation in Norway, Paper presented at the *Nordic Hydrological Conference* (Reykjavik, Iceland, August – 1986)

Telford, T., 1994, *Reservoir Safety and the Environment*, London, 260,

Telford, T., 1996, *Floods and reservoir safety: 3rd edition*. Institution of Civil Engineers, 63.

Vedin, H. and Eriksson, B. 1988. Extrem arealneder-bord i Sverige: 1881-1988 (Extreme areal precipitation in Sweden: 1881-1988). Report No. 76, *Swedish Meteorol. & Hydrol. Institute*, Norrköping.

Vassdragsdirektorat 1986. Beregning av dimensjon-Hydrol., **17**, 217-228. erende og paregnelig maksimal flomar, etningslinjer (*Guidelines for the design estimation of*

probable maximum flood), V-informasjon nr. 1, Norges Vassdrags og Elektrisitetsvesen, Oslo.

Wiesner, C. J., 1970, *Hydrometeorology*, Chapman and Hall Ltd, London

Wigley, T. M. L., Lough, J. M. and Jones, P. D., 1984. Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series. *International Journal of Climatology*, 4, 1-25.

Wilby, R.L., Tomlinson, O.J. and Dawson, C.W. 2003. Multi-site simulation of precipitation by conditional resampling. *Climate Research*, **23**, 183-194.

Wilson, P.S. and Toumi, R. 2005. A fundamental probability distribution for heavy rainfall. *Geophysical Research Letters*, 32, L14812.

Wright, C.E., 1994: "UK Reservoir failures and safety legislation". Dams and Reservoirs, October, 4pp.

National Research Council, 1988, *Estimating Probabilities of Extreme Floods—Methods and Recommended Research*, National Academy Press, Washington, D.C., 141 pp

WMO,1986: Manual for estimation of probable maximum precipitation. *Operational hydrology*, Rep.1, WMO-No.332, 269 pp.

Zhai P. M., A. Sun, F. Ren, X. Liu, B. Gao, and Q. Zhang, 1999: Changes of climate extremes in China. *Climatic Change*, 42, 203-218

This page is intentionally blank.

Appendices

This page is intentionally blank.

Table of Appendices:

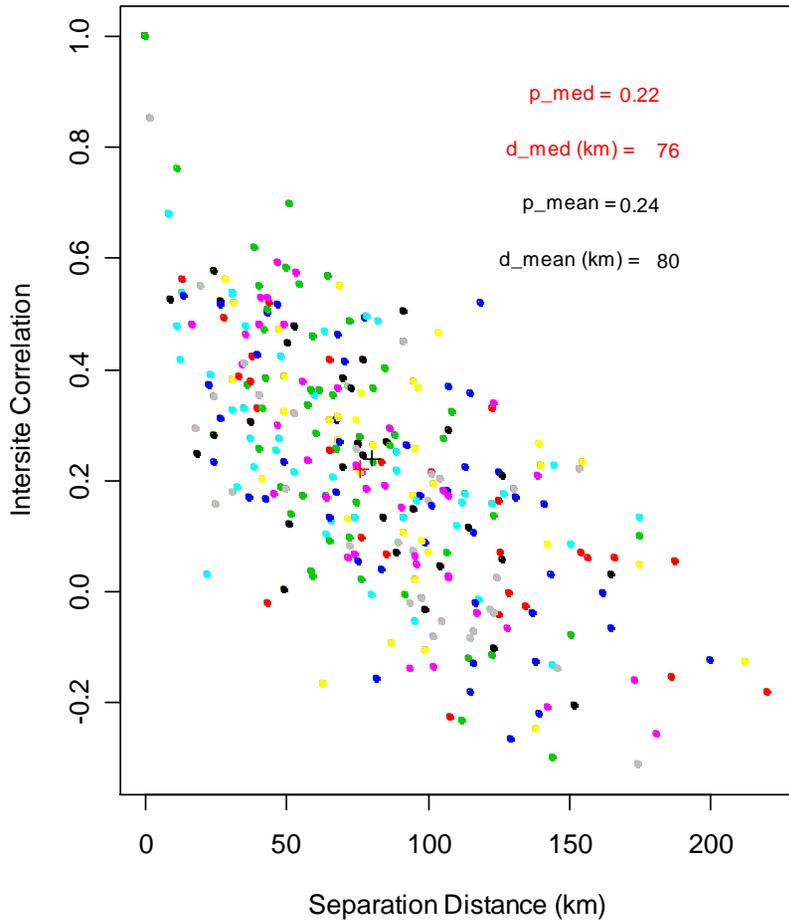
Appendix 1 – Regional Correlation and Effective Number of Sites, Results	148
Appendix 2 - Regional Homogeneity Results	183
Appendix 3 – Sample of ‘R’ routine for Homogeneity testing	194
Appendix 4 - Statistical Terminology and Concepts	198
1. Basics of the technique	199
2. Applications	199
3. Uses	199
4. Laws of probability	200
5. Terminology and concepts	200
6. Statistical parameters	201
Appendix 5 – Complete routines for use within ‘R’	204
A5.1 - Introduction	205
A5.2 – Explanation of Routine 1	205
A5.3 - Routine 1	207

This page is intentionally blank.

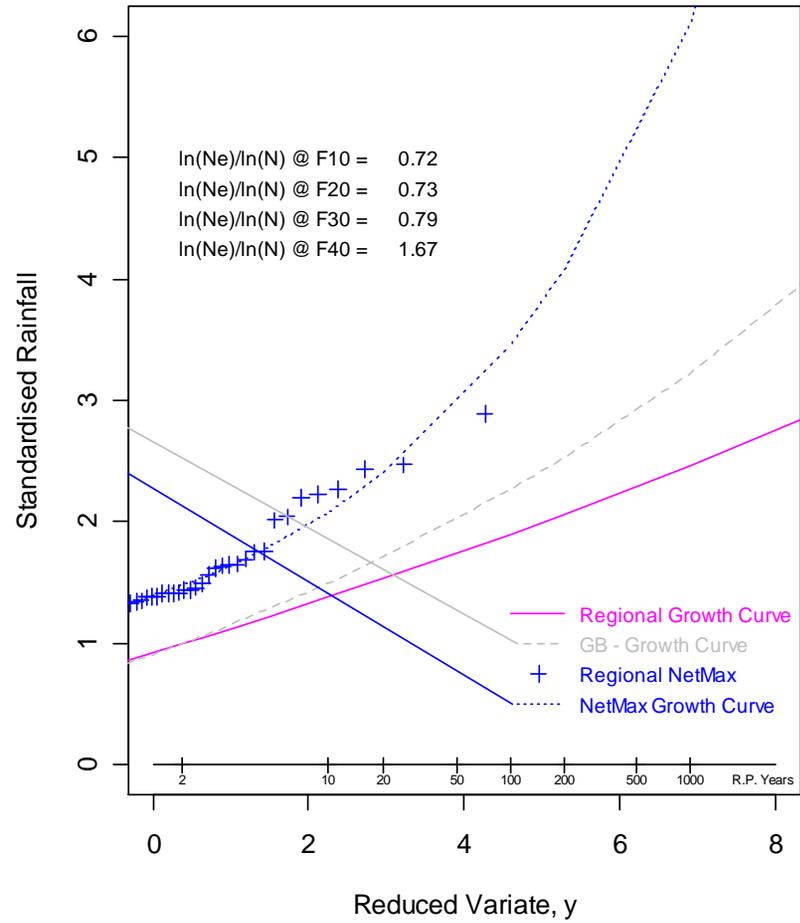
**Appendix 1 – Regional Correlation and Effective Number of Sites,
Results**

This page is intentionally blank.

SS - Distance Correlation Relationship

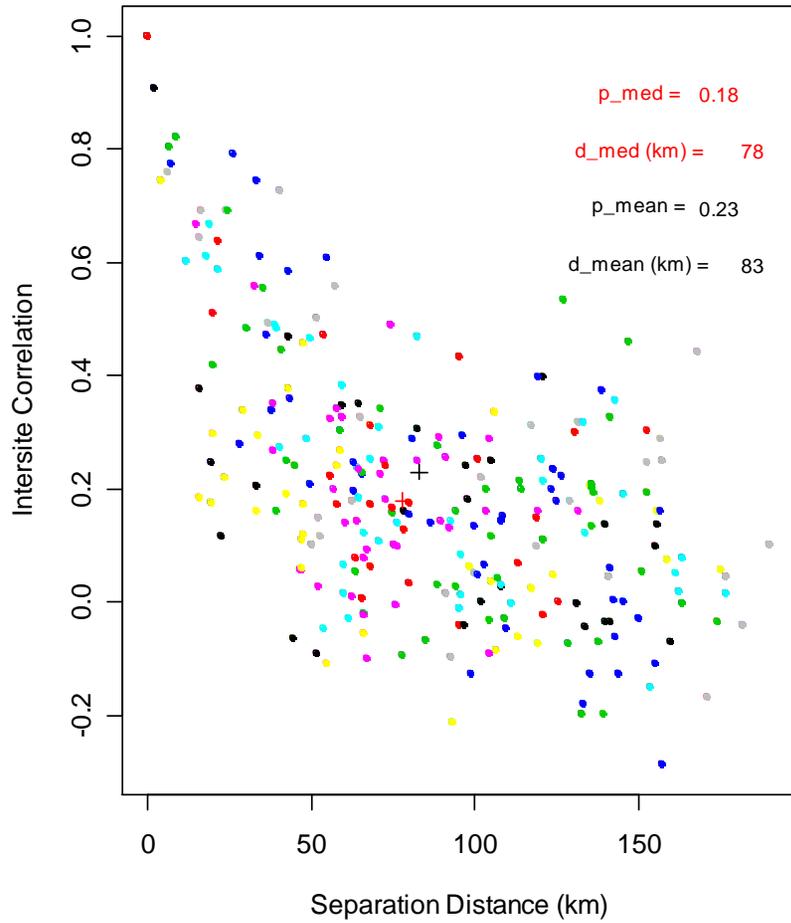


SS - Return Level Plot

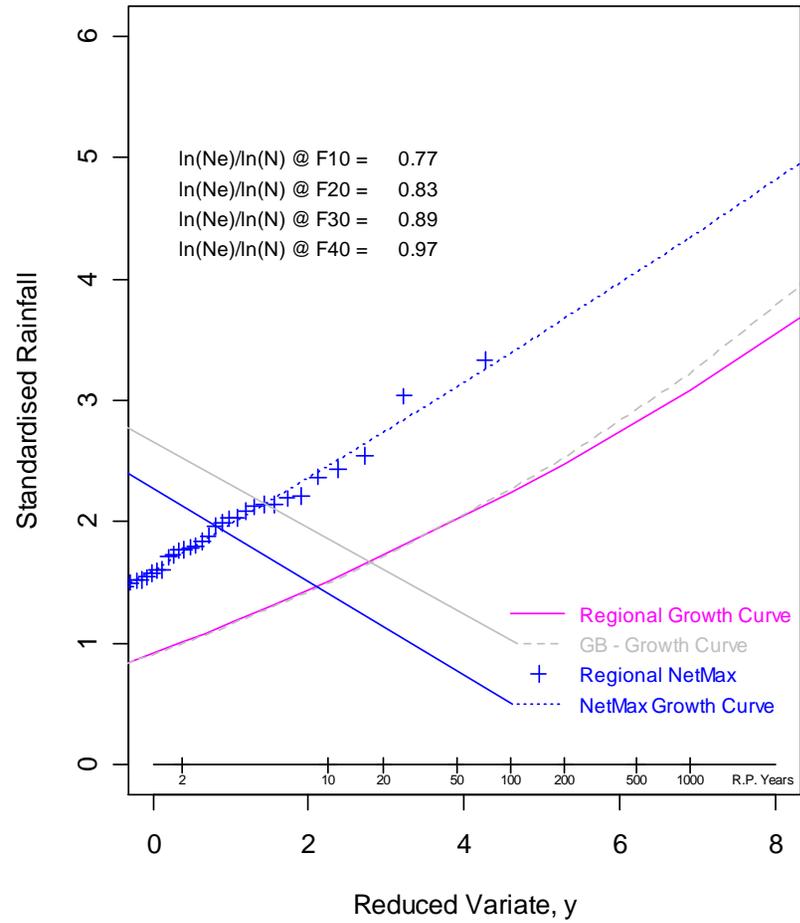


1 Day Annual Maxima

ES - Distance Correlation Relationship

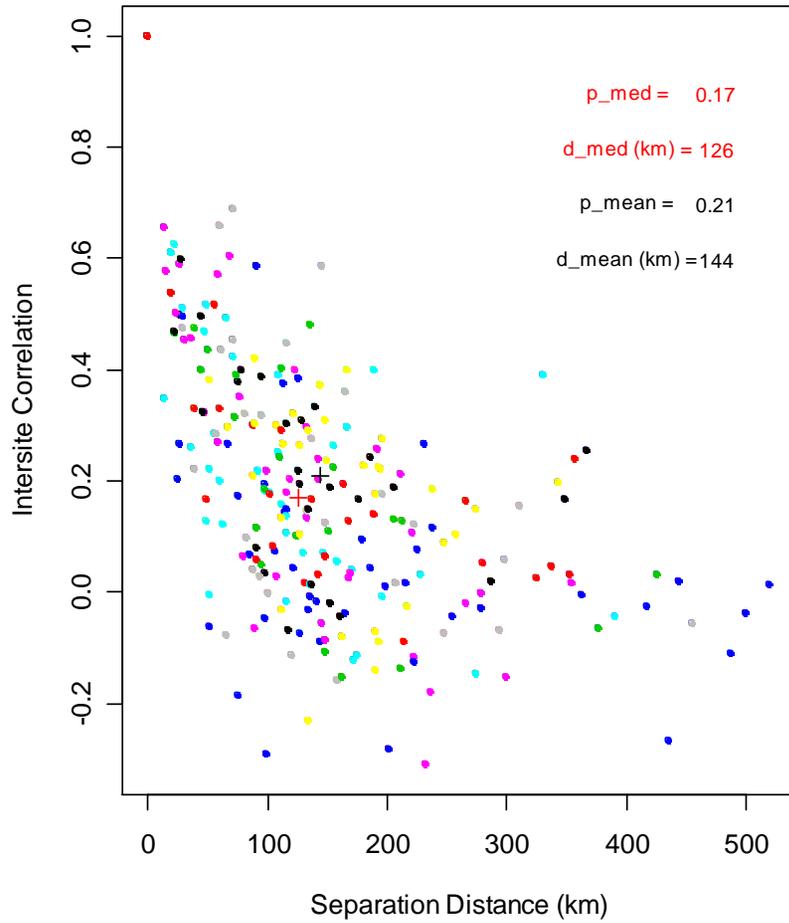


ES - Return Level Plot

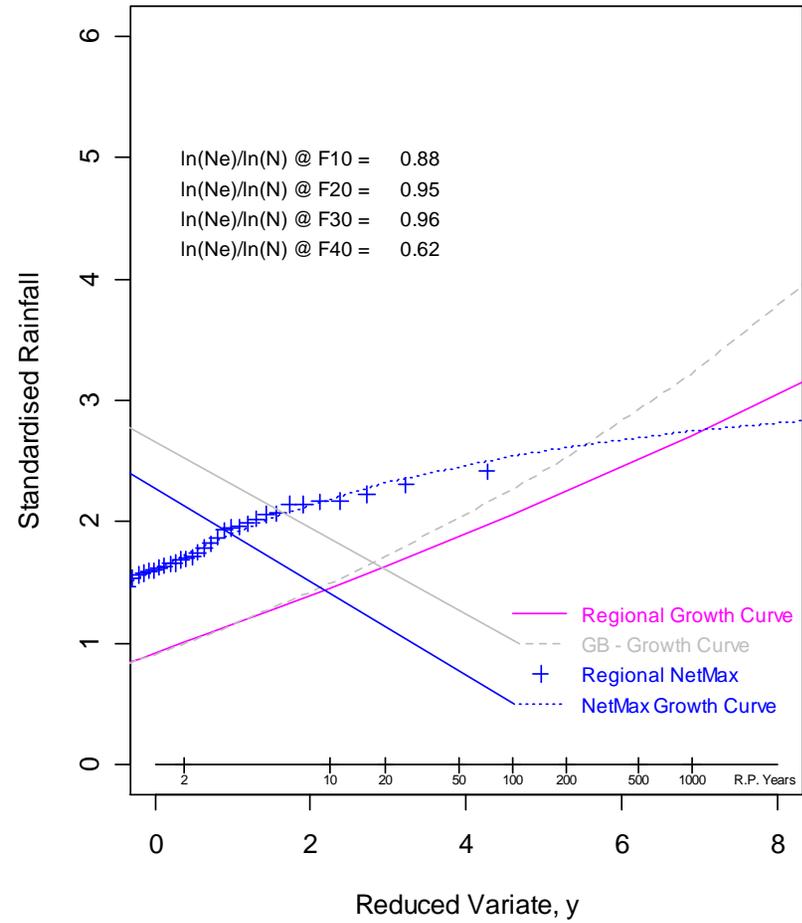


1 Day Annual Maxima

NS - Distance Correlation Relationship

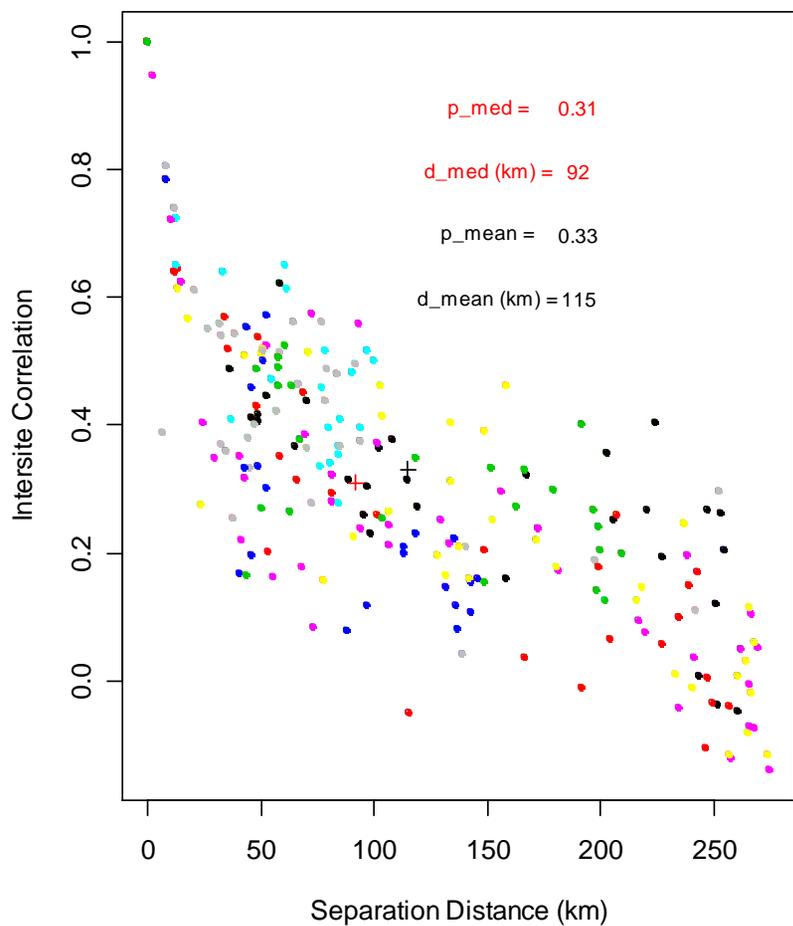


NS - Return Level Plot

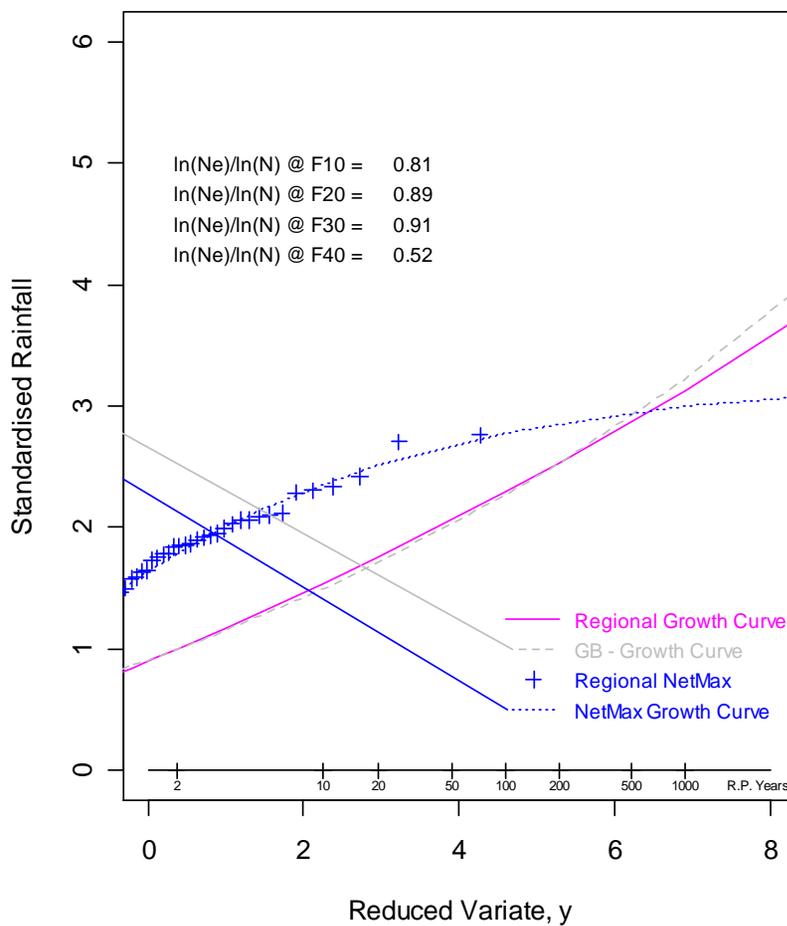


1 Day Annual Maxima

NEE - Distance Correlation Relationship

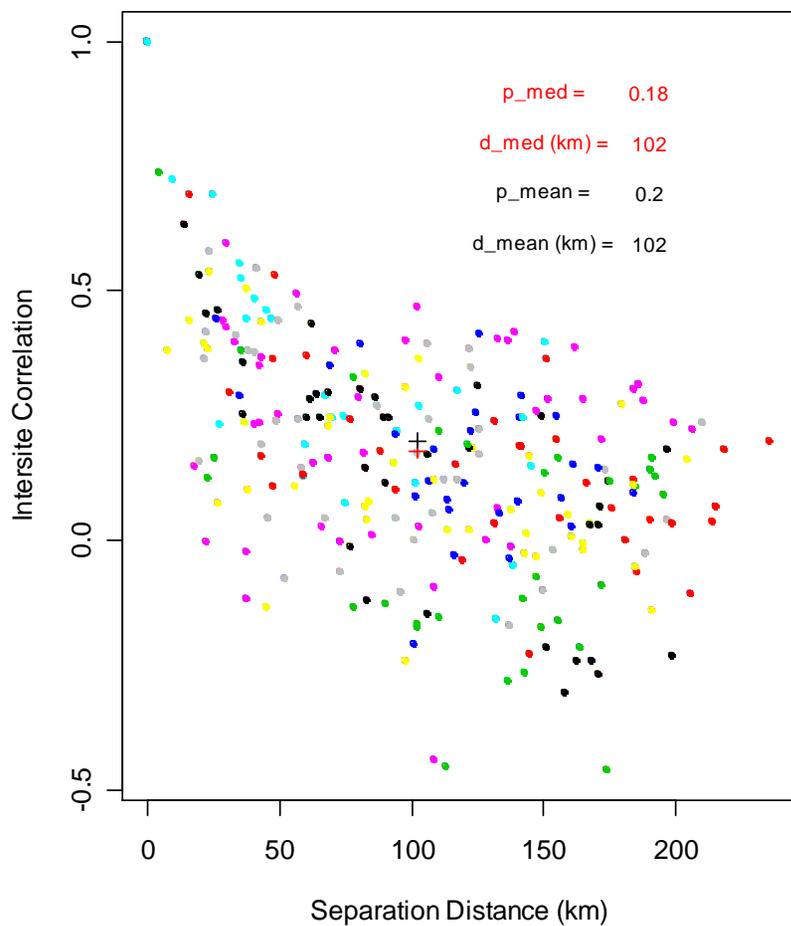


NEE - Return Level Plot



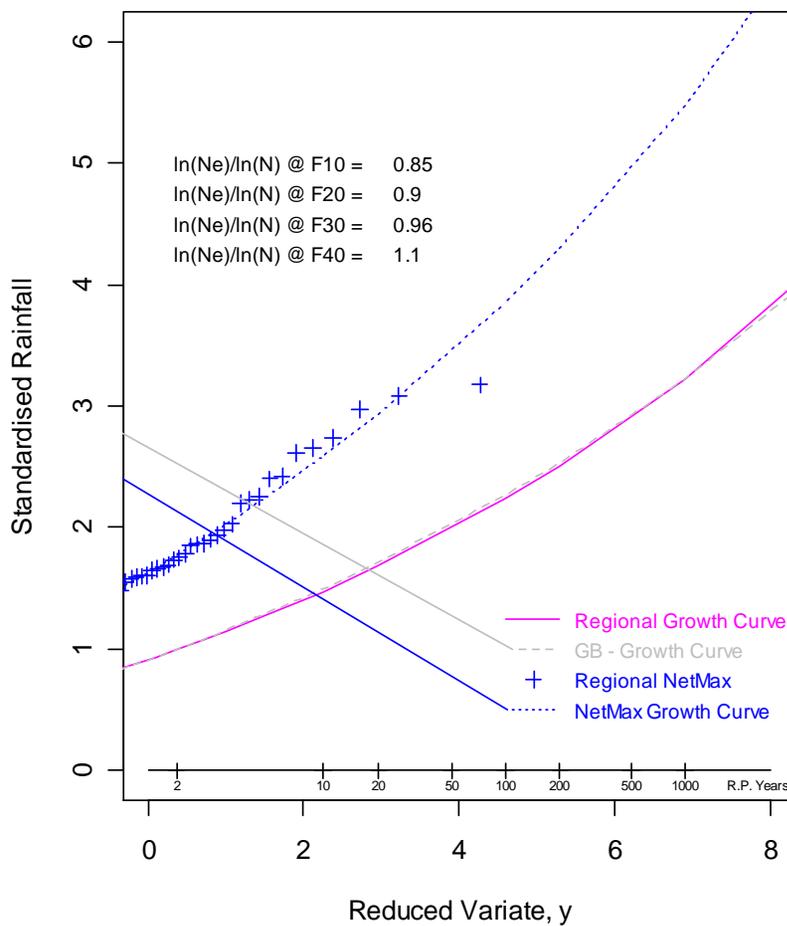
1 Day Annual Maxima

NWE - Distance Correlation Relationship

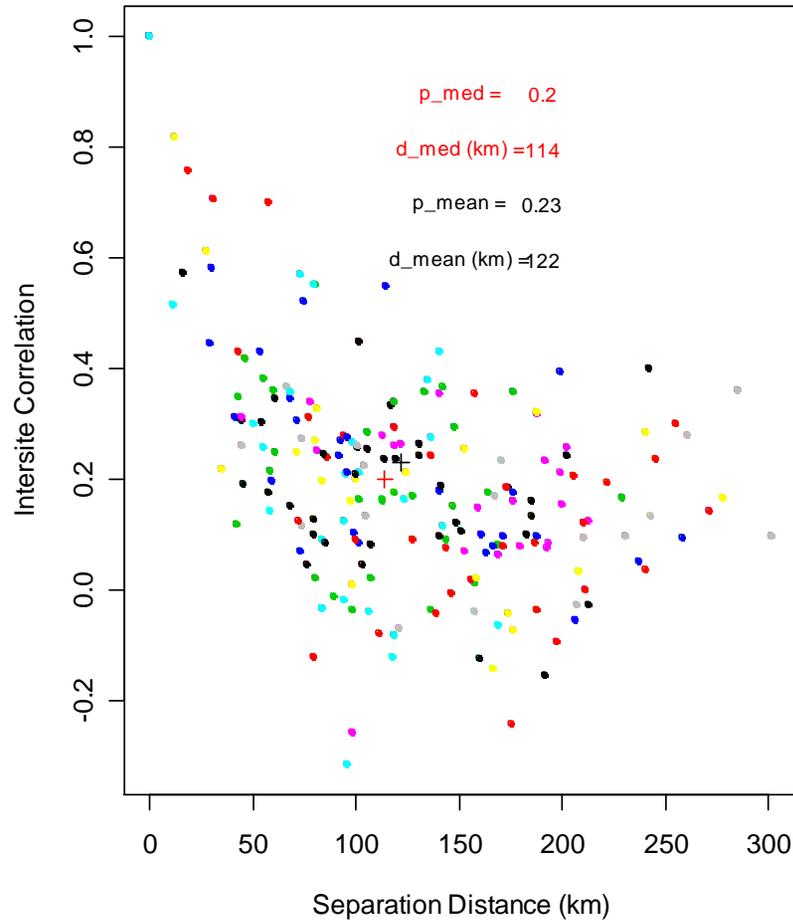


1 Day Annual Maxima

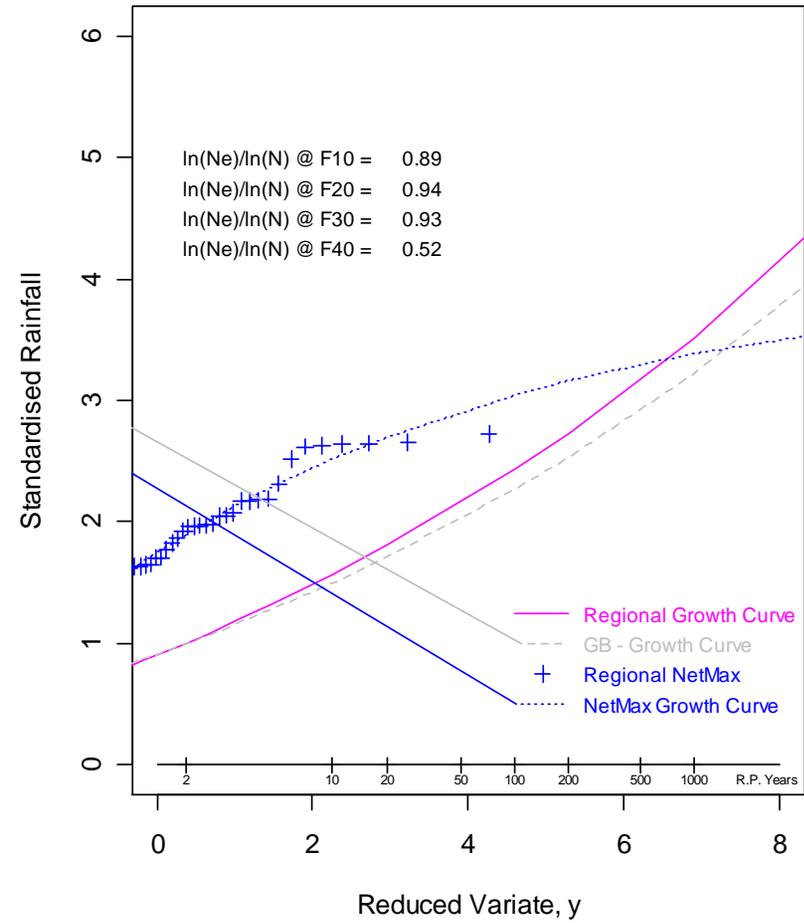
NWE - Return Level Plot



CEE - Distance Correlation Relationship

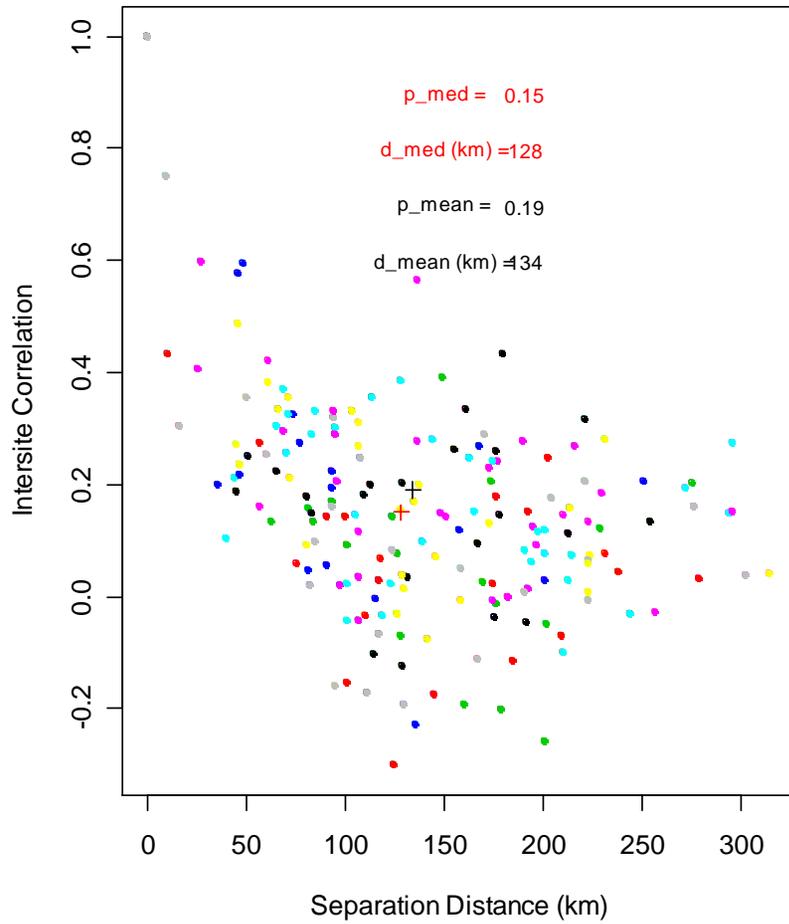


CEE - Return Level Plot



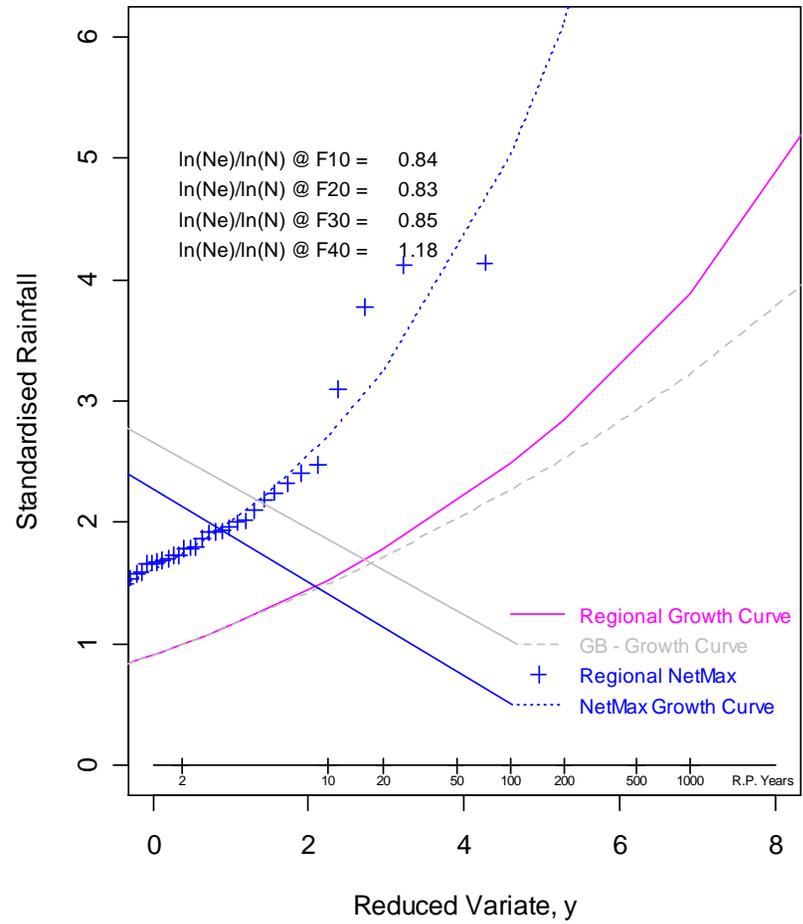
1 Day Annual Maxima

SWE - Distance Correlation Relationship

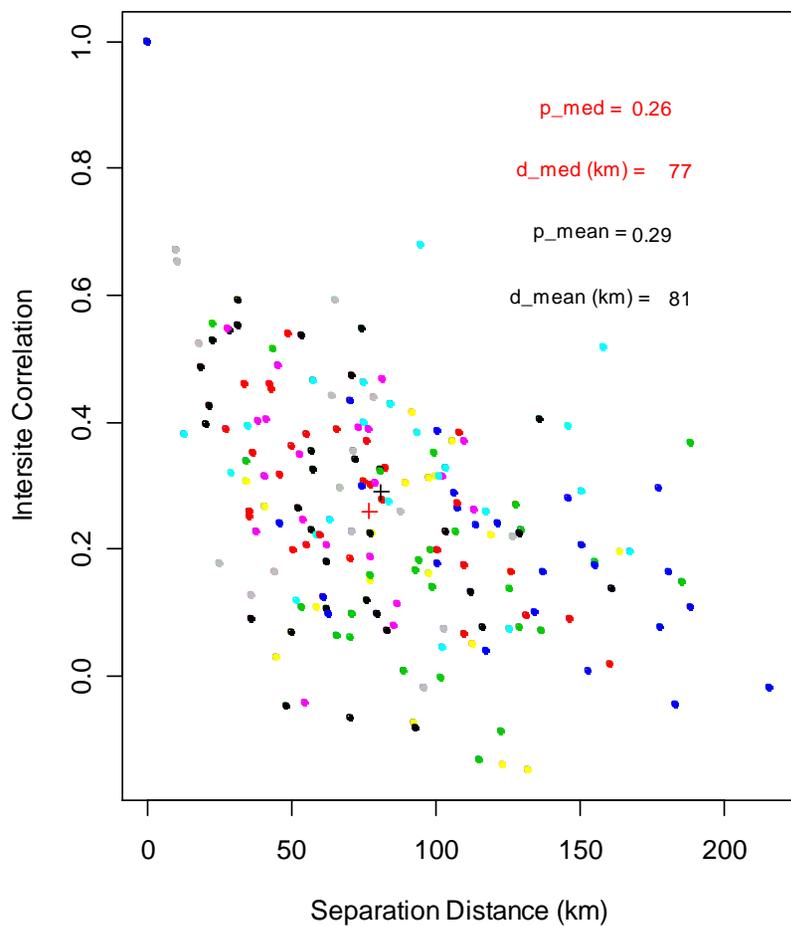


1 Day Annual Maxima

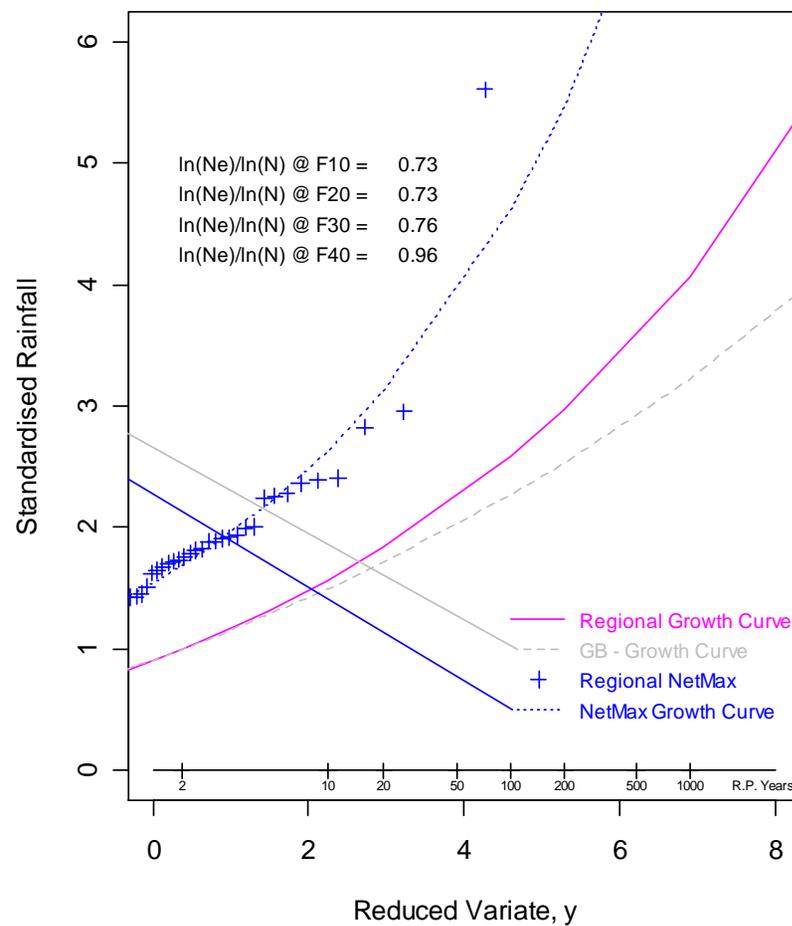
SWE - Return Level Plot



SEE - Distance Correlation Relationship

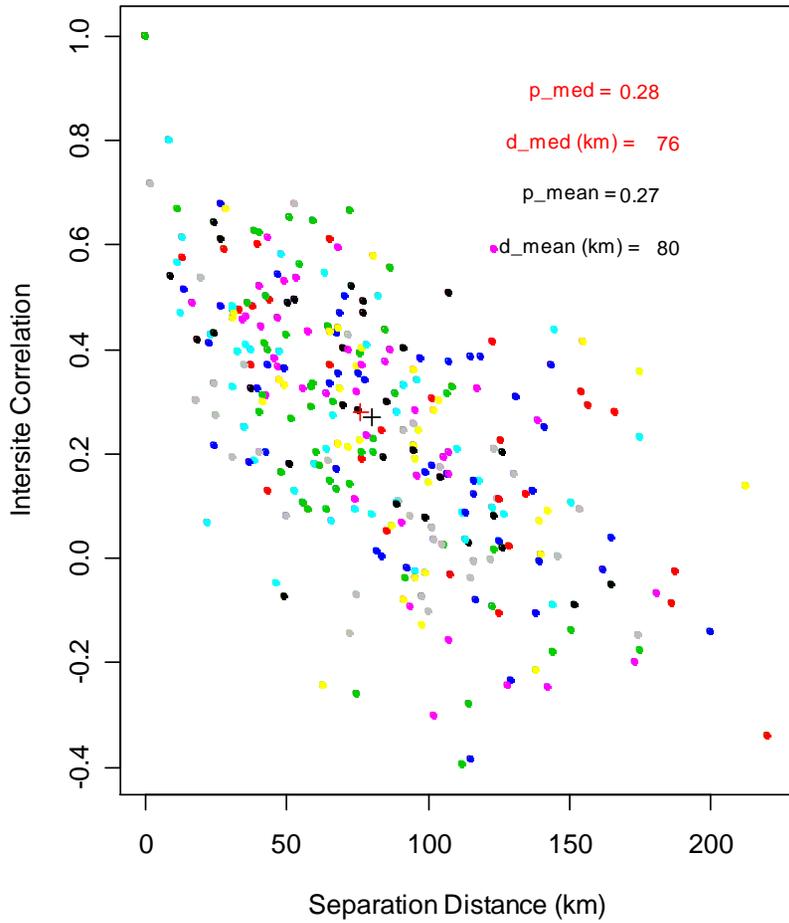


SEE - Return Level Plot



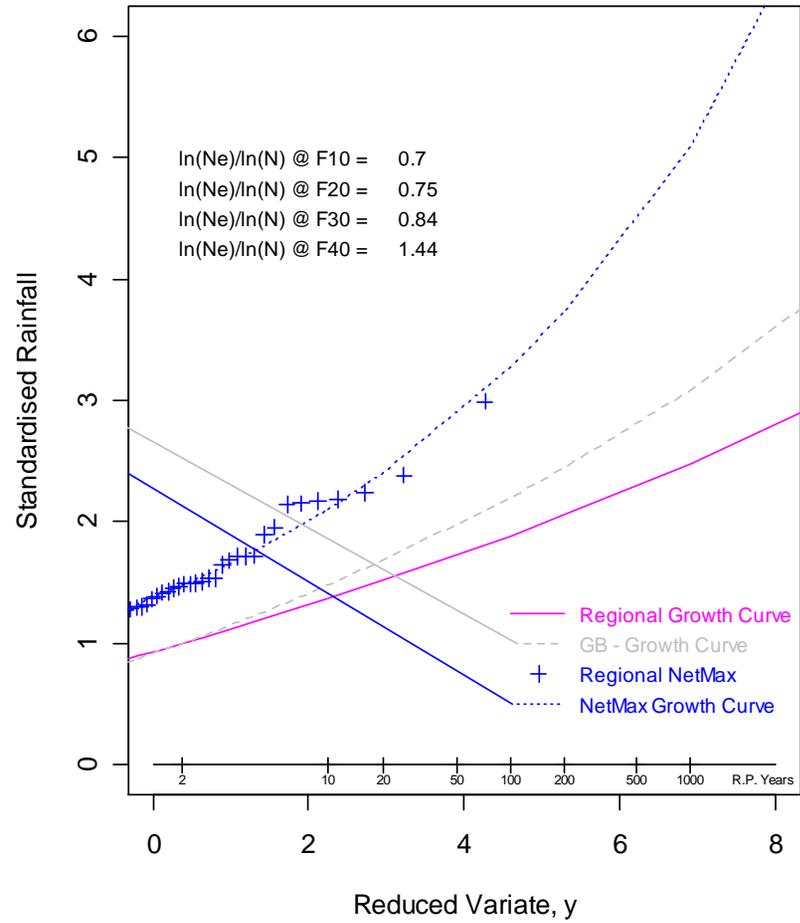
1 Day Annual Maxima

SS - Distance Correlation Relationship

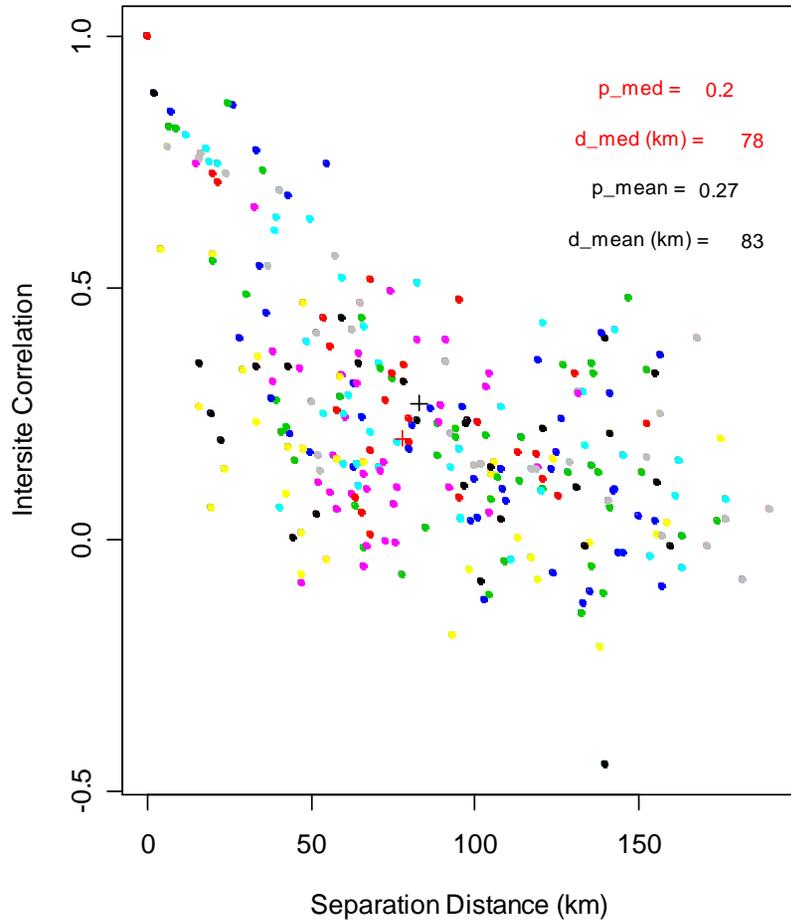


2 Day Annual Maxima

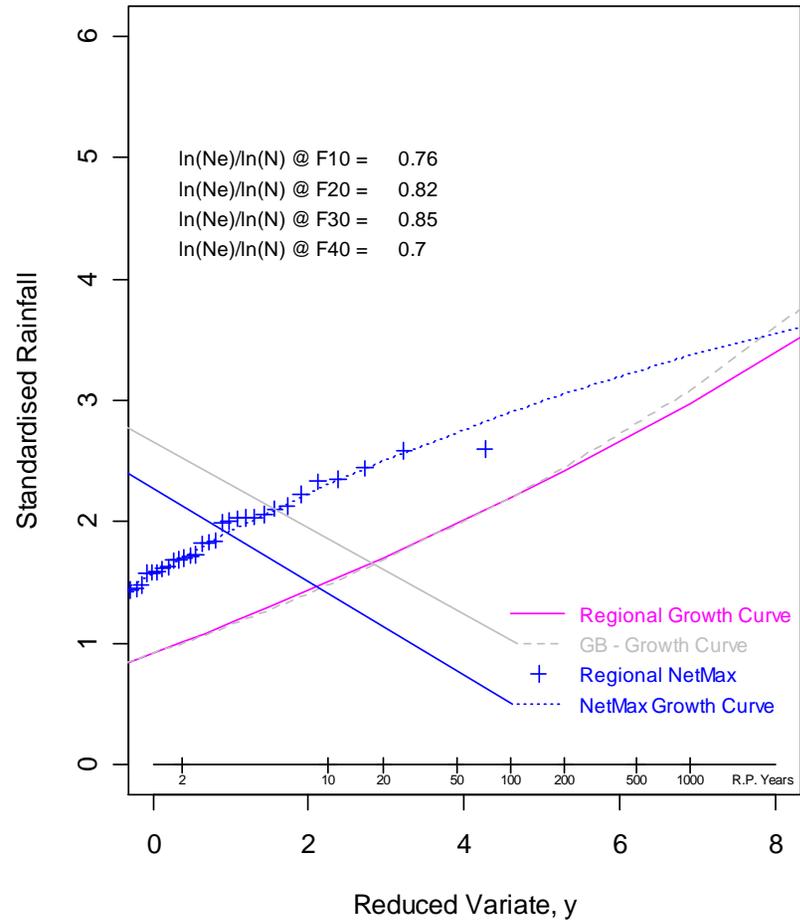
SS - Return Level Plot



ES - Distance Correlation Relationship

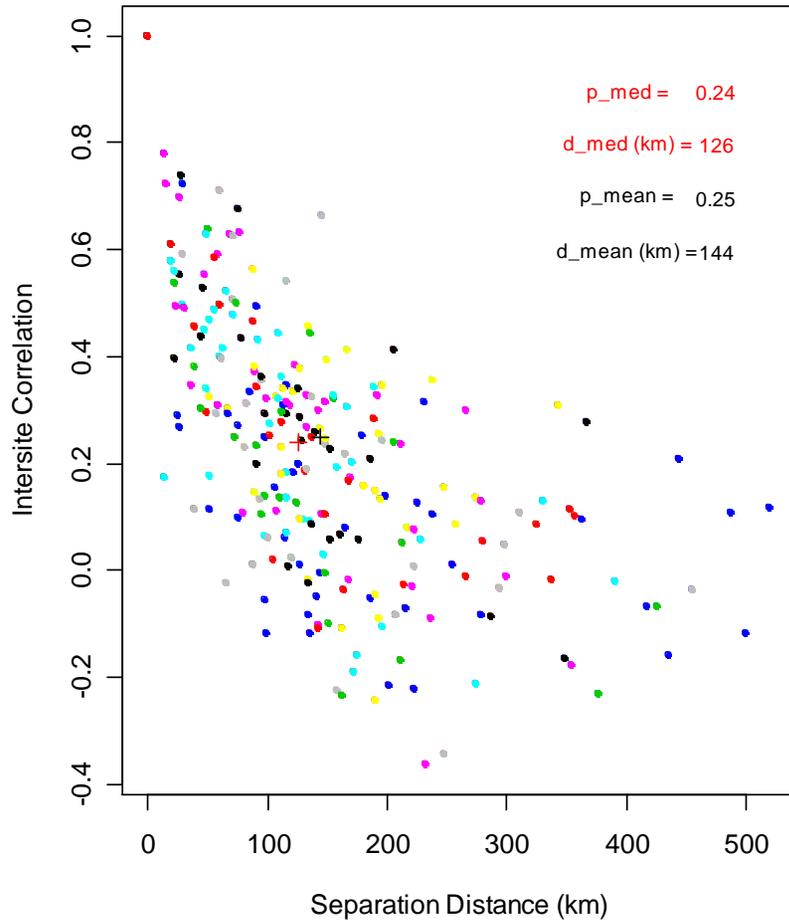


ES - Return Level Plot

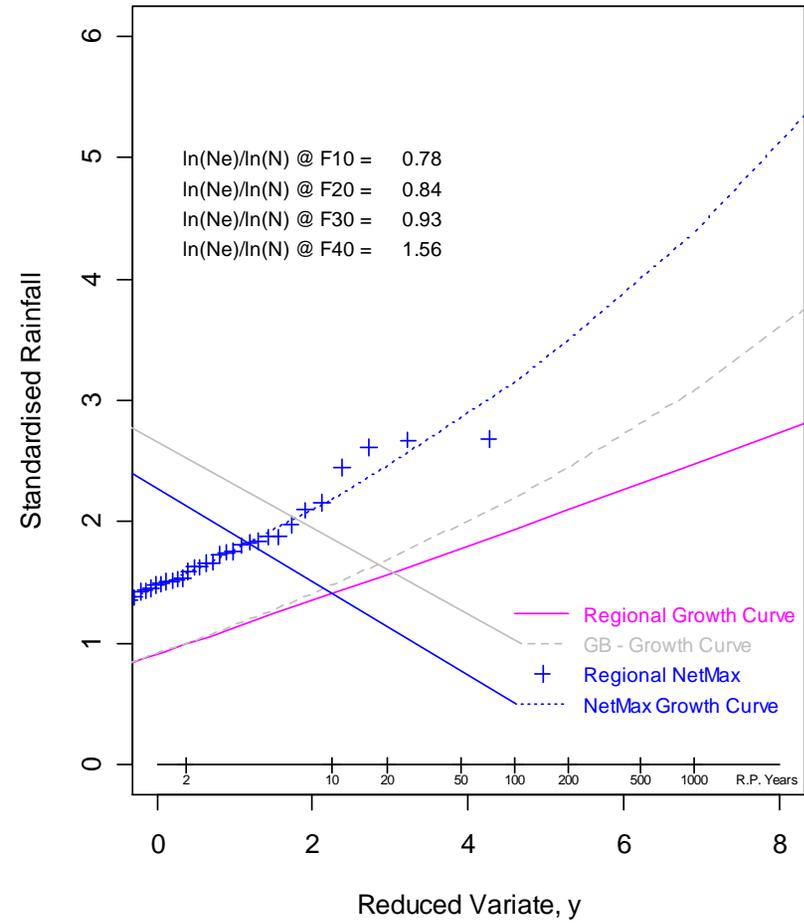


2 Day Annual Maxima

NS - Distance Correlation Relationship

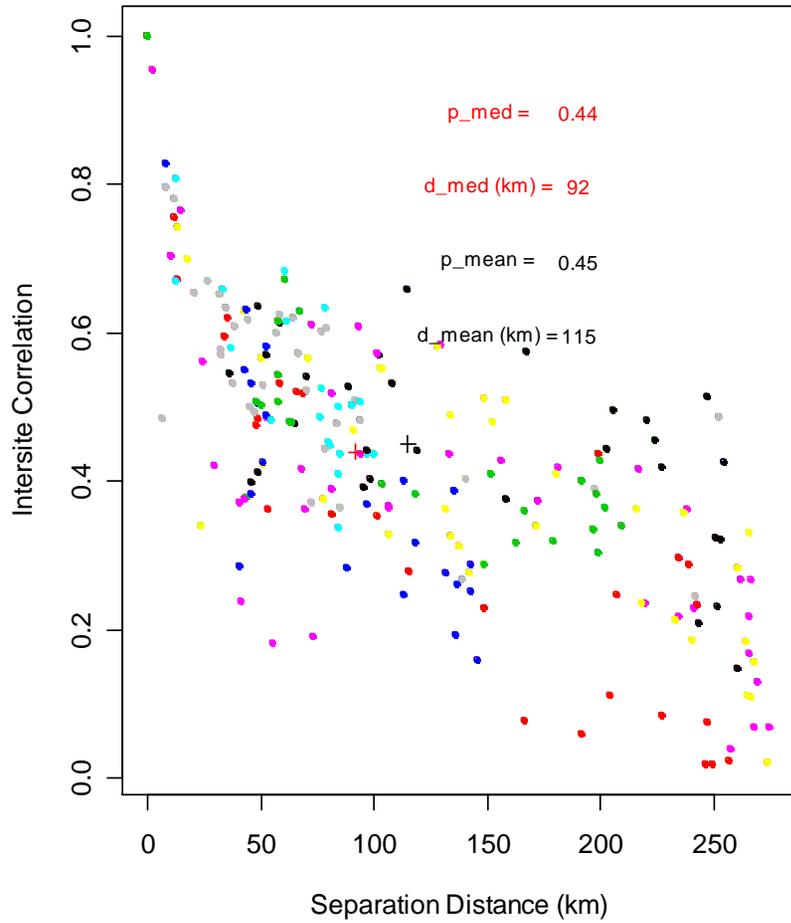


NS - Return Level Plot



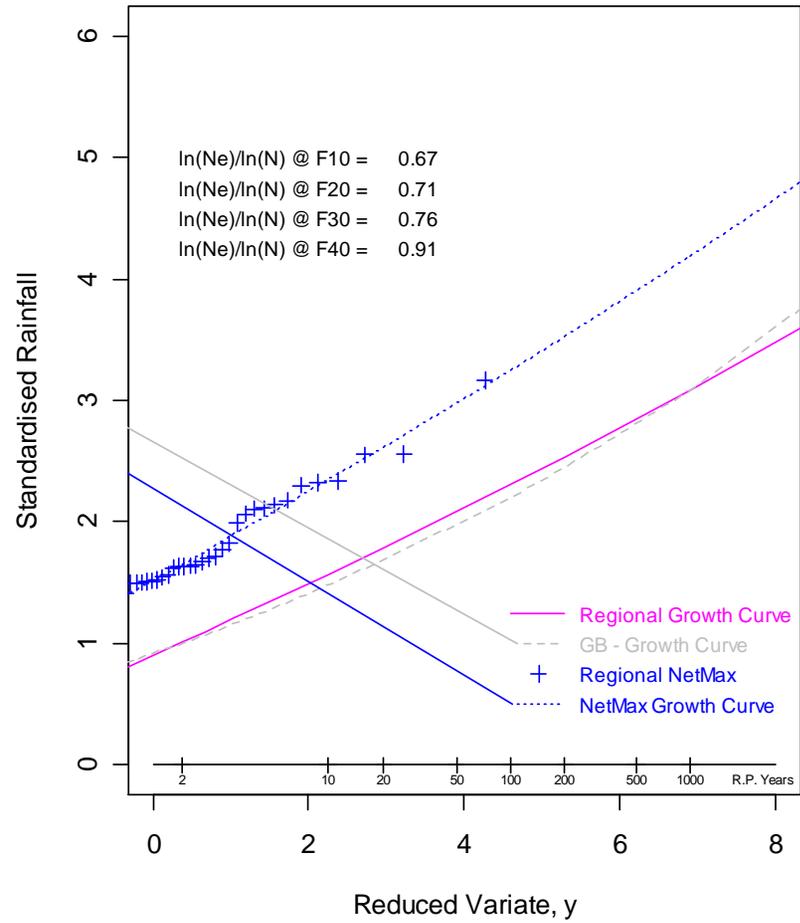
2 Day Annual Maxima

NEE - Distance Correlation Relationship

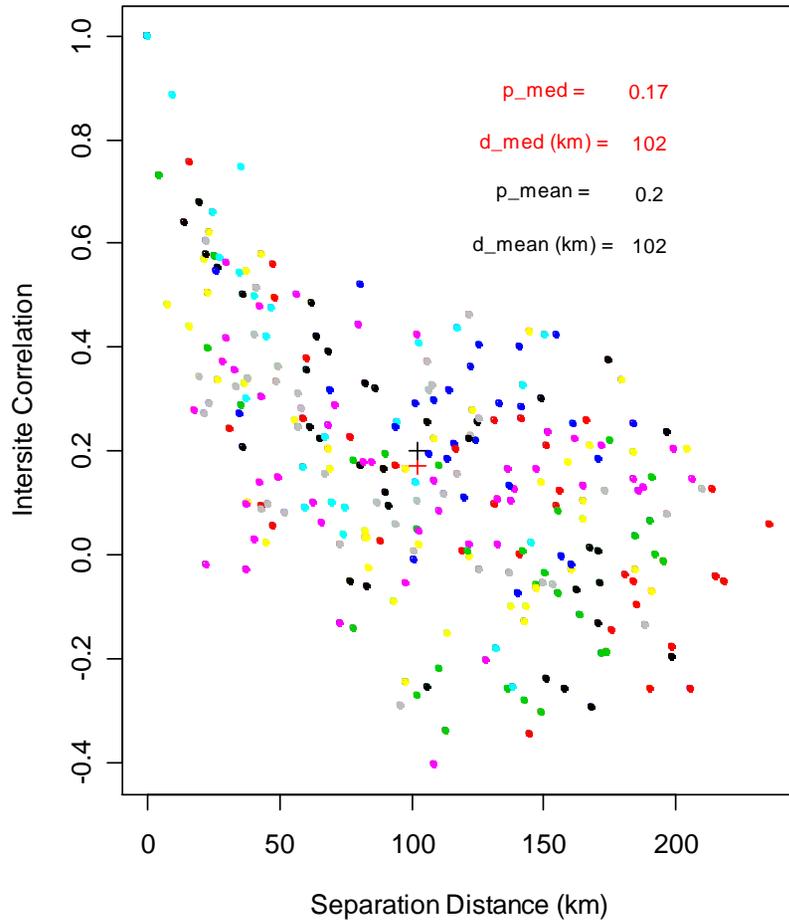


2 Day Annual Maxima

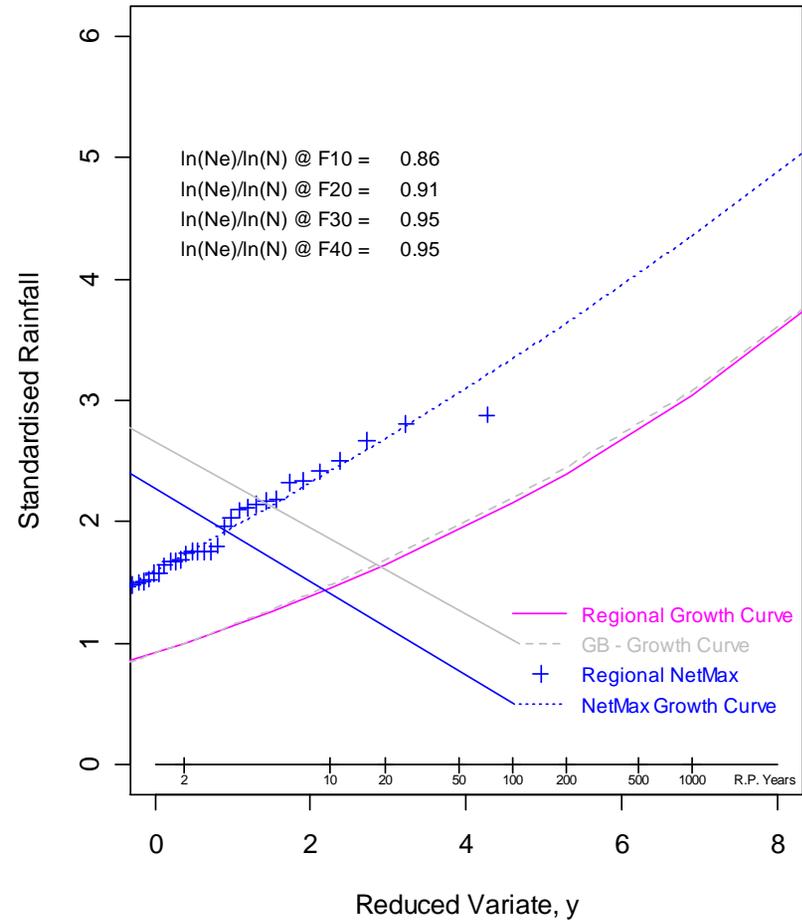
NEE - Return Level Plot



NWE - Distance Correlation Relationship

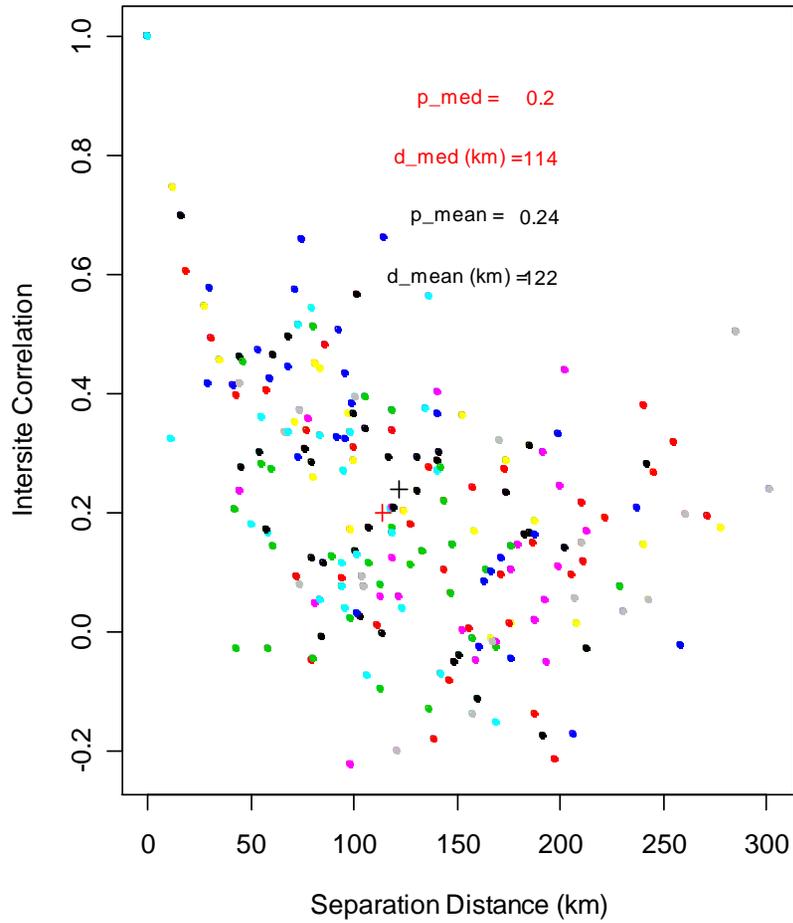


NWE - Return Level Plot



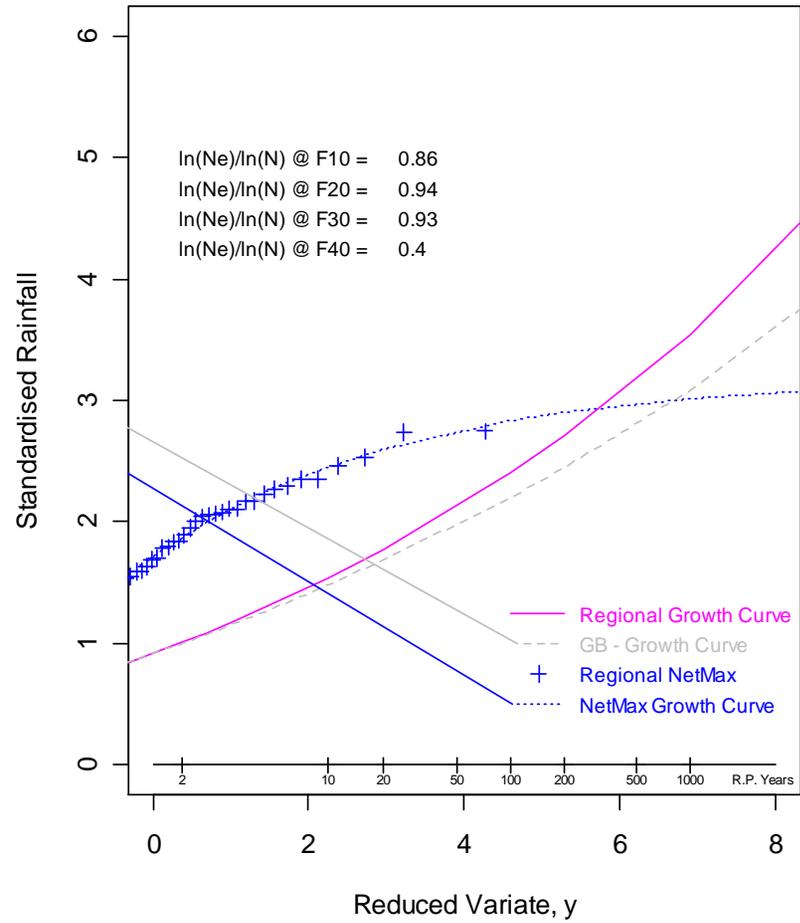
2 Day Annual Maxima

CEE - Distance Correlation Relationship

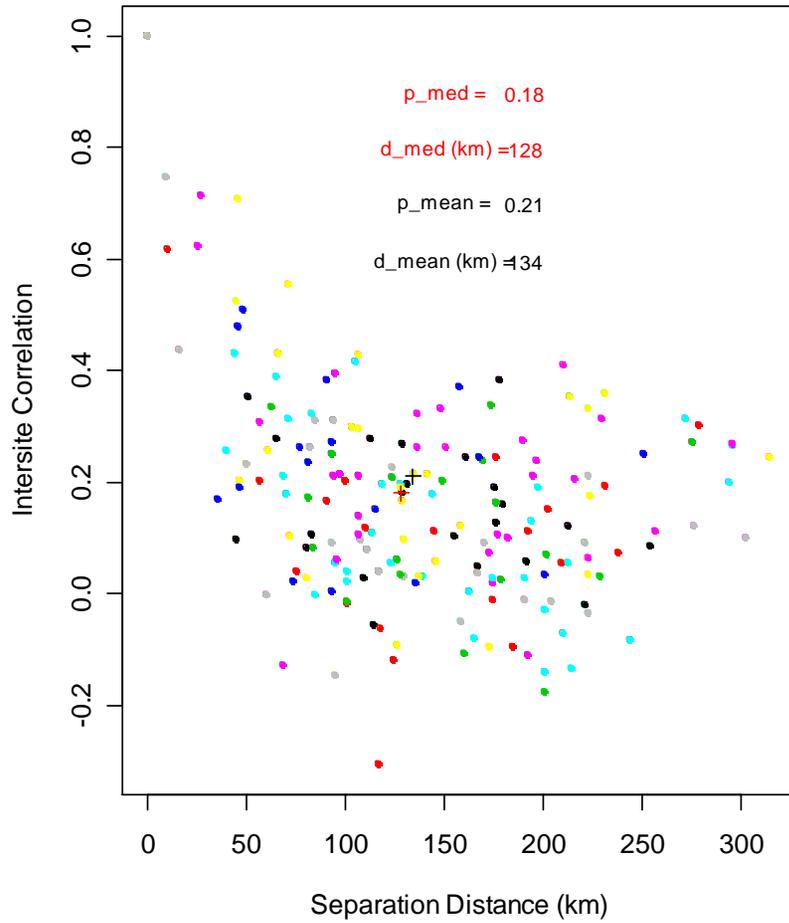


2 Day Annual Maxima

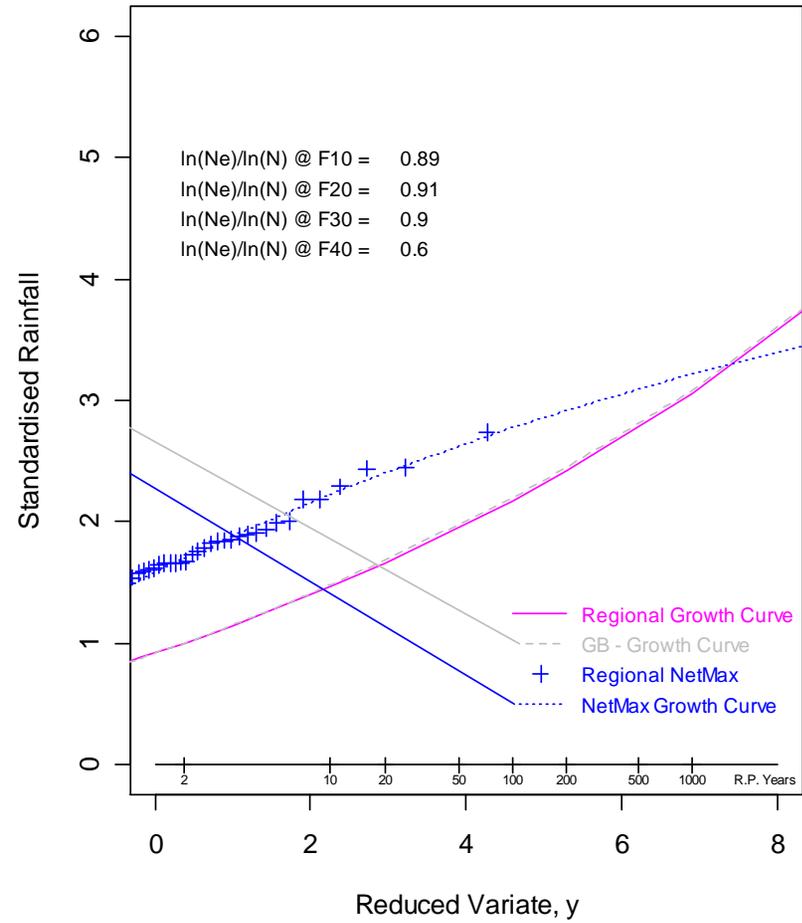
CEE - Return Level Plot



SWE - Distance Correlation Relationship

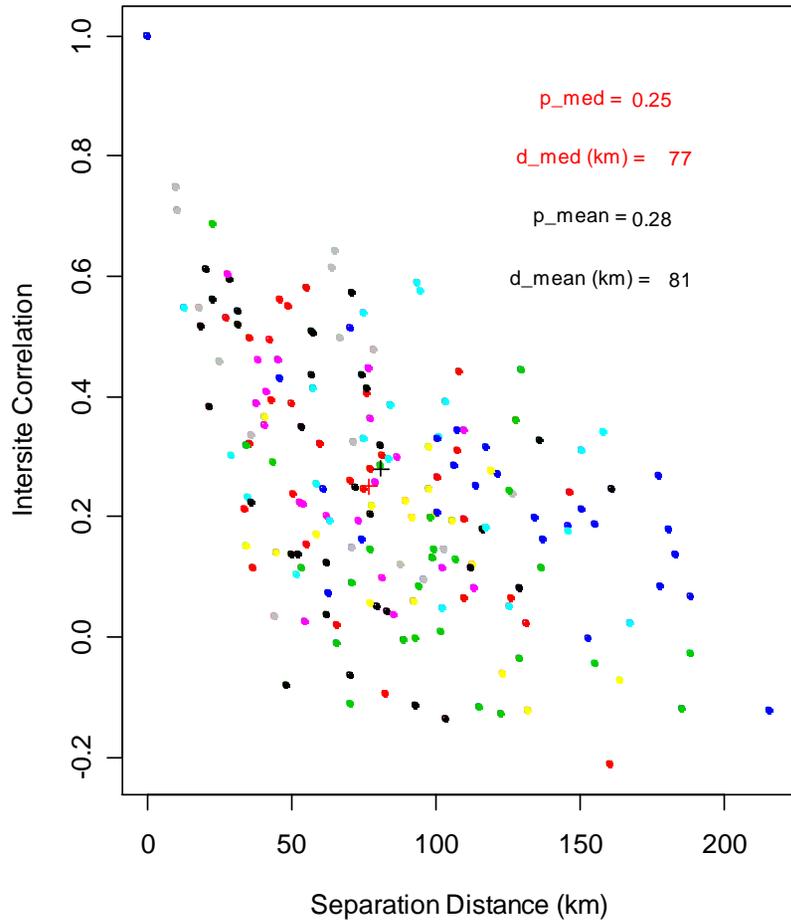


SWE - Return Level Plot

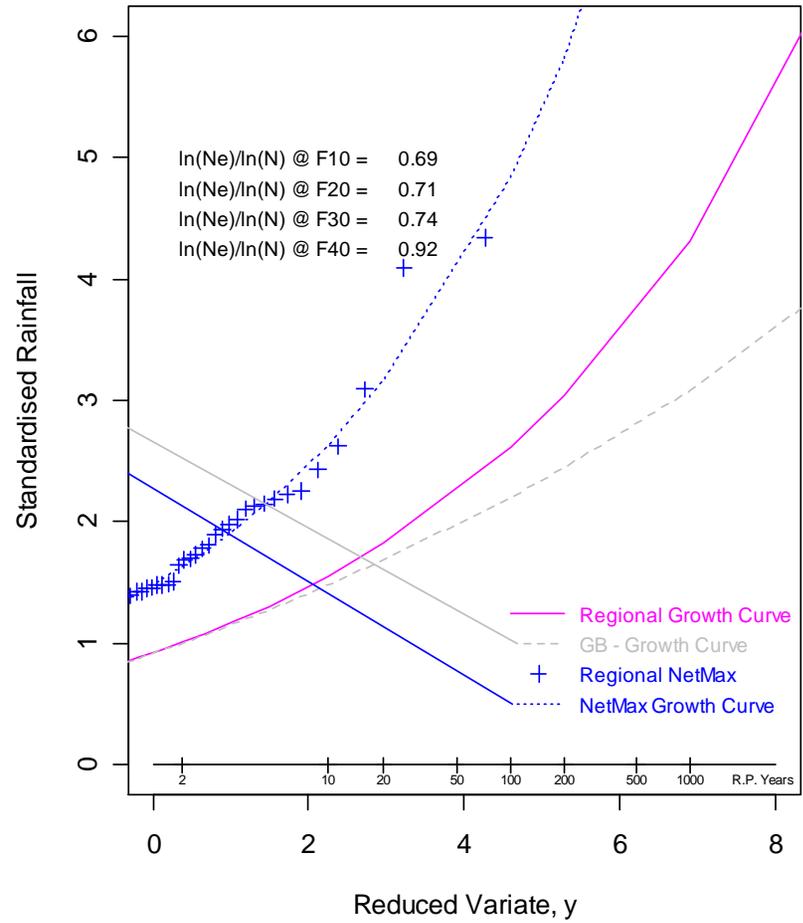


2 Day Annual Maxima

SEE - Distance Correlation Relationship

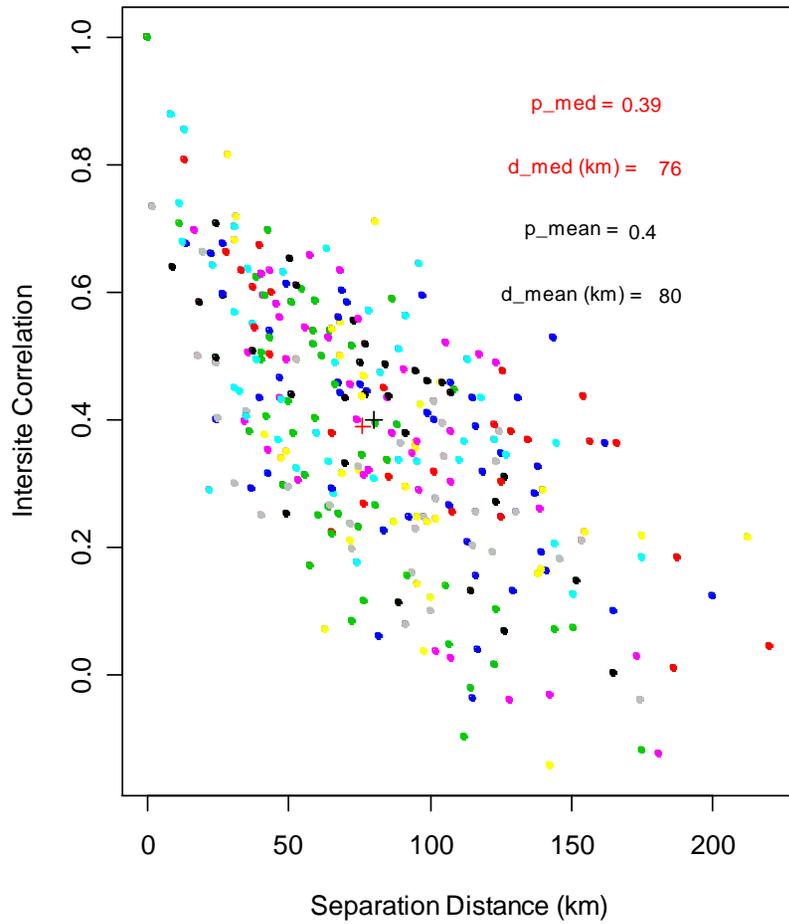


SEE - Return Level Plot

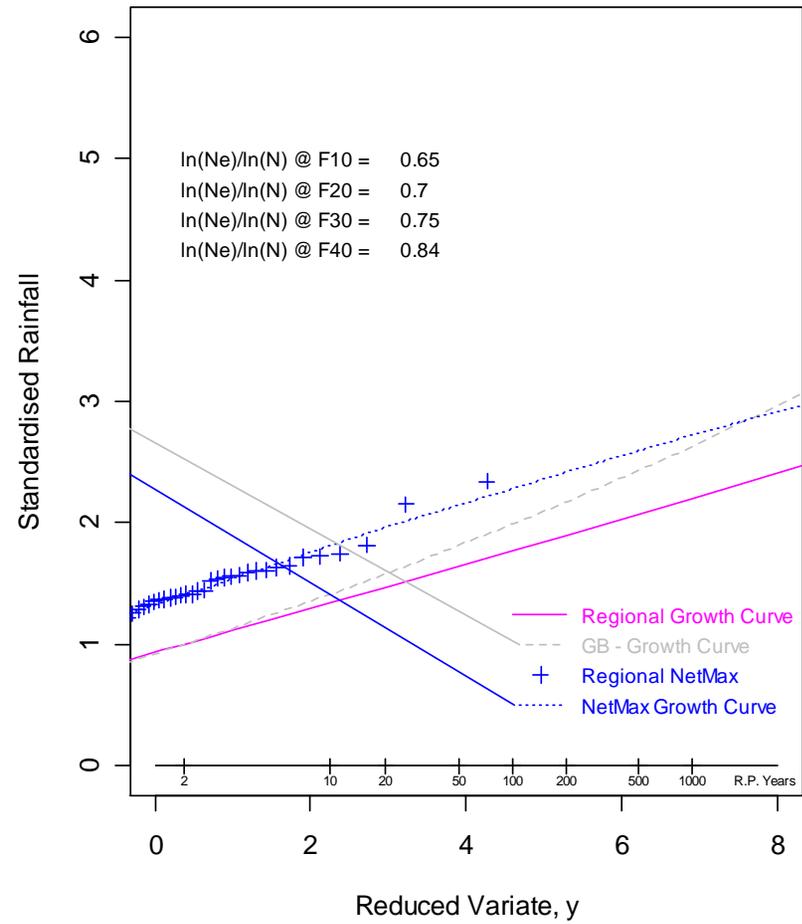


2 Day Annual Maxima

SS - Distance Correlation Relationship

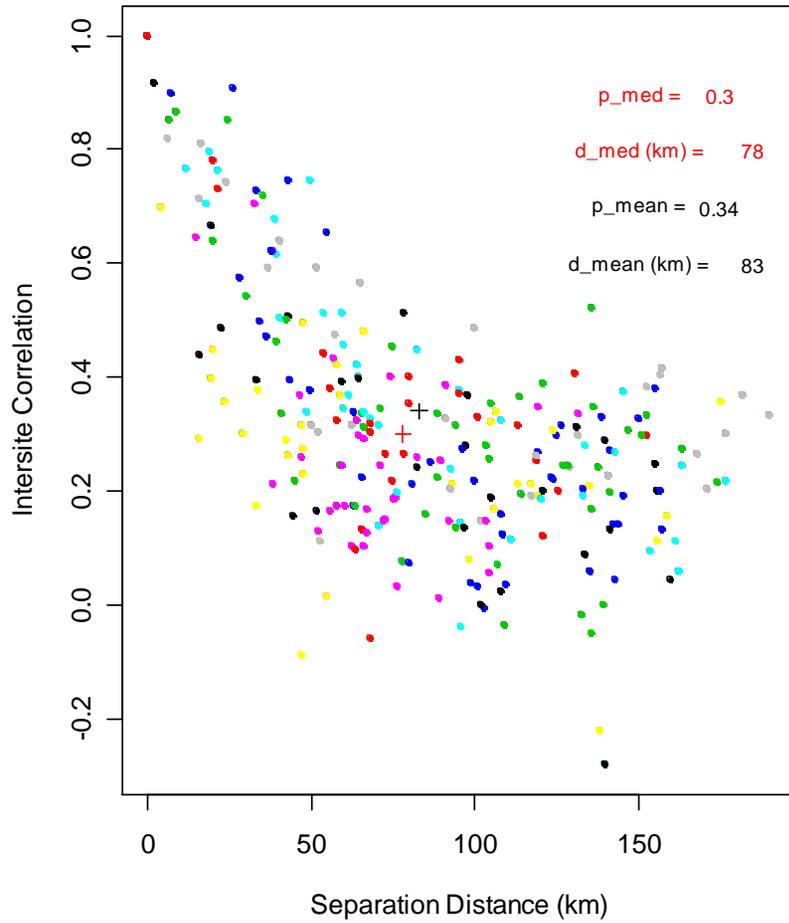


SS - Return Level Plot

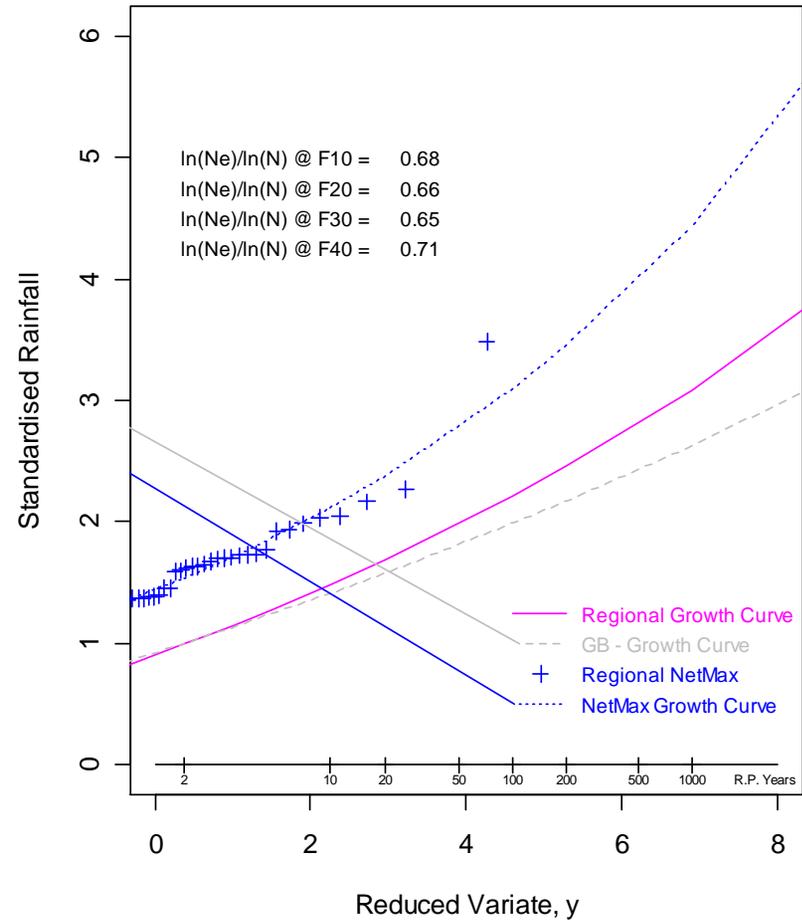


5 Day Annual Maxima

ES - Distance Correlation Relationship

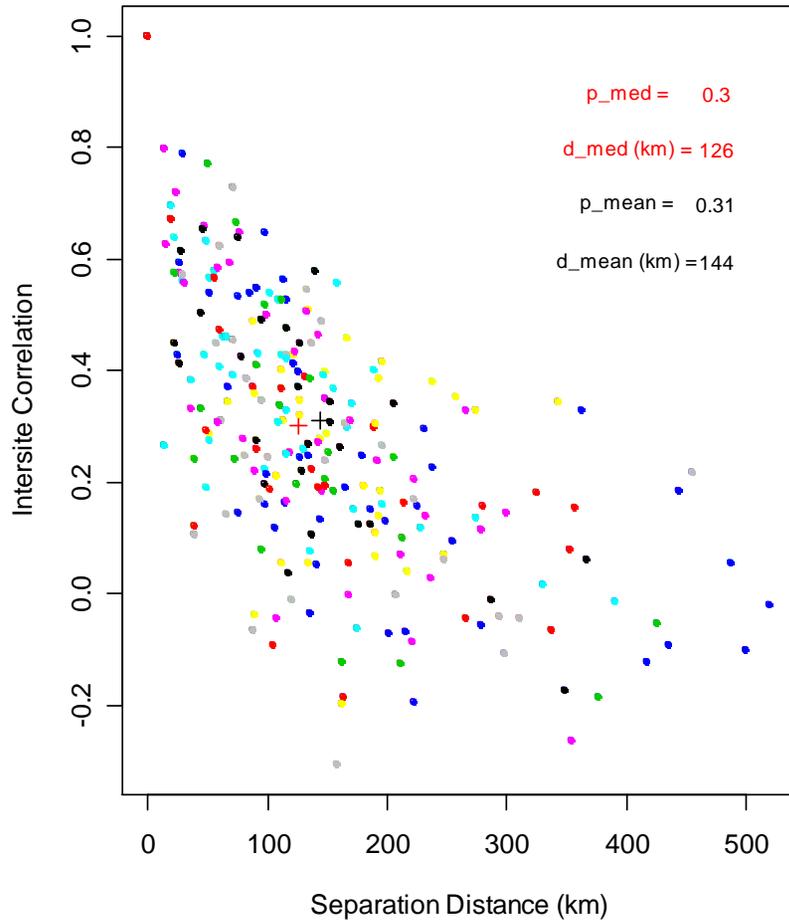


ES - Return Level Plot

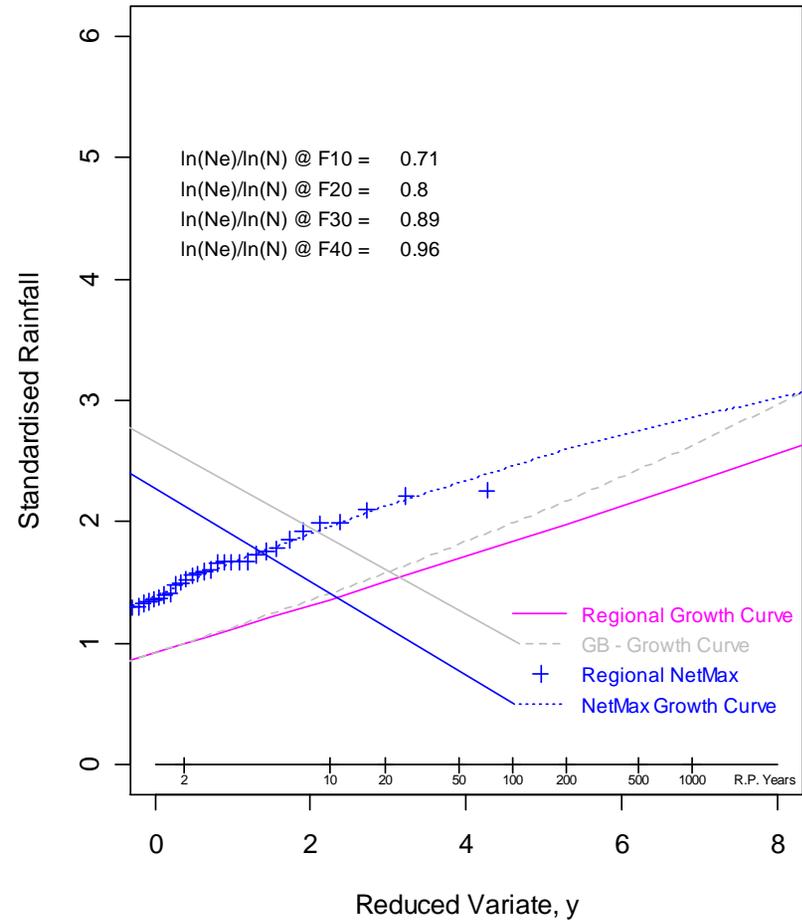


5 Day Annual Maxima

NS - Distance Correlation Relationship

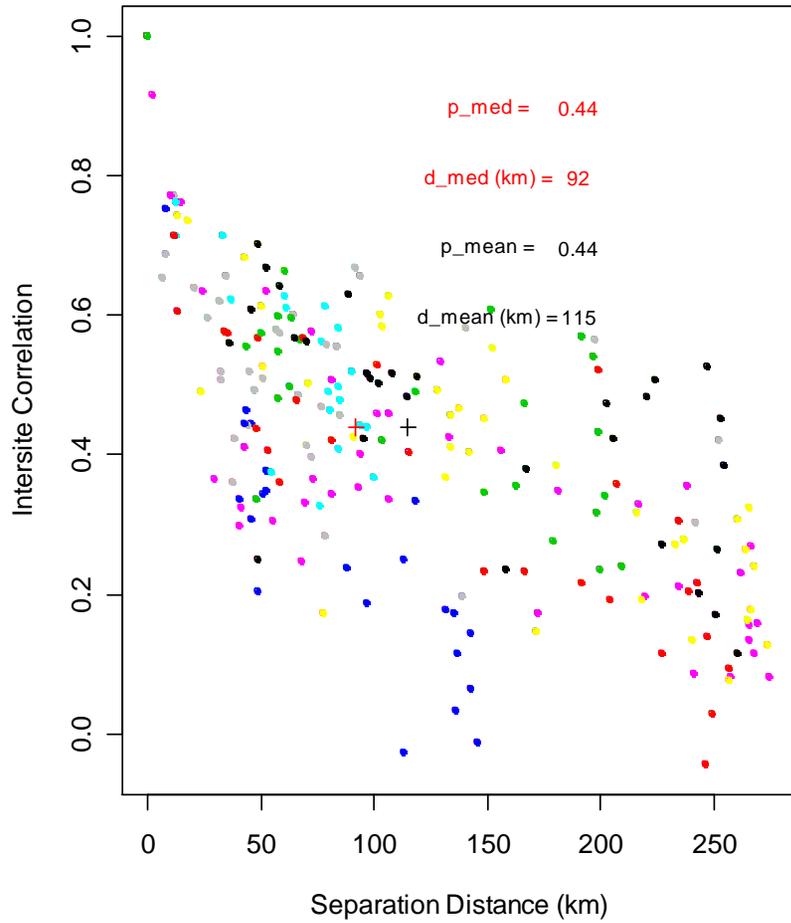


NS - Return Level Plot

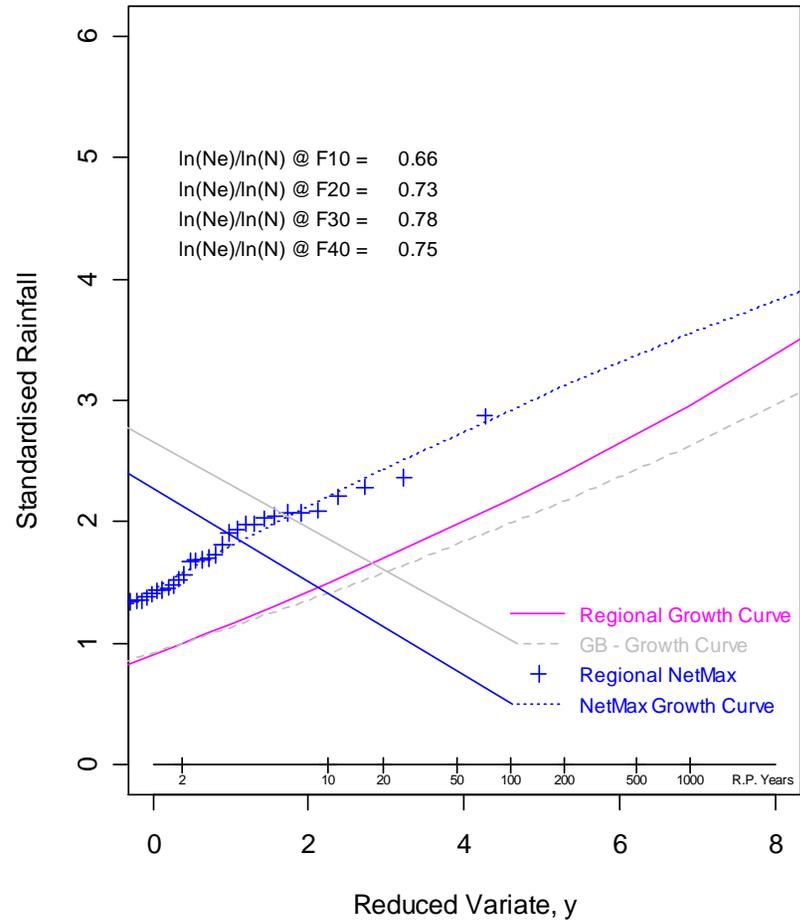


5 Day Annual Maxima

NEE - Distance Correlation Relationship

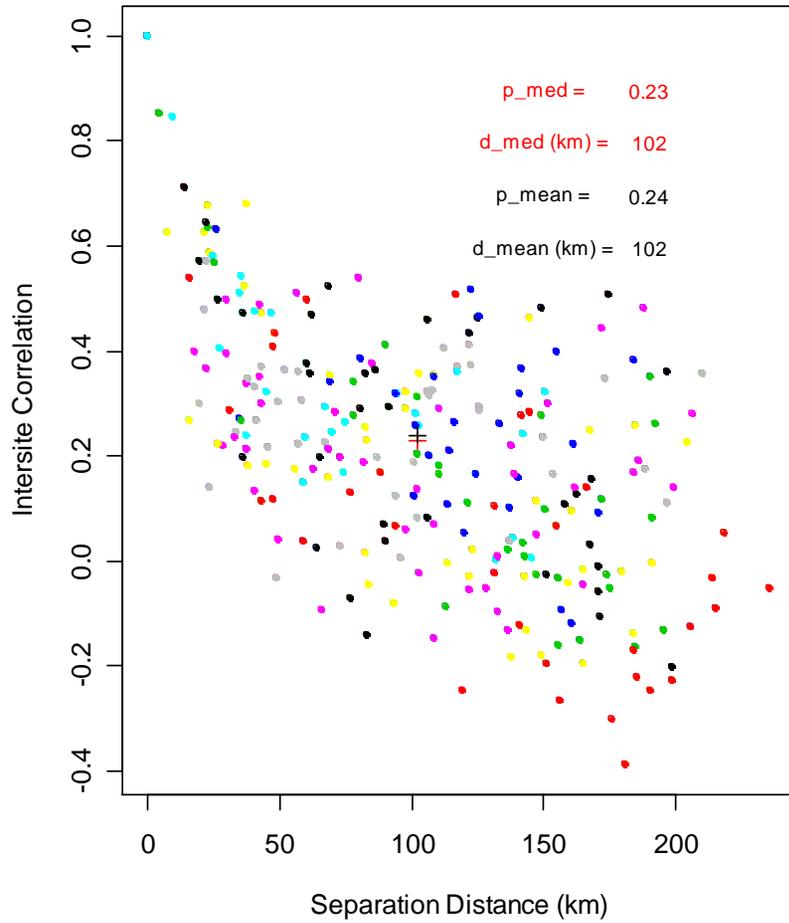


NEE - Return Level Plot

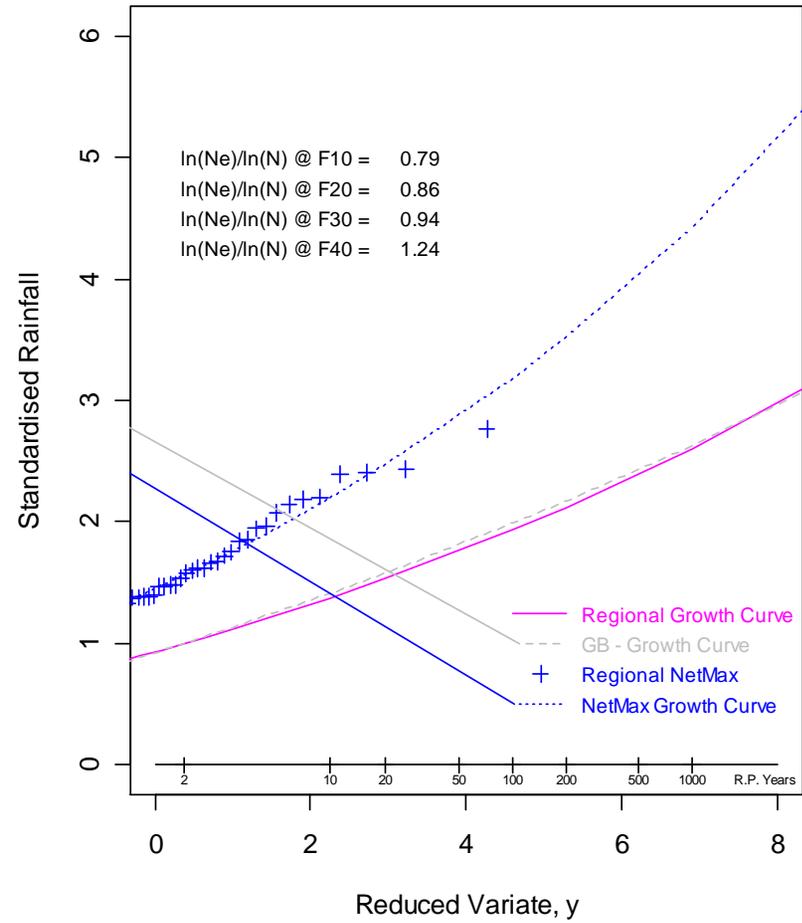


5 Day Annual Maxima

NWE - Distance Correlation Relationship

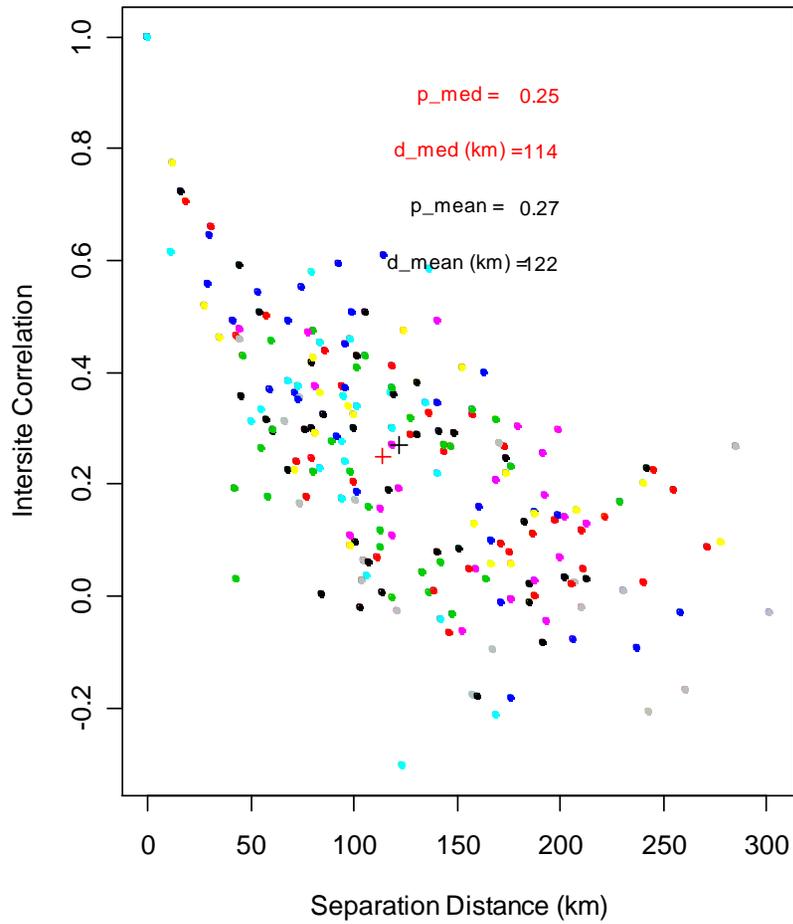


NWE - Return Level Plot

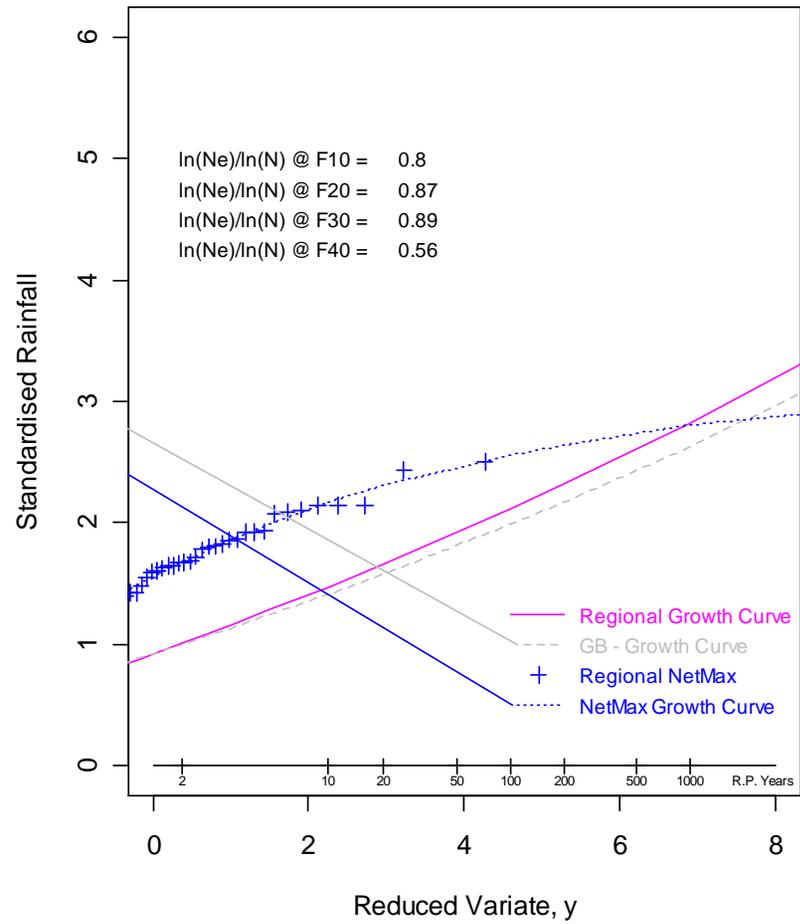


5 Day Annual Maxima

CEE - Distance Correlation Relationship

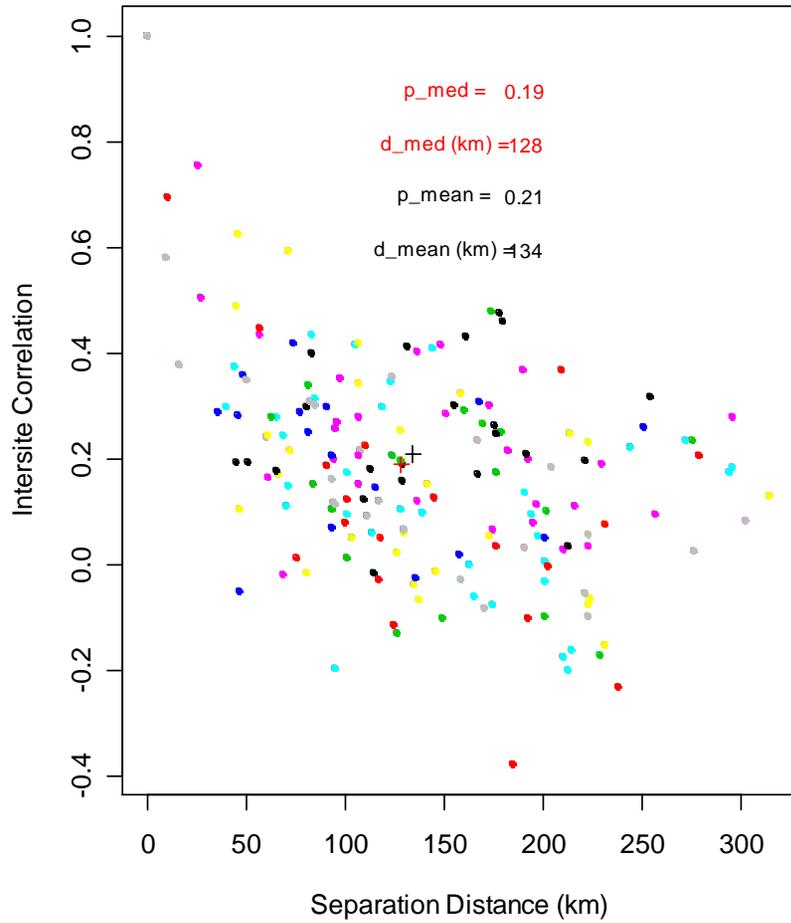


CEE - Return Level Plot

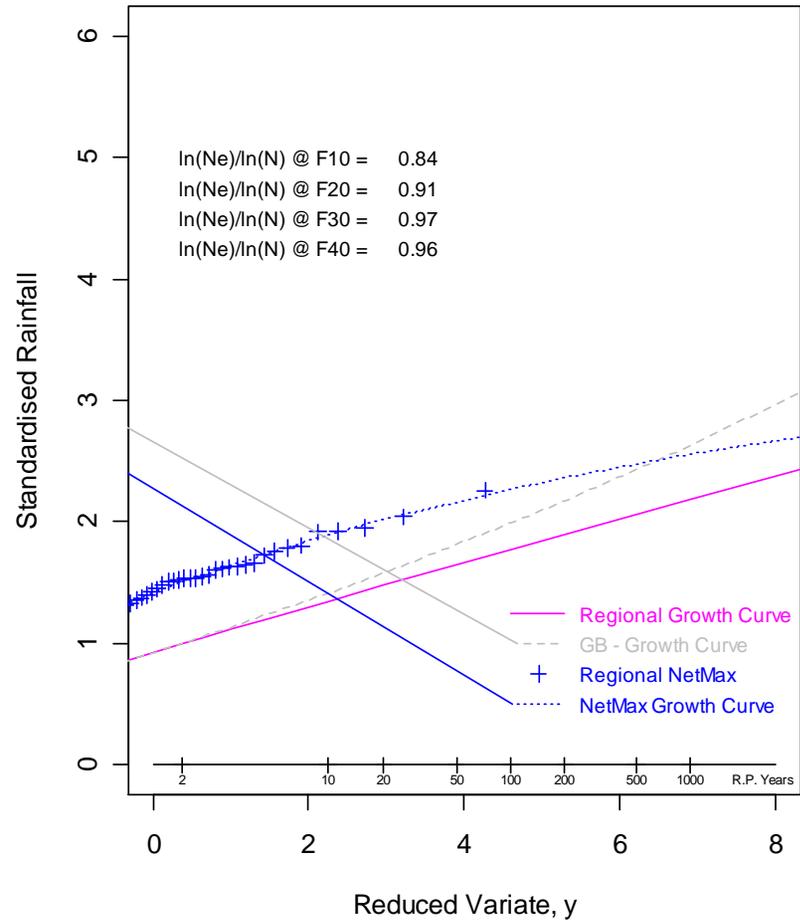


5 Day Annual Maxima

SWE - Distance Correlation Relationship

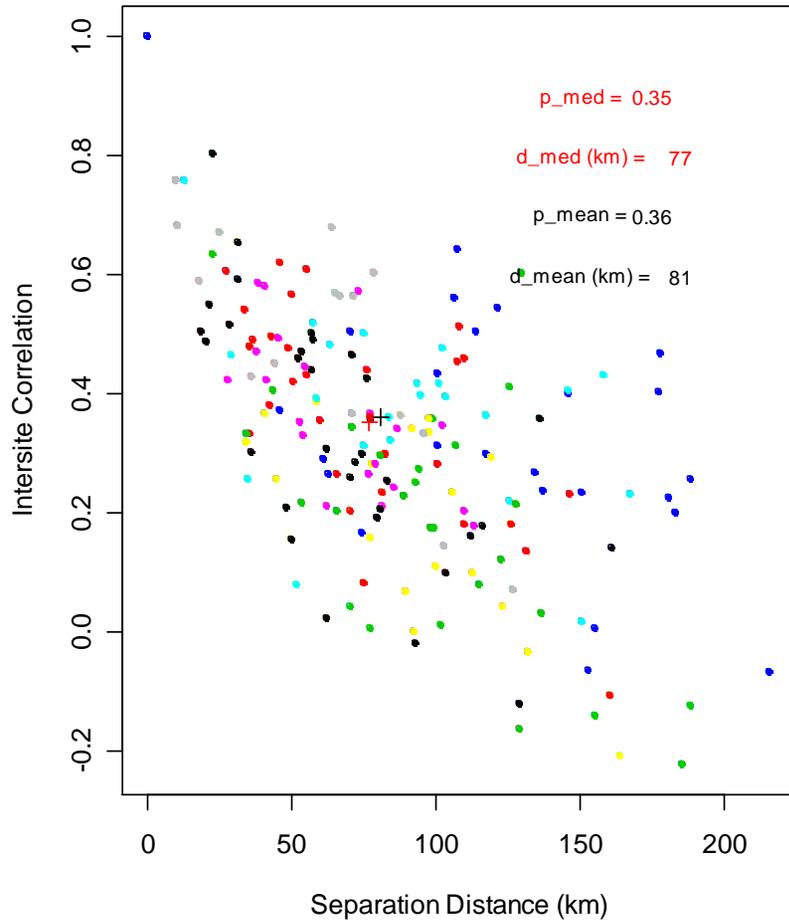


SWE - Return Level Plot



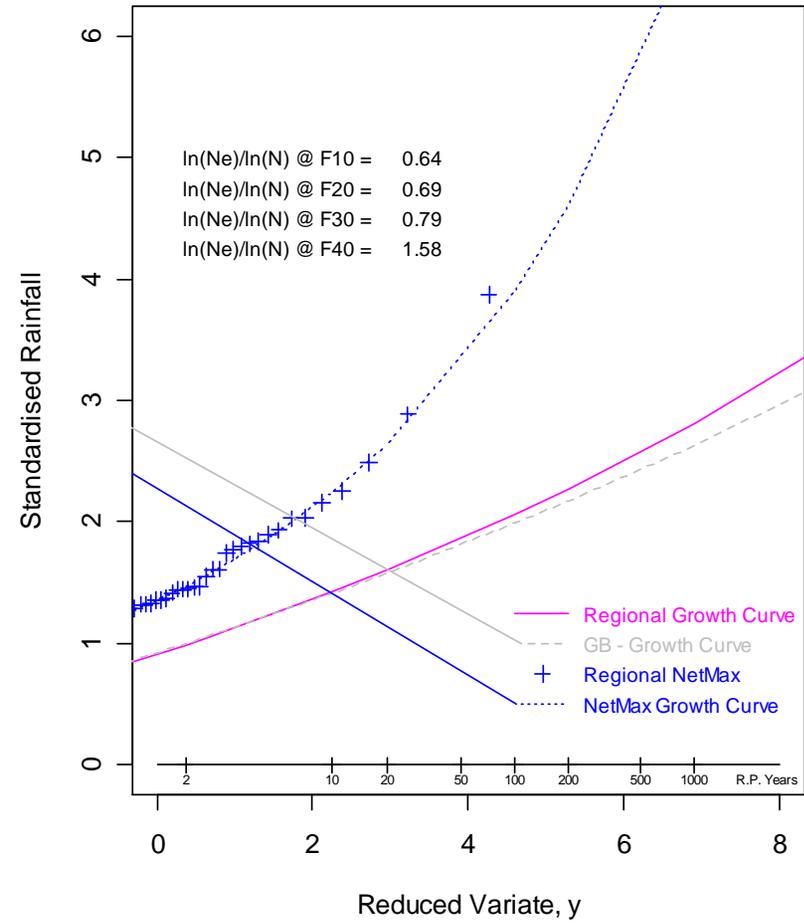
5 Day Annual Maxima

SEE - Distance Correlation Relationship

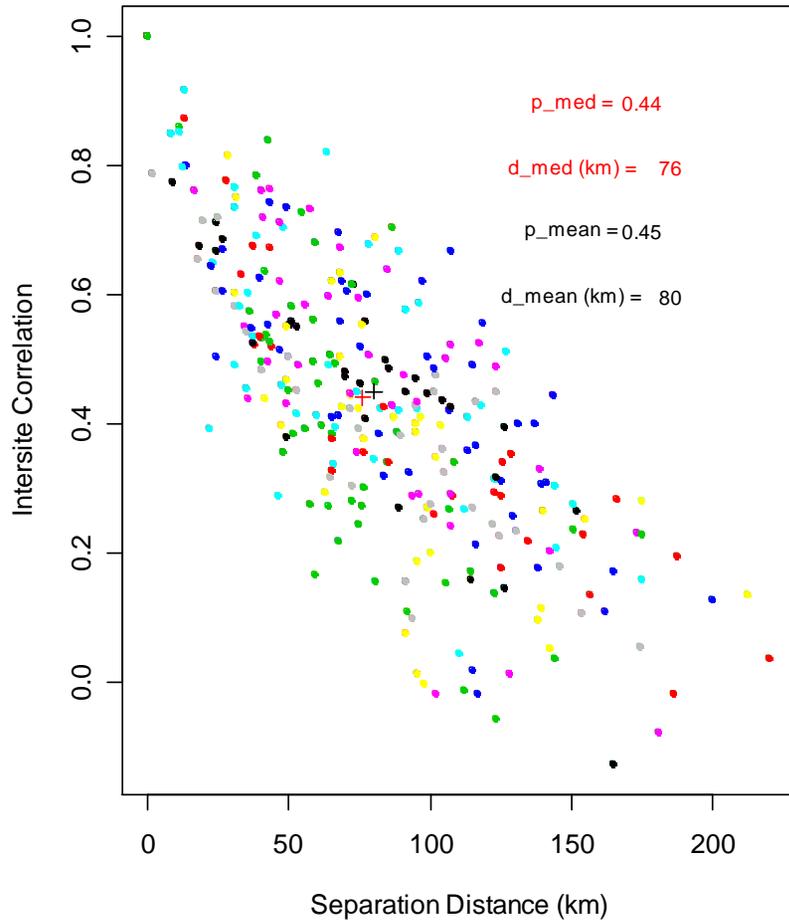


5 Day Annual Maxima

SEE - Return Level Plot

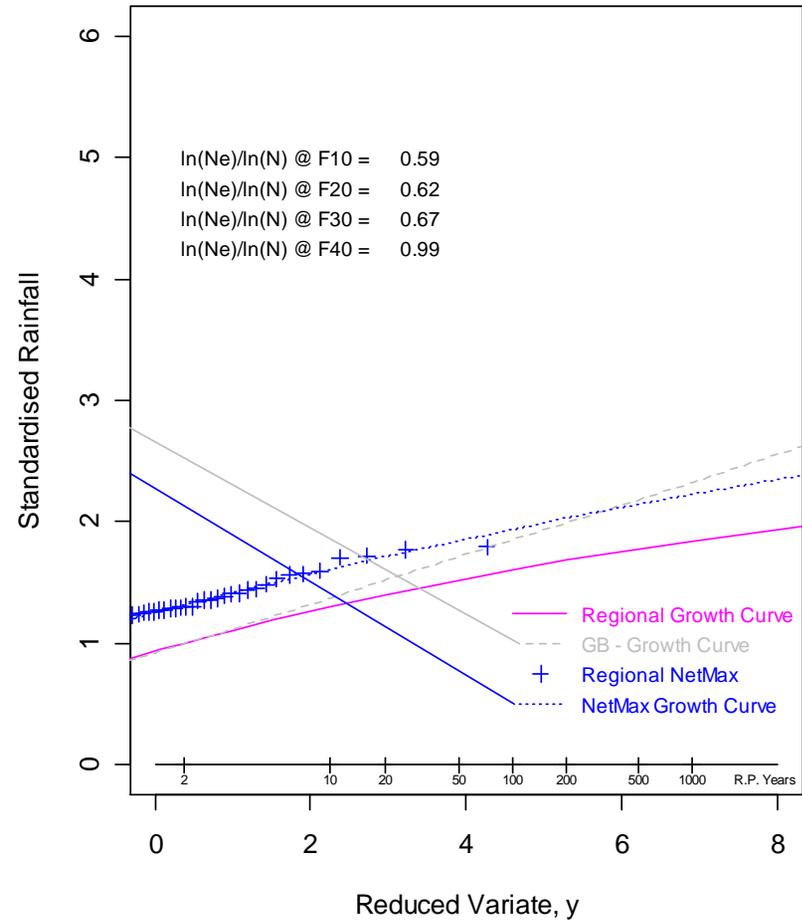


SS - Distance Correlation Relationship

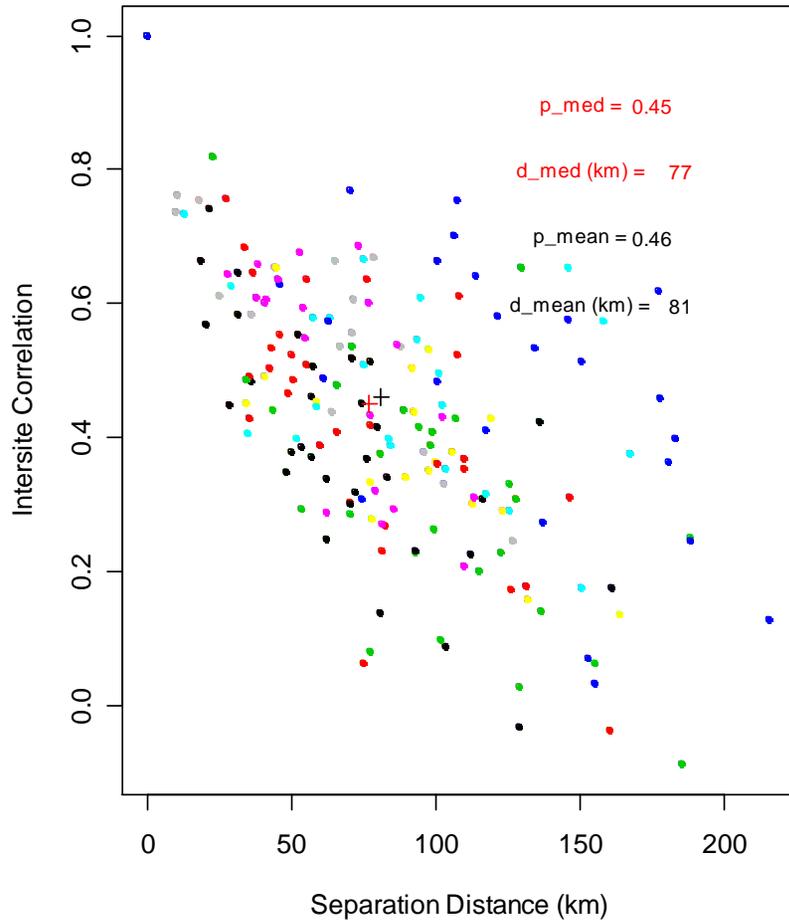


10 Day Annual Maxima

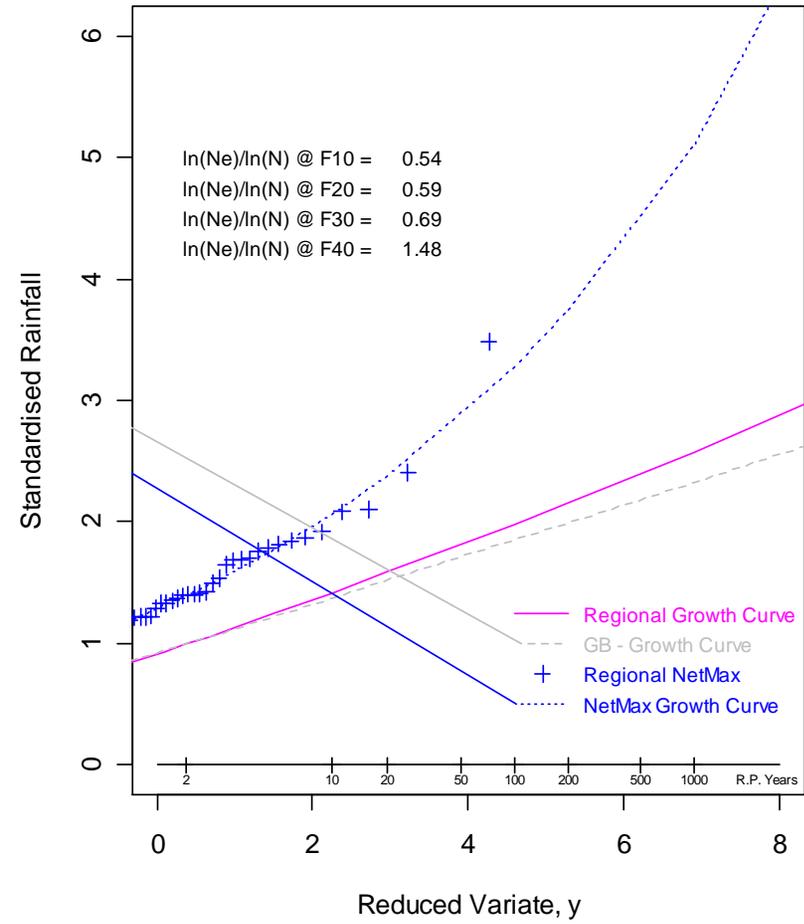
SS - Return Level Plot



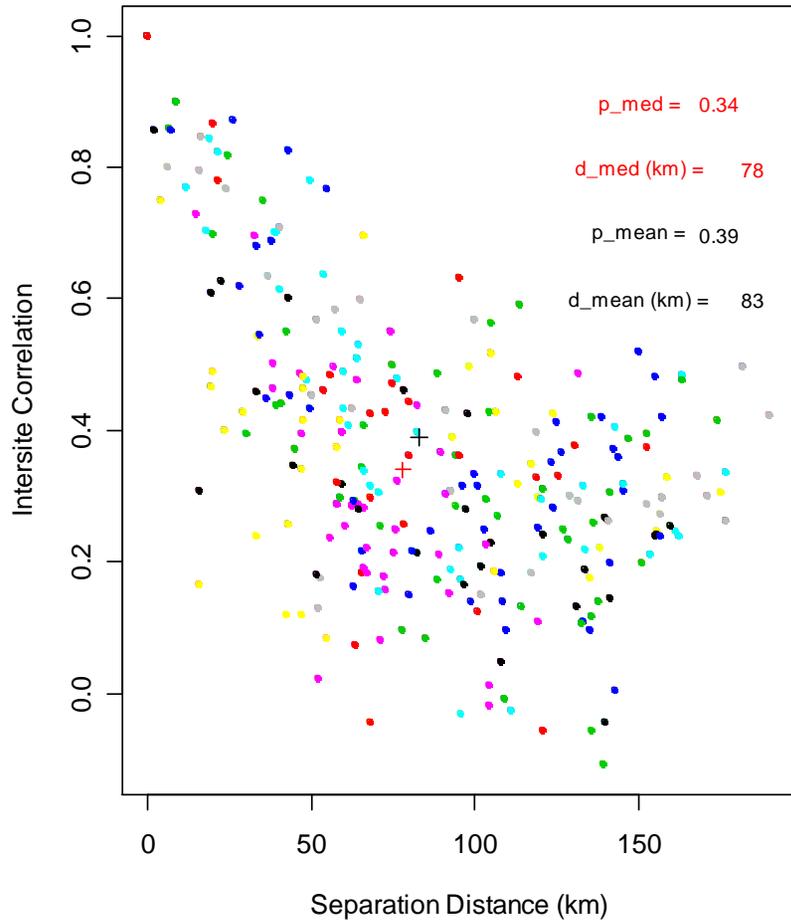
SEE - Distance Correlation Relationship



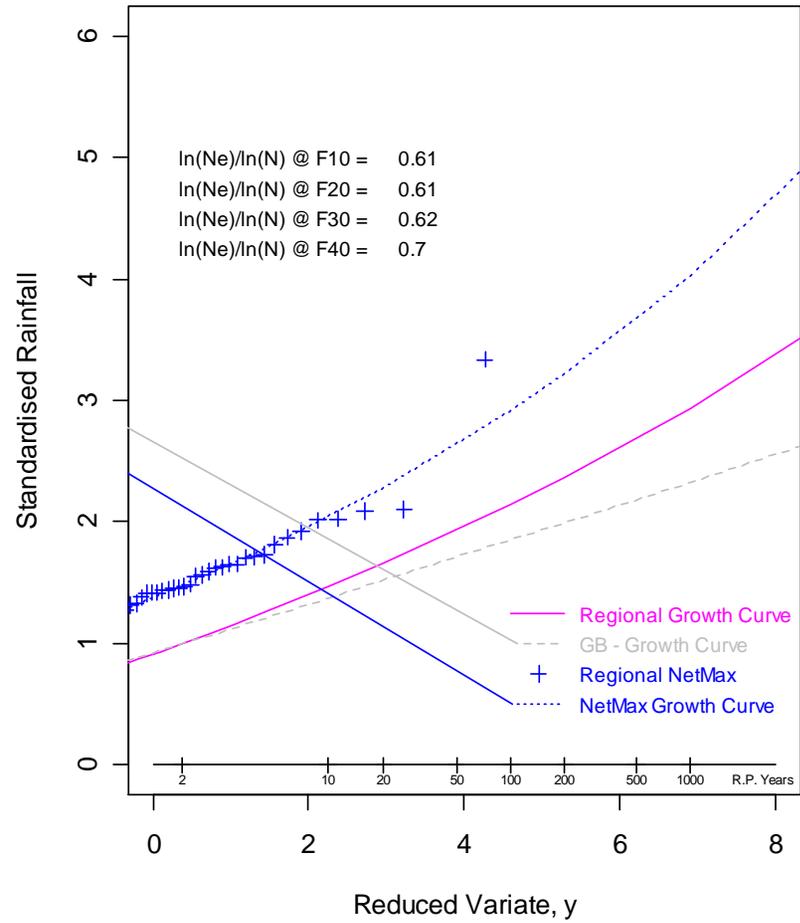
SEE - Return Level Plot



ES - Distance Correlation Relationship

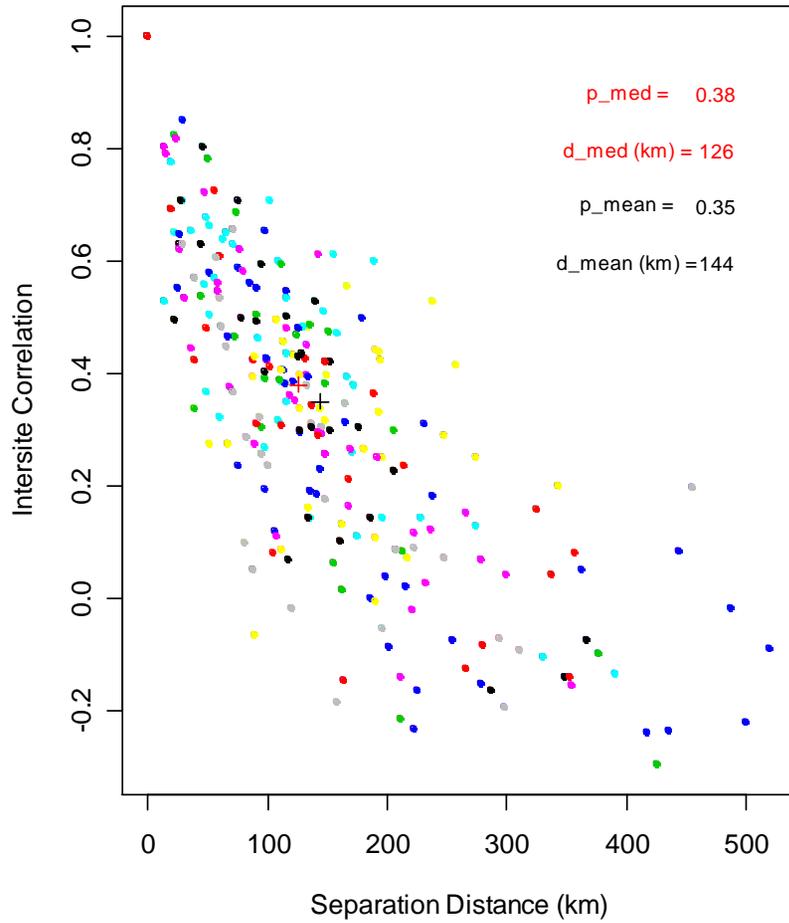


ES - Return Level Plot



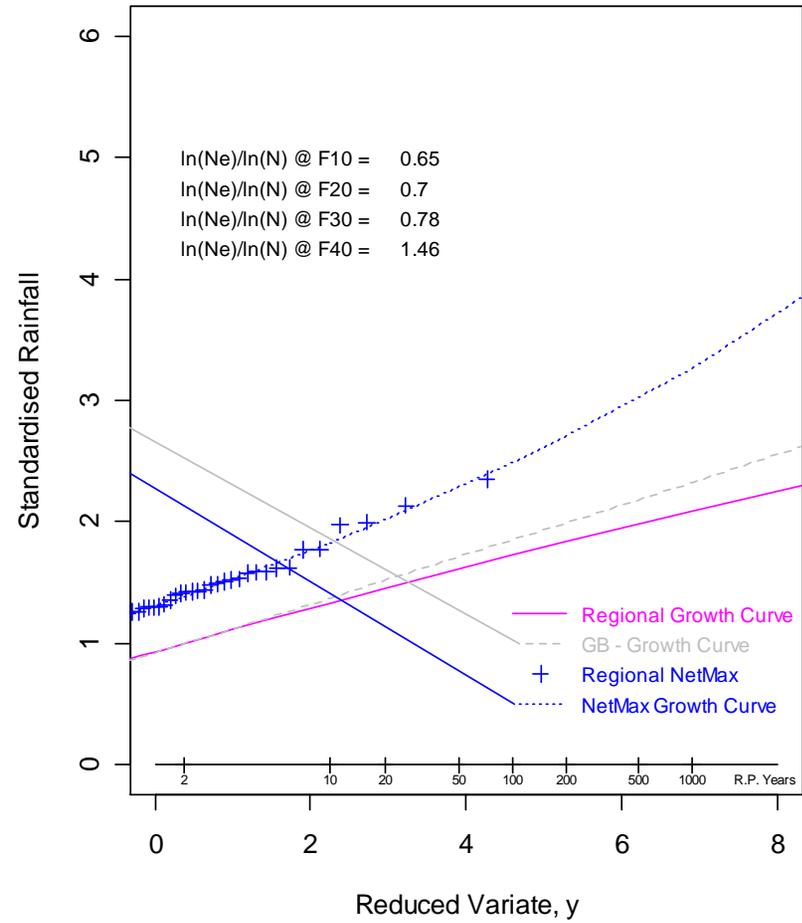
10 Day Annual Maxima

NS - Distance Correlation Relationship

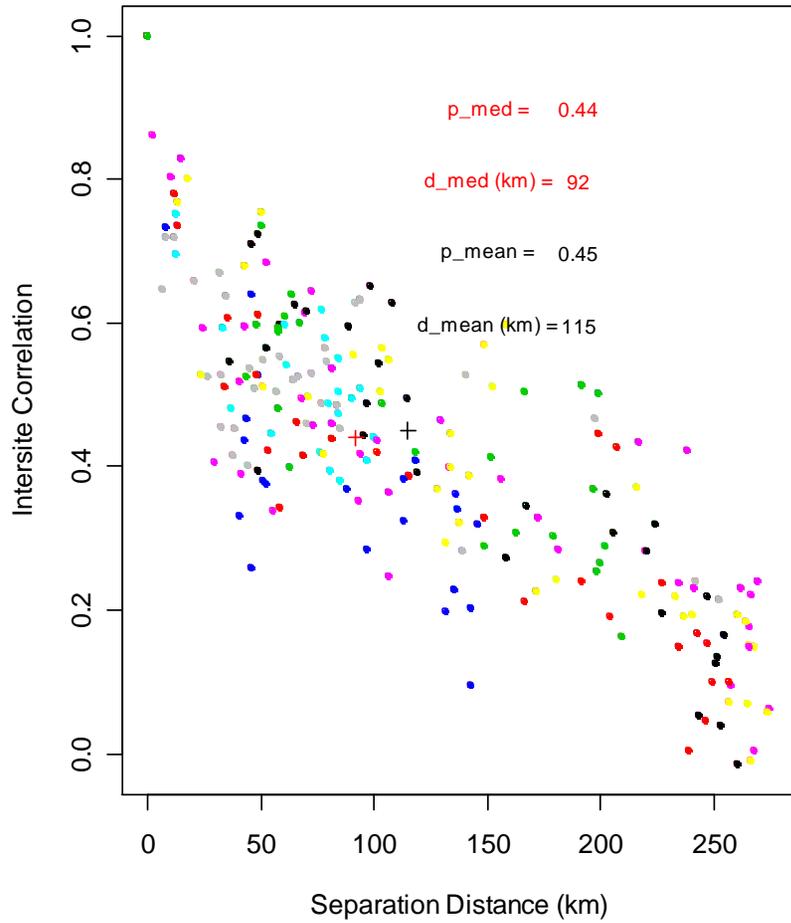


10 Day Annual Maxima

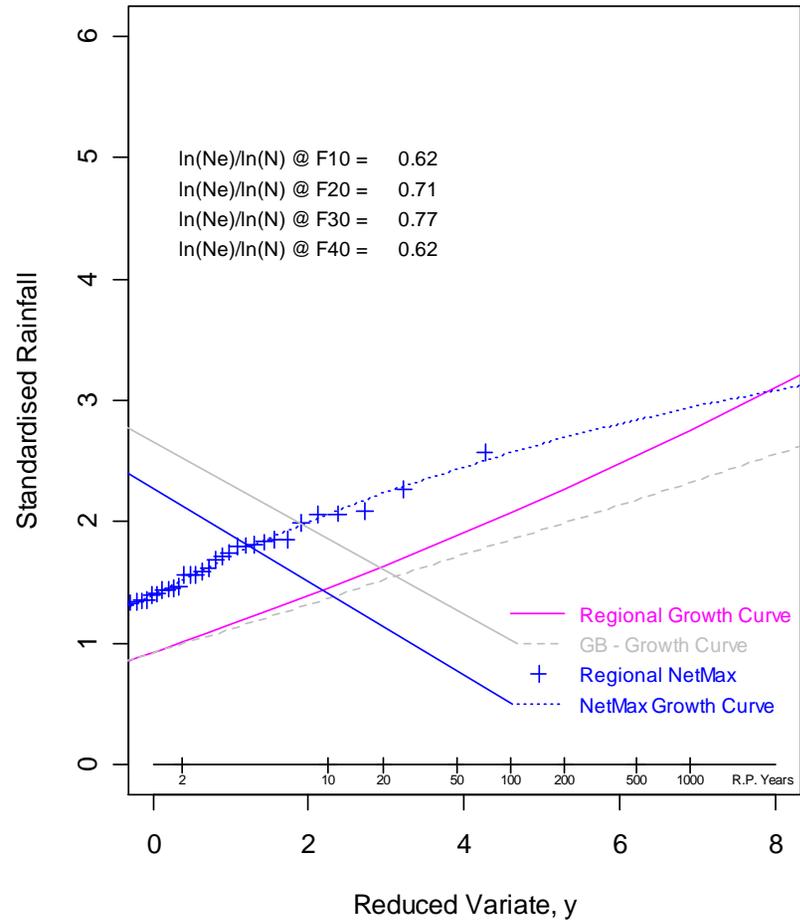
NS - Return Level Plot



NEE - Distance Correlation Relationship

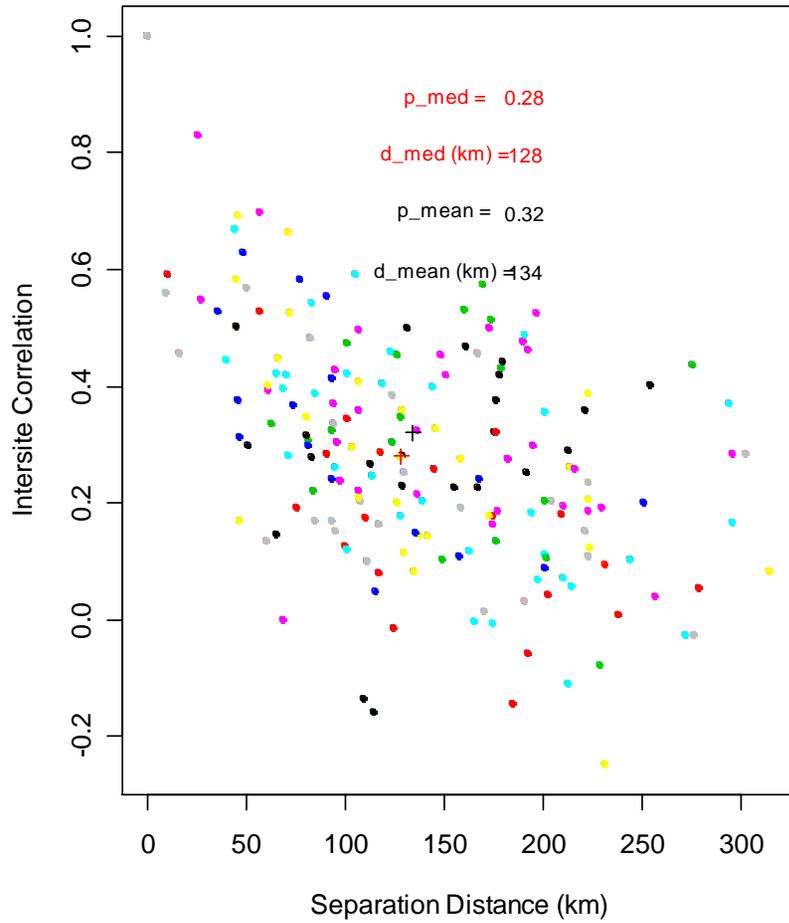


NEE - Return Level Plot



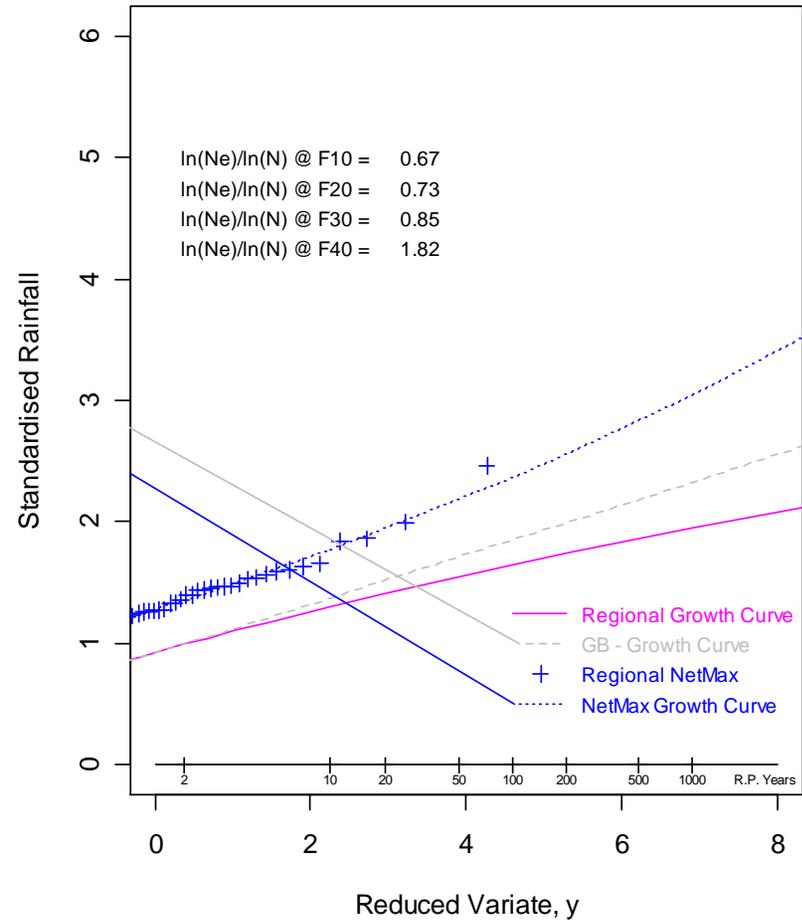
10 Day Annual Maxima

SWE - Distance Correlation Relationship

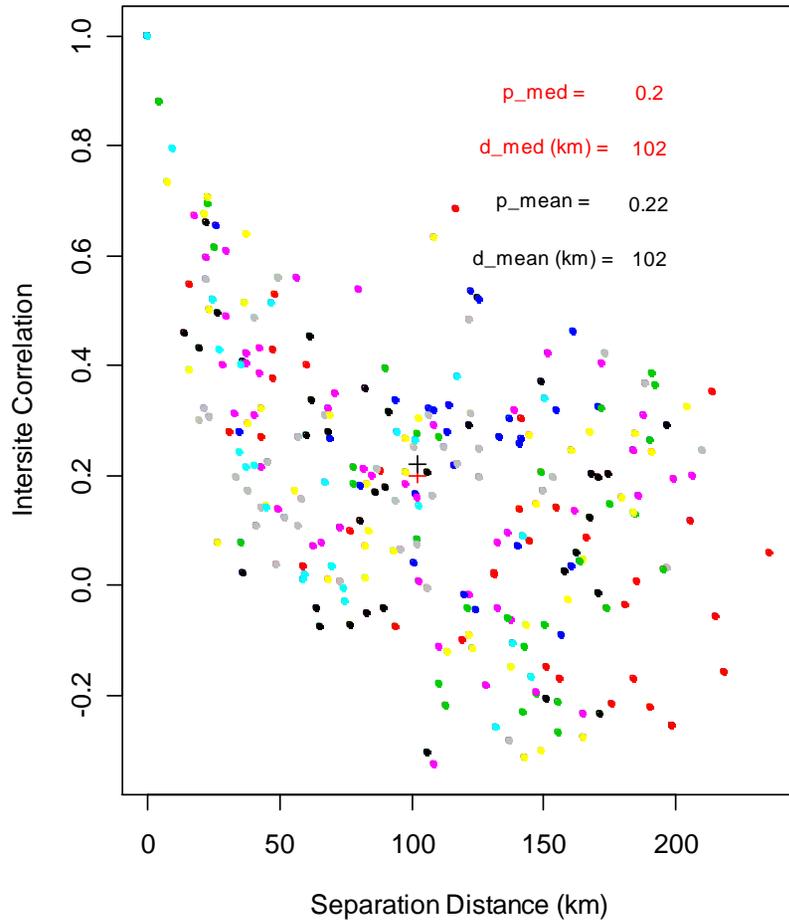


10 Day Annual Maxima

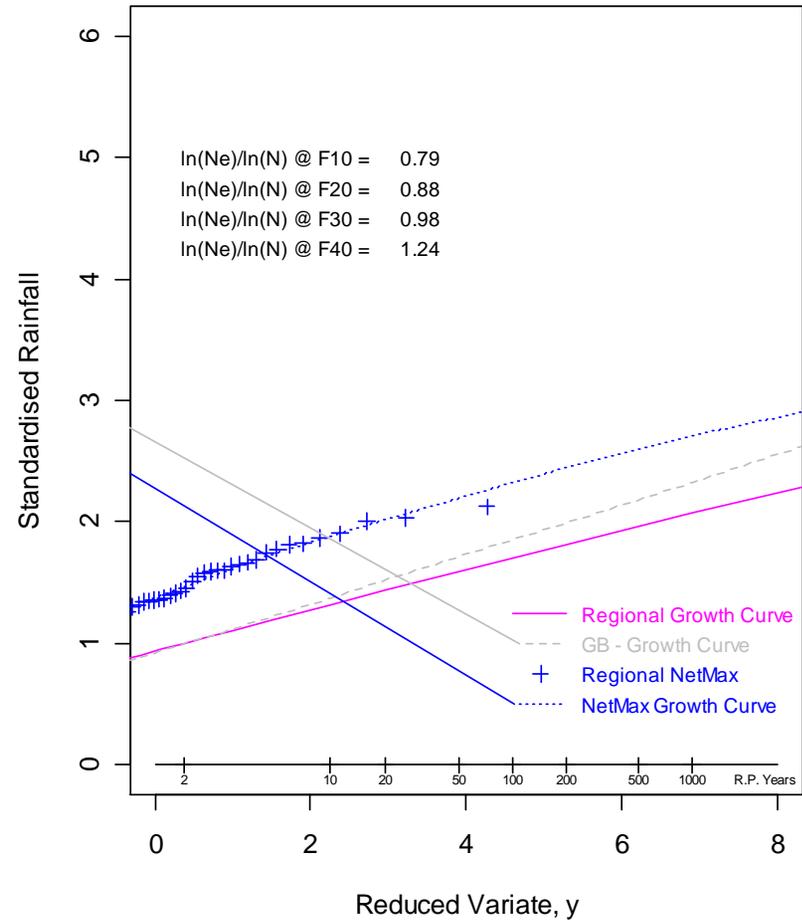
SWE - Return Level Plot



NWE - Distance Correlation Relationship

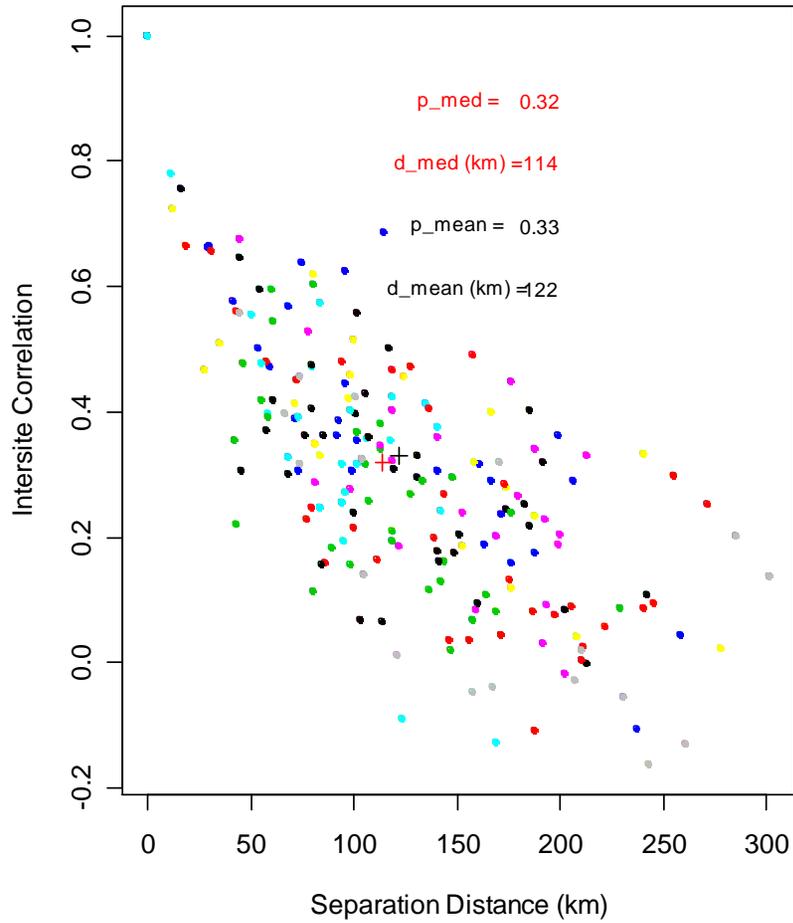


NWE - Return Level Plot

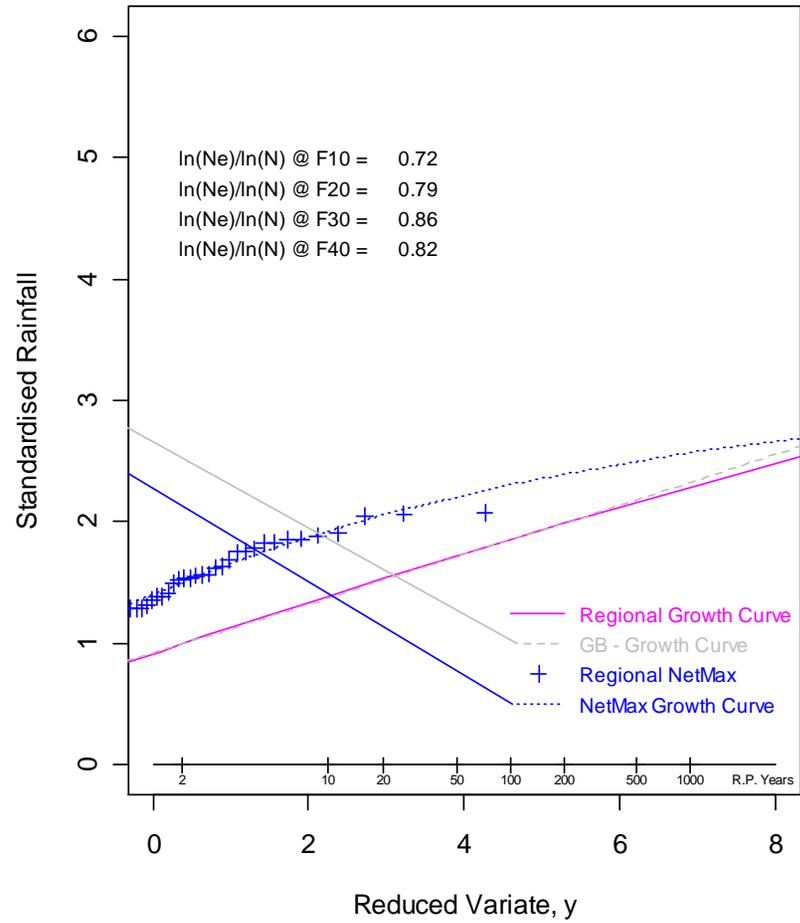


10 Day Annual Maxima

CEE - Distance Correlation Relationship

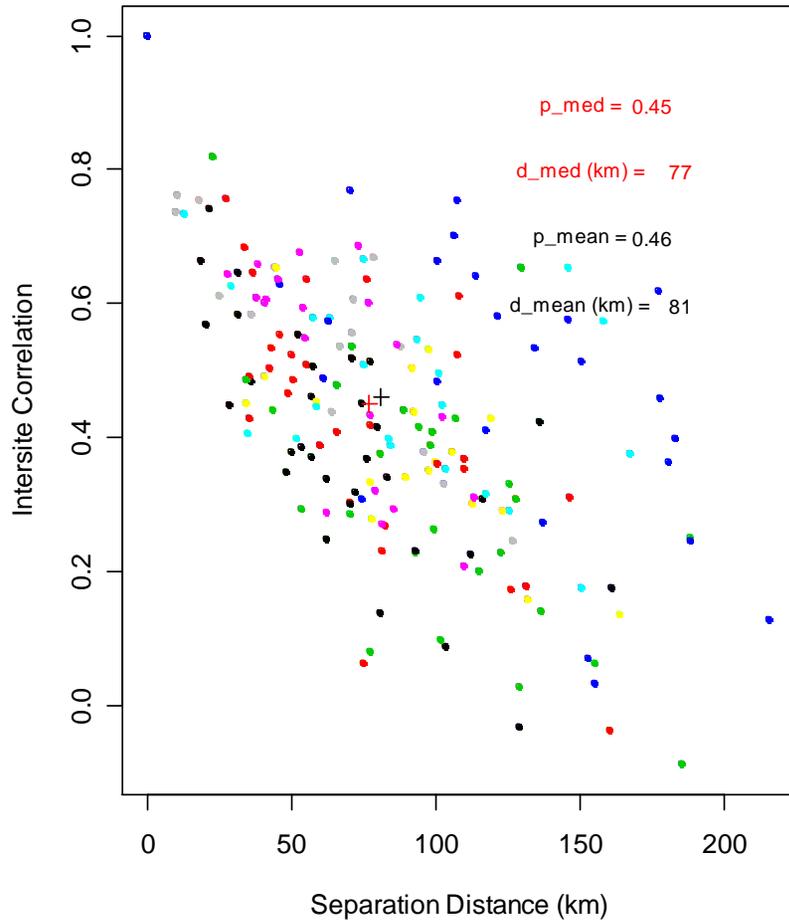


CEE - Return Level Plot



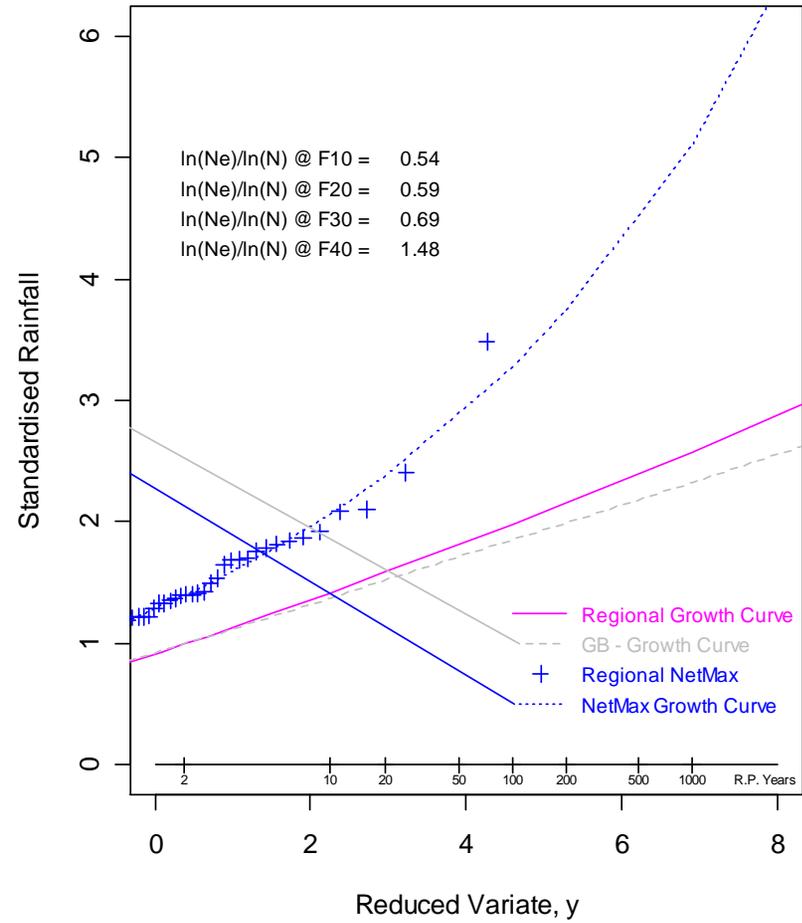
10 Day Annual Maxima

SEE - Distance Correlation Relationship



10 Day Annual Maxima

SEE - Return Level Plot



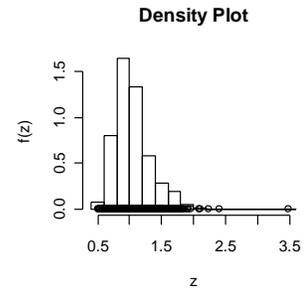
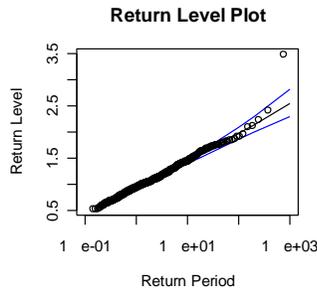
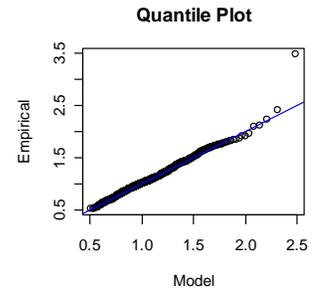
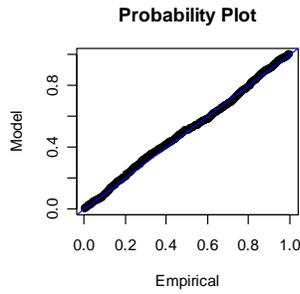
Appendix 2 - Regional Homogeneity Results

This page is intentionally blank.

South East England – 1 Day

Results for the South East of England.

u	a	k	L
0.932774	0.222635	0.04447	-3.6358
0.89935	0.244093	0.067983	-7.74763
0.872194	0.225846	-0.13177	-0.34965
0.907365	0.246507	-0.23456	-2.23582
0.924218	0.20363	0.074848	-0.9981
0.958119	0.166485	0.094999	5.967554
0.875429	0.208834	-0.15595	3.062091
0.957191	0.26367	-0.15562	-6.03652
0.899831	0.195064	-0.04329	2.938675
0.915367	0.194487	0.12659	-0.67357
0.880931	0.231147	0.28071	-10.4426
0.927126	0.214044	-0.1759	2.767844
0.888453	0.206067	-0.05273	1.337721
0.923215	0.155987	0.246487	5.542211
0.938524	0.230097	0.032458	-4.77746
0.94808	0.187561	0.069986	2.080728
0.955957	0.258834	0.059374	-9.84363
0.892858	0.215372	0.119037	-4.58358
0.929551	0.201486	-0.03807	1.748607
0.925636	0.171924	0.127133	4.35611
NA	NA	NA	-21.5228
NA	NA	NA	61.16587



Test Statistic value

Here we see the Test Statistic value = 61

Regional Pool values:

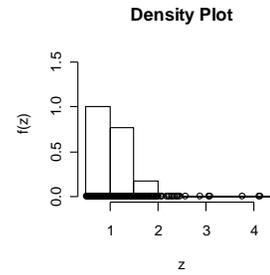
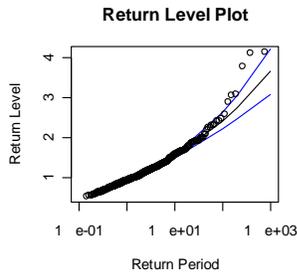
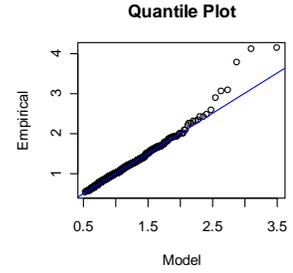
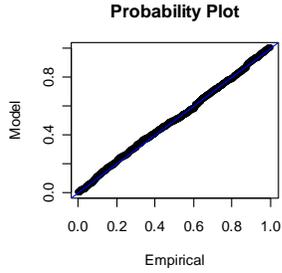
u	a	k	l
0.914419	0.218632	-0.023272	-52.1057

Chi Square Test Statistic with 20 sites in total, $\nu = 57$ @ 5% significance = 75.62

Test Statistic = 61, therefore homogeneity is likely.

South West England – 1 Day

u	a	k	l
0.942036	0.211624	0.093996	-3.1992
0.942192	0.211769	0.192607	-5.53006
0.911459	0.224329	0.340342	-11.133
0.911124	0.267258	0.005082	-10.6507
0.896275	0.233196	0.273198	-11.1267
0.952809	0.227815	0.35171	-12.0498
0.913875	0.212925	-0.39129	6.811892
0.915441	0.211287	0.295503	-7.75602
0.889045	0.248756	0.178687	-11.4995
0.855755	0.228604	0.205212	-8.39799
0.913484	0.216958	0.13232	-4.74053
0.887935	0.275808	-0.07215	-9.73624
0.911084	0.208854	0.171804	-4.15513
0.916082	0.200704	0.299894	-5.7368
0.901103	0.190262	0.141279	-0.0196
0.948006	0.199842	-0.0044	1.490193
0.940956	0.185781	0.164788	0.335779
0.935837	0.157831	0.433141	0.983086
0.935659	0.219414	0.172069	-6.53008
0.928031	0.285398	0.039427	-12.7036
NA	NA	NA	-115.344
NA	NA	NA	67.01819



Test Statistic value

Regional Pool values:

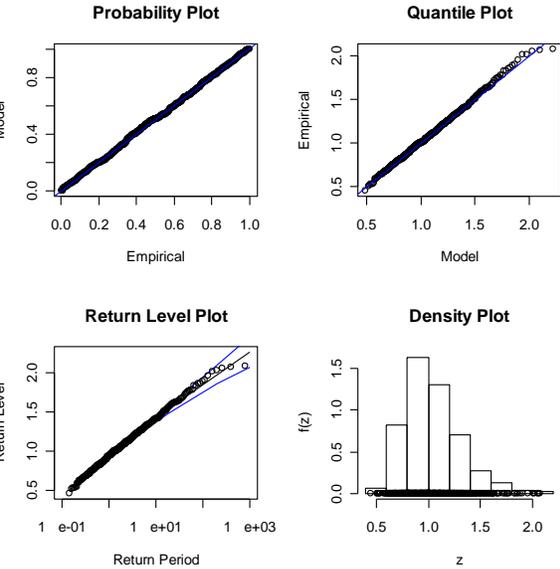
u	a	k	L
0.916089	0.230015	-0.145857	-148.853

Chi Square Test Statistic with 20 sites in total, $\nu = 57$ @ 5% significance = 75.62

Test Statistic = 67, therefore homogeneity is likely.

Central Eastern England – 1 Day

u	a	k	l
0.899759	0.212445	-0.03683	-0.60179
0.909038	0.228752	0.111828	-6.47524
0.864789	0.162739	0.214134	4.635733
0.95293	0.193697	-0.26748	8.458981
0.927076	0.237719	-0.03287	-4.99186
0.883705	0.239639	-0.16878	-2.05173
0.935604	0.151759	-0.00986	12.2212
0.921923	0.230132	-0.01723	-4.18706
0.941666	0.212872	0.129546	-4.31911
0.940511	0.257573	-0.01279	-6.82057
0.920056	0.163839	0.048643	6.741978
0.902815	0.216644	-0.03697	-0.61802
0.890979	0.192307	0.102826	0.365507
0.96	0.210295	-0.32252	6.316616
0.927436	0.23961	-0.13947	-2.85301
0.928653	0.182728	0.125476	1.754144
0.895838	0.218525	-0.19626	2.013733
0.893254	0.21594	0.087398	-3.924
0.919228	0.237341	-0.27093	0.18713
0.985237	0.232833	0.048431	-6.16186
0.9122	0.192191	-0.05164	3.839074
NA	NA	NA	3.529842
NA	NA	NA	58.39992



Test Statistic value

Regional Pool values:

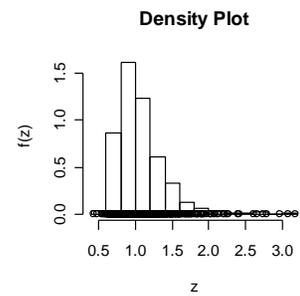
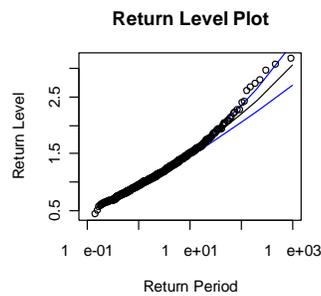
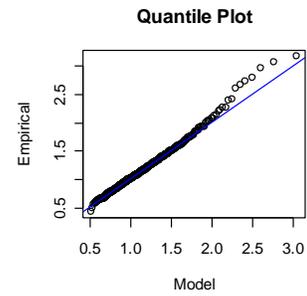
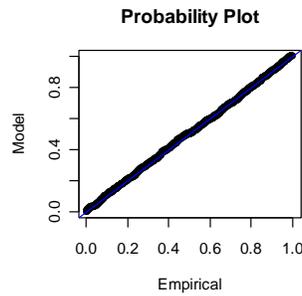
u	a	k	l
0.917964	0.217116	0.03288	-25.6701

Chi Square Test Statistic with 21 sites in total, $\nu = 60$ @ 5% significance = 70.98

Test Statistic = 58, therefore homogeneity is likely.

North West England – 1 Day

u	a	k	l
0.94553	0.188749	0.190615	-0.67333
0.907328	0.216784	0.030413	-2.51238
0.898369	0.249187	0.021219	-7.29694
0.930154	0.181573	0.206503	0.349248
0.926403	0.192171	0.097917	0.437488
0.8975	0.266808	-0.0073	-10.0391
0.91251	0.221872	0.419032	-12.2688
0.96692	0.22013	0.126421	-5.57423
0.922807	0.238798	-0.08935	-4.26598
0.902326	0.225672	0.02628	-4.46473
0.890462	0.229307	-0.23308	0.839642
0.891126	0.246529	0.419742	-16.5299
0.921972	0.190293	0.183044	-0.89032
0.926668	0.175995	-0.01911	6.537236
0.953762	0.171764	0.179888	2.916546
0.905469	0.236731	0.243489	-10.2326
0.909837	0.201044	-0.03885	1.610441
0.904128	0.185789	0.052009	3.195819
0.885019	0.220295	0.047079	-3.81472
0.912519	0.223499	0.093202	-5.1437
0.885504	0.21074	0.07849	-2.8039
0.947227	0.236704	0.192705	-9.90341
0.918037	0.23535	0.096646	-7.63595
0.928965	0.184621	0.118597	1.545812
NA	NA	NA	-86.6177
NA	NA	NA	70.93757



Test Statistic value

Regional Pool values:

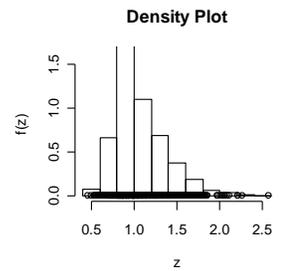
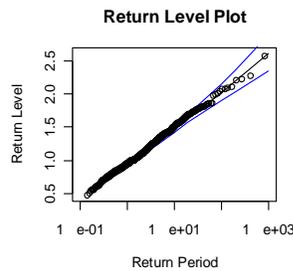
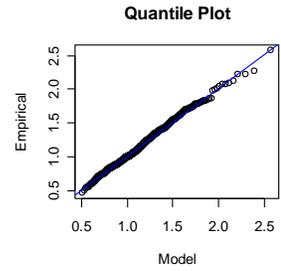
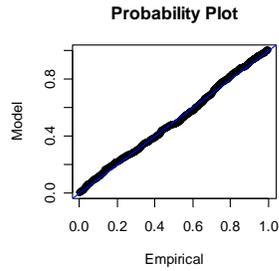
u	a	k	l
0.915916	0.223018	-0.093177	-122.087

Chi Square Test Statistic with 24 sites in total, $\nu = 69$ @ 5% significance = 89.39

Test Statistic = 71, therefore homogeneity is likely.

North East England – 1 Day

u	a	k	l
0.968935	0.225573	0.003119	-3.79085
0.945197	0.238804	-0.06554	-4.57351
0.943873	0.164749	0.177523	4.919153
0.9115	0.216679	-0.12998	1.198109
0.933695	0.254468	-0.1428	-5.04443
0.959522	0.227181	0.16445	-7.73734
0.939142	0.174808	-0.15834	9.797928
0.897273	0.157419	0.401558	1.889436
0.958943	0.183264	-0.01212	4.836467
0.875111	0.268815	-0.20217	-6.35646
0.914081	0.198274	-0.00086	1.558655
0.94677	0.261287	0.260435	-14.9971
0.938954	0.178544	0.127099	2.796644
0.921443	0.185516	0.222843	-0.68347
0.907342	0.205366	0.08528	-1.85812
0.945068	0.199901	0.138125	-2.0729
0.938516	0.182782	0.110998	2.264898
0.934061	0.308688	0.044488	-16.6959
0.877806	0.250919	0.177848	-11.9788
0.941886	0.193709	0.159347	-1.17022
0.929485	0.197261	0.1264	-1.21054
0.921401	0.21953	-0.00267	-2.40675
NA	NA	NA	-51.315
NA	NA	NA	81.56493



Test Statistic value

Regional Pool values:

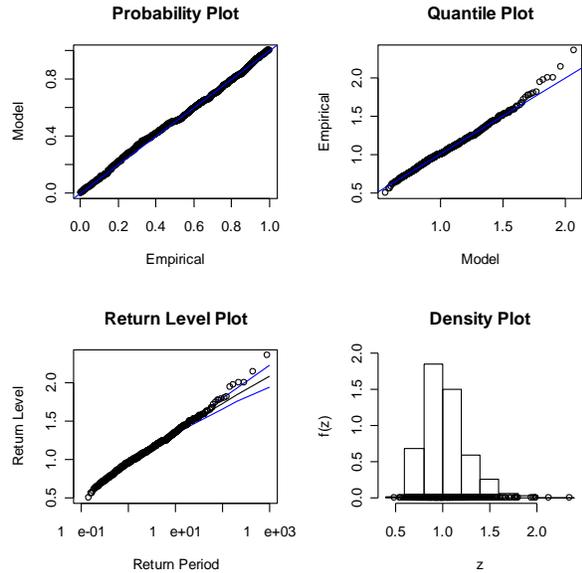
u	a	k	l
0.932311	0.22663	-0.021528	-92.0975

Chi Square Test Statistic with 22 sites in total, $\nu = 63$ @ 5% significance = 82.52

Test Statistic = 81.56, therefore homogeneity is likely.

Northern Scotland – 1 Day

u	a	k	l
0.943749	0.223581	-0.01733	-3.02923
0.947778	0.203802	0.059936	-0.82022
0.930816	0.172614	-0.01624	7.258746
0.928102	0.199755	-0.09948	3.16946
0.929598	0.195707	0.088267	0.184645
0.934393	0.226132	-0.03866	-2.65212
0.944829	0.158863	0.074988	8.364004
0.921562	0.187722	-0.00137	3.617643
0.913388	0.200692	-0.15418	4.436191
0.93294	0.254688	-0.00412	-7.59641
0.954854	0.189902	-0.21206	7.712822
0.891252	0.233049	0.135402	-7.82599
0.917912	0.164129	-0.19242	12.75885
0.913612	0.181948	-0.07266	6.855754
0.91739	0.163867	-0.06564	10.23955
0.969395	0.179624	-0.24288	9.91506
0.94444	0.15978	0.013144	10.04166
0.94594	0.141299	-0.04402	16.25909
0.951959	0.156646	-0.04941	11.78025
0.964893	0.159352	-0.15548	12.82084
0.954753	0.143574	-0.369	22.2755
0.948994	0.126085	-0.18411	23.79358
0.954053	0.199454	-0.20648	5.553423
NA	NA	NA	155.1131
NA	NA	NA	89.0895



Test Statistic value

Regional Pool values:

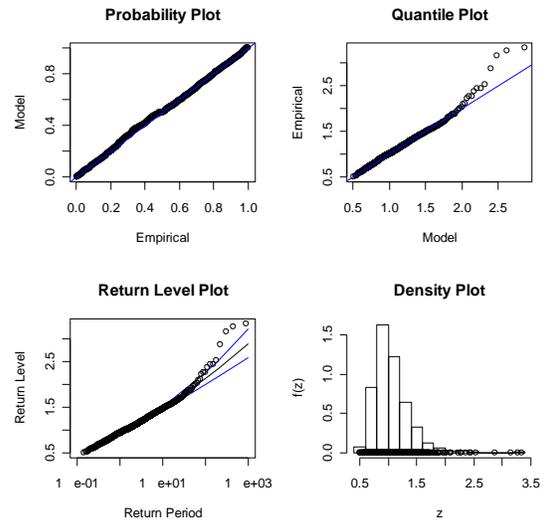
u	a	k	l
0.932685	0.186673	0.03303	110.5684

Chi Square Test Statistic with 23 sites in total, $\nu = 66$ @ 5% significance = 85.96

Test Statistic = 89, therefore appears to be non-homogeneous.

Eastern Scotland – 1 Day

u	a	k	l
0.938597	0.232714	-0.02937	-3.71209
0.907734	0.242429	0.089846	-7.82926
0.933306	0.240213	0.294998	-12.545
0.914276	0.238529	0.065731	-7.33169
0.949034	0.19687	-0.32062	8.155842
0.904175	0.226133	-0.12686	-1.07946
0.921487	0.235912	0.029518	-5.82612
0.944492	0.216206	0.130151	-4.86773
0.909186	0.199756	0.083445	-0.44637
0.938759	0.246101	0.029714	-7.19759
0.921963	0.199326	0.262719	-4.17106
0.948159	0.2055	0.049673	-0.94637
0.896062	0.240439	-0.19597	-1.82426
0.92998	0.193163	0.13425	-0.06651
0.899638	0.1791	-0.14992	8.925604
0.883742	0.274169	0.138652	-14.2453
0.867462	0.213082	0.274879	-7.47895
0.936554	0.196758	-0.18865	5.970558
0.894003	0.257095	0.265481	-9.67044
0.942718	0.194706	0.073956	0.487422
0.90405	0.235975	-0.09791	-3.42436
0.91257	0.275899	0.158513	-14.4163
0.920666	0.187635	-0.05209	5.166909
0.948148	0.188351	0.011598	3.310533
NA	NA	NA	-75.0619
NA	NA	NA	72.85796



Test Statistic value

Regional Pool values:

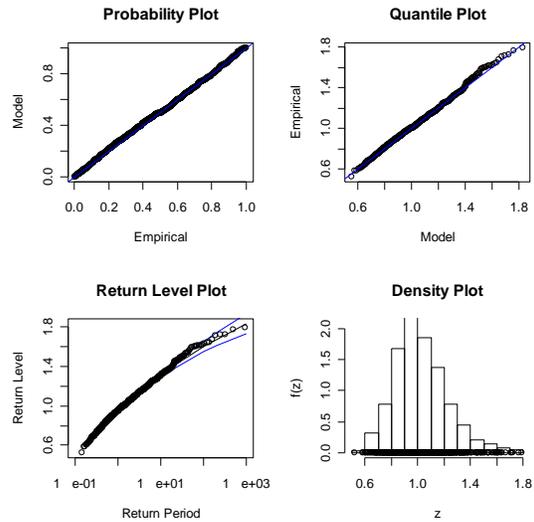
u	a	k	l
0.915434	0.224929	0.069214	-111.491

Chi Square Test Statistic with 24 sites in total, $\nu = 69$ @ 5% significance = 89.39

Test Statistic = 73, therefore homogeneity is likely.

Southern Scotland – 1 Day

u	a	k	l
0.967236	0.164928	0.046388	7.490773
0.952565	0.119673	0.014863	21.7669
0.935666	0.194543	-0.17986	6.269138
0.932005	0.195289	-0.08032	4.011154
0.937603	0.182905	-0.33836	11.87947
0.95449	0.189428	-0.16183	6.455731
0.961783	0.186087	-0.18036	8.066889
0.945557	0.201476	-0.1454	3.898922
0.923371	0.19836	-0.15127	5.002378
0.931499	0.193647	-0.23466	7.754474
0.946396	0.167676	-0.05898	9.717948
0.913258	0.173311	-0.28985	13.58012
0.903742	0.193253	0.013344	2.275558
0.918037	0.201063	-0.18014	5.259407
0.938475	0.160237	-0.14715	13.53653
0.92552	0.169988	-0.17466	11.67853
0.940065	0.151893	-0.20565	16.30748
0.959672	0.168349	0.175777	4.004063
0.967778	0.109926	0.001175	25.37232
0.922052	0.171969	0.11941	4.461126
0.947594	0.153373	0.038624	10.12649
0.942261	0.164014	-0.21747	13.56026
0.955832	0.147576	-0.06384	14.74242
0.930203	0.164669	0.008535	8.750192
0.964896	0.205793	-0.27637	6.002971
NA	NA	NA	241.9712
NA	NA	NA	76.76215



Test Statistic value

Regional Pool values:

u	a	k	l
0.938547	0.177985	-0.09821	203.5902

Chi Square Test Statistic with 25 sites in total, $\nu = 72$ @ 5% significance = 92.80

Test Statistic = 76, therefore homogeneity is likely.

Appendix 3

This page is intentionally blank.

Appendix 3 – Sample of ‘R’ routine for Homogeneity testing

Portion of the fitting routine for model 1:

```
# Model 1.      ## Stationary - GEV - parameter estimation
par2 <- sqrt(6 * var(xdat))/pi
par1 <- mean(xdat) - 0.57722 * par2      # 0.577 = euler's constant
par3 <- 0.1
mu <- mumat %>% (a[1])      # parameter 1 - Location
Sc <- Scmat %>% (a[seq(2, length=1)])  # parameter 2 - Scale
xi <- ximat %>% (a[seq(3, length=1)])  # parameter 3 - Shape
y <- (xdat - mu)/Sc
y <- 1 + xi * y
```

Portion of the fitting routine for model 2:

```
# Model 2. ## GEV - Test for trend in Scale
in4 <- 0.1
in3 <- 0.01
in2 <- sqrt(6 * var(xdat))/pi
in1 <- mean(xdat) - 0.57722 * in2 # 0.577 = euler's constant
mu <- mumat %>% (a[1])      # parameter 1 – Location
B <- Bmat %>% (a[seq(2, length=1)])  # the Scale parameter starting value.
a <- amat %>% (a[seq(3, length=1)])  # Incremental increase (trend) in the Scale Parameter
xi <- xmat %>% (a[seq(4, length=1)])  # parameter 3 – Shape,  $\xi$ 
y <- (xdat - mu)/(B + (a*ydat))
y <- 1 + xi * y
```

ydat – is the time index for the data set.

The covariate is a time dependent parameter which is associated with each of the GEV parameters in turn.

This page is intentionally blank.

Appendix 4

This page is intentionally blank.

Appendix 4 - Statistical Terminology and Concepts

1. Basics of the technique

- (1) Evaluate past data using statistical methods and probabilistic approaches to prepare a probability distribution curve, i.e. discharge-probability relationship, **(Q vs. P)_{past}**. The curve indicates the magnitude of a flood (peak flow) or drought (lowflow) for a given return period (frequency) or a cumulative probability of occurrence.
- (2) Predict future events, assuming they follow the same laws of probability.

$$\mathbf{(Q\ vs.\ P)_{future} = (Q\ vs.\ P)_{past}}$$

2. Applications

Flood (peak flow), drought (lowflow), rainfall, evaporation, and infiltration.

3. Uses

- (1) Statistics and stochastic approaches
- (2) Frequency (probability) analysis approaches - to find the magnitude and/or frequency of a given event
- (3) Economic evaluation of expected benefits and costs

4. Laws of probability

(1) Total probability always equals to 1.0:

$$P(-\infty < X < +\infty) = \sum_{i=1}^N p(x_i) = 1 \quad \text{for discrete variable}$$

$$P(-\infty < X < +\infty) = \int_{-\infty}^{+\infty} p_x(x) dx = 1 \quad \text{for continuous variable}$$

(2) Probability that an event occurs is 1 minus probability that it does not occur:

$$P = 1 - (1 - P)$$

(3) Joint probability--probability that event A and B both occur is: $P(A \cap B) = P(A) \cdot P(B)$, assuming A and B are statistically independent.

(4) Conditional probability--probability that event B occurs given that event A has already occurred is $P(B/A) = P(A \cap B)/P(A)$.

5. Terminology and concepts

Random variable: A variable governed by a probability distribution function. That is, the value obtained is somewhat dependent on probability. Flood discharge is a random variable. Two types of variables: discrete (sample) and continuous (population).

Variate: An individual observation or value of a variable. A flood peak flow is known as a variate.

Time series (sample): An array of variates, representing a sample of population of peak flows (discharges) recorded in the past and to be observed in the future at the study site. A collection of discharge representing the process.

Sample space: All possible values, zero to infinity for flood discharge: $0 < Q < +\infty$

Event: A subset of the sample space, X. The probability that event $(X \geq x)$ occurs is $P(X \geq x)$, where x is a assigned value of the variable and X is values which occur.

Frequency: The number of items in a class (number of occurrence of a variate) within the entire data base. Plotted in a frequency histogram (bar chart).

Relative frequency: Frequency divided by the total number of items in all classes:

$$f_i = \frac{n_i}{N}, \quad \text{where } n_i \text{ is the number of items in the } i^{\text{th}} \text{ class.}$$

Probability: $p(x_i)$: Relative frequency when N tends to infinity, i.e.,

$$p(x_i) = \lim_{N \rightarrow \infty} \frac{n_i}{N}, \quad 0 \leq p(x_i) \leq 1$$

Probability density function (PDF): Theoretical (mathematical) distribution functions of probability density: *normal, Pearson, extreme value and log-normal*. PDF's are very convenient because of known solutions. Its integration gives the cumulative probability.

$$p_x(x) = \frac{dF(x)}{dx}, \quad p_x(x) = \lim_{\substack{N \rightarrow \infty \\ \Delta x \rightarrow 0}} \frac{f(x)}{\Delta x}$$

Probability distribution function or Cumulative distribution function (CDF): Integration of PDF (*from continuous population*)

Nonexceedance probability:

$$P(X \leq x) = \int_{-\infty}^x p_x(x) dx = F(x)$$

Exceedance probability:

$$P(X \geq x) = \int_x^{+\infty} p_x(x) dx = 1 - F(x)$$

6. Statistical parameters

PDF's are defined by population parameters which are estimated from sample data.

There are three main categories:

1. Central tendency, i.e., expected value (mean) – 1st moment about the mean;
2. Variability, i.e., variance and standard deviation – 2nd moment about the mean;
and,
3. Symmetry, i.e., coefficient of skewness – 3rd moment about the mean.

Appendix 5

This page is intentionally blank.

Appendix 5 – Complete routines for use within ‘R’

A5.1 - Introduction

See <http://www.r-project.org/> for a complete and free download of this software.

What follows are two of the key routines. Prior to each of these is an explanation of how to use them including any changes/setup requirements that will be required.

A5.2 – Explanation of Routine 1

The routine needs to be copied and pasted into the ‘R’ Environment. Within ‘R’ the working directory must be set to that which contains all of the Annual Maxima files and location reference files. Sites can be added to the list but this will require annual maxima data files being generated, where missing data is represented as ‘NA’, currently all annual maxima files are of length 40 (years). Additionally the spatial location reference files will have to be updated, ensuring the new data is located in the correct position.

At the command prompt, within the ‘R’ environment, call the routine by entering the name followed by open and closed brackets, i.e. `Forge()`.

This will result in the default analysis, which is currently focused on site 41 (Cambridge), having a pooling group of 5 sites and a confidence of 0.95, or 95% (applied to the homogeneity testing).

To change the default analysis there is a text file called ‘Forge Site Ref’ listing all of the sites currently contained within the routine and the supporting files. To change the number of sites, simply choose the number you require based upon desired pooling group size. The final option relates to the probability associated with the homogeneity test, 0.95 gives a 5% significance level, meaning the associated error equates to 5 incorrect results for each 100 tests. A modified analysis might take the form: `Forge(30, 10, 0.95)`

The routine starts by standardising all the of the annual maxima data using the median value at each site. The median annual maxima rainfall depth at the focal site is then used to convert the pooling group growth curve into equivalent rainfall depth estimates for the focal site. The pooling group is calculated based upon spatial location relative to the pooling focal site. A separation (km) matrix is produced from which it is possible to rank in ascending order the separation of all sites from the focal site. The user defines how many sites are included, up to a current maximum of 179 sites for Great Britain.

Each individual site has its GEV parameters estimated using MLE. Then all of the selected sites are pooled and fitted too. This allows homogeneity testing of each site within the pooling group.

The growth curve is multiplied by the R_{med} value for the focal site, allowing rainfall depth estimates to be generated for a number of quantiles. The L-Moment estimate is included for comparison. Finally the confidence interval is calculated and added to the graph, this displays the likely range that the estimate falls in and must be used to highlight the uncertainty associated with the estimate. As the user will discover, the uncertainty is negligible for lesser return periods, but increases considerably for the rarer return period events. This tool also demonstrates that the uncertainty reduces with increased pooling group size.

It is believed that this routine would be useful in the classroom. It can be used to aid comprehension by demonstrating graphically the topics that it covers.

A5.3 - Routine 1

```
Forge <- function(focus=41, len=5, conf=0.95) {  
  
  lines=40      #reducing this value would shorten the length of the time series  
  library(ismev) # library commands required to provide pre-written routines  
  library(stats)  
  library(base)  
  
  ## SEE  
  boxly <- c(scan("boxly.txt", nlines = lines, na.strings = "NA"))  
  dartf <- c(scan("dartf.txt", nlines = lines, na.strings = "NA"))  
  eastb <- c(scan("eastb.txt", nlines = lines, na.strings = "NA"))  
  evert <- c(scan("evert.txt", nlines = lines, na.strings = "NA"))  
  faver <- c(scan("faver.txt", nlines = lines, na.strings = "NA"))  
  gatwk <- c(scan("gatwk.txt", nlines = lines, na.strings = "NA"))  
  hastg <- c(scan("hastg.txt", nlines = lines, na.strings = "NA"))  
  heath <- c(scan("heath.txt", nlines = lines, na.strings = "NA"))  
  kewbg <- c(scan("kewbg.txt", nlines = lines, na.strings = "NA"))  
  Ingst <- c(scan("Ingst.txt", nlines = lines, na.strings = "NA"))  
  #10  
  manst <- c(scan("manst.txt", nlines = lines, na.strings = "NA"))  
  marty <- c(scan("marty.txt", nlines = lines, na.strings = "NA"))  
  oxfor <- c(scan("oxfor.txt", nlines = lines, na.strings = "NA"))  
  rotha <- c(scan("rotha.txt", nlines = lines, na.strings = "NA"))  
  shoeb <- c(scan("shoeb.txt", nlines = lines, na.strings = "NA"))  
  stans <- c(scan("stans.txt", nlines = lines, na.strings = "NA"))  
  winds <- c(scan("winds.txt", nlines = lines, na.strings = "NA"))  
  wisle <- c(scan("wisle.txt", nlines = lines, na.strings = "NA"))  
  writt <- c(scan("writt.txt", nlines = lines, na.strings = "NA"))  
}
```

```
wyedr <- c(scan("wyedr.txt", nlines = lines, na.strings = "NA"))  
#20
```

```
## SWE
```

```
BOSCM <- c(scan("BOSCM.txt", nlines = lines, na.strings = "NA"))  
BUDEE <- c(scan("BUDEE.txt", nlines = lines, na.strings = "NA"))  
CHELT <- c(scan("CHELT.txt", nlines = lines, na.strings = "NA"))  
CWMYS <- c(scan("CWMYS.txt", nlines = lines, na.strings = "NA"))  
DALEF <- c(scan("DALEF.txt", nlines = lines, na.strings = "NA"))  
GOGER <- c(scan("GOGER.txt", nlines = lines, na.strings = "NA"))  
HURNN <- c(scan("HURNN.txt", nlines = lines, na.strings = "NA"))  
LNGAS <- c(scan("LNGAS.txt", nlines = lines, na.strings = "NA"))  
LYNEH <- c(scan("LYNEH.txt", nlines = lines, na.strings = "NA"))  
LYONS <- c(scan("LYONS.txt", nlines = lines, na.strings = "NA"))  
#10
```

```
PLYMO <- c(scan("PLYMO.txt", nlines = lines, na.strings = "NA"))  
PRESW <- c(scan("PRESW.txt", nlines = lines, na.strings = "NA"))  
RHOOS <- c(scan("RHOOS.txt", nlines = lines, na.strings = "NA"))  
SIDMT <- c(scan("SIDMT.txt", nlines = lines, na.strings = "NA"))  
STANN <- c(scan("STANN.txt", nlines = lines, na.strings = "NA"))  
SWANS <- c(scan("SWANS.txt", nlines = lines, na.strings = "NA"))  
TRAWS <- c(scan("TRAWS.txt", nlines = lines, na.strings = "NA"))  
TRENG <- c(scan("TRENG.txt", nlines = lines, na.strings = "NA"))  
USKKK <- c(scan("USKKK.txt", nlines = lines, na.strings = "NA"))  
YEOTN <- c(scan("YEOTN.txt", nlines = lines, na.strings = "NA"))  
#20
```

```
## CEE
```

```
cambn <- c(scan("cambn.txt", nlines = lines, na.strings = "NA"))  
colti <- c(scan("colti.txt", nlines = lines, na.strings = "NA"))  
cranw <- c(scan("cranw.txt", nlines = lines, na.strings = "NA"))
```

```

elmdo <- c(scan("elmdo.txt", nlines = lines, na.strings = "NA"))
eyebr <- c(scan("eyebr.txt", nlines = lines, na.strings = "NA"))
hulll <- c(scan("hulll.txt", nlines = lines, na.strings = "NA"))
keele <- c(scan("keele.txt", nlines = lines, na.strings = "NA"))
lowes <- c(scan("lowes.txt", nlines = lines, na.strings = "NA"))
newpo <- c(scan("newpo.txt", nlines = lines, na.strings = "NA"))
persh <- c(scan("persh.txt", nlines = lines, na.strings = "NA"))
#10
santd <- c(scan("santd.txt", nlines = lines, na.strings = "NA"))
shawb <- c(scan("shawb.txt", nlines = lines, na.strings = "NA"))
silso <- c(scan("silso.txt", nlines = lines, na.strings = "NA"))
skegn <- c(scan("skegn.txt", nlines = lines, na.strings = "NA"))
strat <- c(scan("strat.txt", nlines = lines, na.strings = "NA"))
suttb <- c(scan("suttb.txt", nlines = lines, na.strings = "NA"))
terri <- c(scan("terri.txt", nlines = lines, na.strings = "NA"))
warso <- c(scan("warso.txt", nlines = lines, na.strings = "NA"))
watti <- c(scan("watti.txt", nlines = lines, na.strings = "NA"))
welle <- c(scan("welle.txt", nlines = lines, na.strings = "NA"))
#20
wobur <- c(scan("wobur.txt", nlines = lines, na.strings = "NA"))

## NWE
apple <- c(scan("apple.txt", nlines = lines, na.strings = "NA"))
askhl <- c(scan("askhl.txt", nlines = lines, na.strings = "NA"))
bodnt <- c(scan("bodnt.txt", nlines = lines, na.strings = "NA"))
buxto <- c(scan("buxto.txt", nlines = lines, na.strings = "NA"))
carli <- c(scan("carli.txt", nlines = lines, na.strings = "NA"))
daleh <- c(scan("daleh.txt", nlines = lines, na.strings = "NA"))
dougl <- c(scan("dougl.txt", nlines = lines, na.strings = "NA"))
dunhm <- c(scan("dunhm.txt", nlines = lines, na.strings = "NA"))
formb <- c(scan("formb.txt", nlines = lines, na.strings = "NA"))

```

```

glasc <- c(scan("glasc.txt", nlines = lines, na.strings = "NA"))
#10
llanu <- c(scan("llanu.txt", nlines = lines, na.strings = "NA"))
logge <- c(scan("logge.txt", nlines = lines, na.strings = "NA"))
lymep <- c(scan("lymep.txt", nlines = lines, na.strings = "NA"))
newto <- c(scan("newto.txt", nlines = lines, na.strings = "NA"))
penyf <- c(scan("penyf.txt", nlines = lines, na.strings = "NA"))
ringw <- c(scan("ringw.txt", nlines = lines, na.strings = "NA"))
stmic <- c(scan("stmic.txt", nlines = lines, na.strings = "NA"))
sunny <- c(scan("sunny.txt", nlines = lines, na.strings = "NA"))
sutth <- c(scan("sutth.txt", nlines = lines, na.strings = "NA"))
thirl <- c(scan("thirl.txt", nlines = lines, na.strings = "NA"))
#20
valle <- c(scan("valle.txt", nlines = lines, na.strings = "NA"))
vivod <- c(scan("vivod.txt", nlines = lines, na.strings = "NA"))
voela <- c(scan("voela.txt", nlines = lines, na.strings = "NA"))
worth <- c(scan("worth.txt", nlines = lines, na.strings = "NA"))
#24

## NEE
turnh <- c(scan("turnh.txt", nlines = lines, na.strings = "NA"))
edinb <- c(scan("edinb.txt", nlines = lines, na.strings = "NA"))
birds <- c(scan("birds.txt", nlines = lines, na.strings = "NA"))
bushh <- c(scan("bushh.txt", nlines = lines, na.strings = "NA"))
dunbr <- c(scan("dunbr.txt", nlines = lines, na.strings = "NA"))
locht <- c(scan("locht.txt", nlines = lines, na.strings = "NA"))
haydo <- c(scan("haydo.txt", nlines = lines, na.strings = "NA"))
barnd <- c(scan("barnd.txt", nlines = lines, na.strings = "NA"))
bradf <- c(scan("bradf.txt", nlines = lines, na.strings = "NA"))
cockl <- c(scan("cockl.txt", nlines = lines, na.strings = "NA"))
#10

```

```

birst <- c(scan("birst.txt", nlines = lines, na.strings = "NA"))
durhm <- c(scan("durhm.txt", nlines = lines, na.strings = "NA"))
gantn <- c(scan("gantn.txt", nlines = lines, na.strings = "NA"))
leemg <- c(scan("leemg.txt", nlines = lines, na.strings = "NA"))
highm <- c(scan("highm.txt", nlines = lines, na.strings = "NA"))
bramm <- c(scan("bramm.txt", nlines = lines, na.strings = "NA"))
askbr <- c(scan("askbr.txt", nlines = lines, na.strings = "NA"))
lockw <- c(scan("lockw.txt", nlines = lines, na.strings = "NA"))
mulgv <- c(scan("mulgv.txt", nlines = lines, na.strings = "NA"))
scamp <- c(scan("scamp.txt", nlines = lines, na.strings = "NA"))
#20
cawoo <- c(scan("cawoo.txt", nlines = lines, na.strings = "NA"))
whitb <- c(scan("whitb.txt", nlines = lines, na.strings = "NA"))

## NS
ardng <- c(scan("ardng.txt", nlines = lines, na.strings = "NA"))
assyn <- c(scan("assyn.txt", nlines = lines, na.strings = "NA"))
bridg <- c(scan("bridg.txt", nlines = lines, na.strings = "NA"))
capew <- c(scan("capew.txt", nlines = lines, na.strings = "NA"))
cassl <- c(scan("cassl.txt", nlines = lines, na.strings = "NA"))
fairb <- c(scan("fairb.txt", nlines = lines, na.strings = "NA"))
fanni <- c(scan("fanni.txt", nlines = lines, na.strings = "NA"))
fasna <- c(scan("fasna.txt", nlines = lines, na.strings = "NA"))
ferst <- c(scan("ferst.txt", nlines = lines, na.strings = "NA"))
frtag <- c(scan("frtag.txt", nlines = lines, na.strings = "NA"))
#10
hoyps <- c(scan("hoyps.txt", nlines = lines, na.strings = "NA"))
inver <- c(scan("inver.txt", nlines = lines, na.strings = "NA"))
irhum <- c(scan("irhum.txt", nlines = lines, na.strings = "NA"))
kinlc <- c(scan("kinlc.txt", nlines = lines, na.strings = "NA"))
kirkw <- c(scan("kirkw.txt", nlines = lines, na.strings = "NA"))

```

```

lchrn <- c(scan("lchrn.txt", nlines = lines, na.strings = "NA"))
lerwk <- c(scan("lerwk.txt", nlines = lines, na.strings = "NA"))
mullw <- c(scan("mullw.txt", nlines = lines, na.strings = "NA"))
rhuba <- c(scan("rhuba.txt", nlines = lines, na.strings = "NA"))
storn <- c(scan("storn.txt", nlines = lines, na.strings = "NA"))
#20
tiree <- c(scan("tiree.txt", nlines = lines, na.strings = "NA"))
ulvah <- c(scan("ulvah.txt", nlines = lines, na.strings = "NA"))
wicka <- c(scan("wicka.txt", nlines = lines, na.strings = "NA"))

## ES
aberd <- c(scan("aberd.txt", nlines = lines, na.strings = "NA"))
balbr <- c(scan("balbr.txt", nlines = lines, na.strings = "NA"))
balmo <- c(scan("balmo.txt", nlines = lines, na.strings = "NA"))
braem <- c(scan("braem.txt", nlines = lines, na.strings = "NA"))
cardy <- c(scan("cardy.txt", nlines = lines, na.strings = "NA"))
clatt <- c(scan("clatt.txt", nlines = lines, na.strings = "NA"))
craib <- c(scan("craib.txt", nlines = lines, na.strings = "NA"))
cromb <- c(scan("cromb.txt", nlines = lines, na.strings = "NA"))
drumm <- c(scan("drumm.txt", nlines = lines, na.strings = "NA"))
dycee <- c(scan("dycee.txt", nlines = lines, na.strings = "NA"))
#10
elgin <- c(scan("elgin.txt", nlines = lines, na.strings = "NA"))
faska <- c(scan("faska.txt", nlines = lines, na.strings = "NA"))
frand <- c(scan("frand.txt", nlines = lines, na.strings = "NA"))
geani <- c(scan("geani.txt", nlines = lines, na.strings = "NA"))
glnqu <- c(scan("glnqu.txt", nlines = lines, na.strings = "NA"))
glnth <- c(scan("glnth.txt", nlines = lines, na.strings = "NA"))
kinls <- c(scan("kinls.txt", nlines = lines, na.strings = "NA"))
lchl v <- c(scan("lchl v.txt", nlines = lines, na.strings = "NA"))
lethn <- c(scan("lethn.txt", nlines = lines, na.strings = "NA"))

```

```

leuch <- c(scan("leuch.txt", nlines = lines, na.strings = "NA"))
#20
mylne <- c(scan("mylne.txt", nlines = lines, na.strings = "NA"))
rocho <- c(scan("rocho.txt", nlines = lines, na.strings = "NA"))
strer <- c(scan("strer.txt", nlines = lines, na.strings = "NA"))
tilli <- c(scan("tilli.txt", nlines = lines, na.strings = "NA"))

## SS
auchi <- c(scan("auchi.txt", nlines = lines, na.strings = "NA"))
benmr <- c(scan("benmr.txt", nlines = lines, na.strings = "NA"))
black <- c(scan("black.txt", nlines = lines, na.strings = "NA"))
blyth <- c(scan("blyth.txt", nlines = lines, na.strings = "NA"))
bowhl <- c(scan("bowhl.txt", nlines = lines, na.strings = "NA"))
buted <- c(scan("buted.txt", nlines = lines, na.strings = "NA"))
carnw <- c(scan("carnw.txt", nlines = lines, na.strings = "NA"))
dumfr <- c(scan("dumfr.txt", nlines = lines, na.strings = "NA"))
dunsd <- c(scan("dunsd.txt", nlines = lines, na.strings = "NA"))
eskdl <- c(scan("eskdl.txt", nlines = lines, na.strings = "NA"))
#10
forre <- c(scan("forre.txt", nlines = lines, na.strings = "NA"))
garls <- c(scan("garls.txt", nlines = lines, na.strings = "NA"))
glass <- c(scan("glass.txt", nlines = lines, na.strings = "NA"))
glnkn <- c(scan("glnkn.txt", nlines = lines, na.strings = "NA"))
glnle <- c(scan("glnle.txt", nlines = lines, na.strings = "NA"))
irvin <- c(scan("irvin.txt", nlines = lines, na.strings = "NA"))
islay <- c(scan("islay.txt", nlines = lines, na.strings = "NA"))
mugdk <- c(scan("mugdk.txt", nlines = lines, na.strings = "NA"))
ormsy <- c(scan("ormsy.txt", nlines = lines, na.strings = "NA"))
paisl <- c(scan("paisl.txt", nlines = lines, na.strings = "NA"))
#20
penwh <- c(scan("penwh.txt", nlines = lines, na.strings = "NA"))

```

```
pulla <- c(scan("pulla.txt", nlines = lines, na.strings = "NA"))
roths <- c(scan("roths.txt", nlines = lines, na.strings = "NA"))
skipn <- c(scan("skipn.txt", nlines = lines, na.strings = "NA"))
threv <- c(scan("threv.txt", nlines = lines, na.strings = "NA"))
```

```
## boxly omitted in reg' analysis due to non-positive sigma.
## elmdo omitted in reg' analysis due to non-positive sigma.
## askhl omitted in reg' analysis due to non-positive sigma.
```

```
data=cbind(boxly, dartf, eastb, evert, faver, gatwk, hastg, heath,
kewbg, Ingst, manst, marty, oxfor, rotha, shoeb, stans,
winds, wisle, writt, wyedr, BOSCM, BUDEE, CHELT, CWMYS,
DALEF, GOGER, HURNN, LNGAS, LYNEH, LYONS, PLYMO, PRESW,
RHOOS, SIDMT, STANN, SWANS, TRAWS, TRENG, USKKK, YEOTN,
cambn, colti, cranw, elmdo, eyebr, hulll, keele, lowes,
newpo, persh, santd, shawb, silso, skegn, strat, suttb,
terri, warso, watti, welle, wobur, apple, askhl, bodnt,
buxto, carli, daleh, dougl, dunhm, formb, glasc, llanu,
logge, lymep, newto, penyf, ringw, stmic, sunny, sutth,
thirl, valle, vivod, voela, worth, turnh, edinb, birds,
bushh, dunbr, locht, haydo, barnd, bradf, cockl, birst,
durhm, gantn, leemg, highm, bramm, askbr, lockw, mulgy,
scamp, cawoo, whitb, ardnng, assyn, bridg, capew, cassl,
fairb, fanni, fasna, ferst, frtag, hoyps, inver, irhum,
kinlc, kirkw, lchn, lerwk, mullw, rhuba, storn, tiree,
ulvah, wicka, aberd, balbr, balmo, braem, cardy, clatt,
craib, cromb, drumm, dycee, elgin, faska, frand, geani,
glnqu, glnth, kins, lchlv, lethn, leuch, mylne, rocho,
strer, tilli, auchi, benmr, black, blyth, bowhl, buted,
carnw, dumfr, dunsd, eskdl, forre, garls, glass, glnkn,
```

glnle, irvin, islay, mugdk, ormsy, paisl, penwh, pulla,
roths, skipn, threv)

```
k <- dim(data)[1]# number of rows
c <- dim(data)[2]# number of col's
L <- c(1:dim(data)[1])# seq' from 1 to total number of rows (or years / length of data set)
n <- c(1:dim(data)[2])# seq' from 1 to total number of col's (or number of sites)
p <- n
q <- n
av <- matrix(nrow = c, ncol = 1)
stand <- matrix(nrow = k, ncol = c)
```

```
  for (s in c(n)) { # col'n loop number for output matrix, read and work on each
col'n in turn.
```

```
  site <- c(data[,s])
  av[s] <- median(site, na.rm=TRUE)
  stand[,s] <- site/av[s] # Use (site/av) if standardizing data.
  }
```

Distance Calcs

```
site <- c(scan("ref.txt"))
east <- c(scan("east.txt"))
north <- c(scan("north.txt"))
ddata <- cbind(site, east, north)
as.matrix(ddata)
```

```
a <- dim(ddata) # gives [rows x col's]
distance <- matrix(nrow = c, ncol = c)
  for (s in c(p)) { # col'n loop number for output matrix
    for (t in c(q)) { # row loop number for output matrix
```

```

        distance[t,s]      <-      (((ddata[s,2]-ddata[t,2])^2)+((ddata[s,3]-
ddata[t,3])^2))^0.5)/1000

    }
}

```

```

diss <- distance[,focus]
tempdata <- cbind(p,diss)# vector with seq' and col'n 'focus'
tempdata[is.na(tempdata[,2]),2]=999
tempdata=tempdata[sort(tempdata[,2],index=T)$ix,]#   arranges   tempdata   's'   into
ascending order
tdata=tempdata
tdata[tdata[,2]==999,2]=NA
disdata <- tdata

```

```

## Data ranked in ascending order, with Col'n 1 containing the site ref' number and col'n
2 the distance from the
## chosen 'focal' site. Standard = site 41 Cambridge.
# If we choose the first 5 rows and save the site numbers to vector, we can use these to
call data from the other
# matrices without modification to them.

```

```

fone <- c(tdata[1:len,1])      # Number of sites chosen by user
test <- fone
seq <- c(1:length(test))
bb <- length(test)
zz <- 1

```

```

### Model 2:
#Fit the GEV parameters for each individual site and record the log-likelihood (llh)

```

```
gev2 <- matrix(nrow = bb+3, ncol = 4)
```

```
  for (s in c(seq)) { # col'n loop number for output matrix, read and work on each  
col'n in turn.
```

```
  b <- fone[s]  
  x <- c(stand[,b])  
  x <- sort(x)  
  y <- gev.fit(x)
```

```
  gev2[s,1] <- y$mle[1]  
  gev2[s,2] <- y$mle[2]  
  gev2[s,3] <- -1 * y$mle[3]  
  gev2[s,4] <- -1 * (y$nullh)  
  }
```

```
### Model 1:
```

```
# Starting with 5 sites (default) and using the focus site and minimum separation  
selection technique.
```

```
# standardise site annual maxima by dividing by the at site sample median  
#####
```

```
pool <- c(stand[,fone])  
pool <- sort(pool)  
z <- gev.fit(pool)
```

```
x <- pool  
n <- length(x)  
l1 <- mean(x)  
jb1 <- c(2:n)
```

```

z1 <- c((jb1-1)/(n-1))
y1 <- c(x[jb1])
q1 <- c(z1*y1)
b01 <- sum(q1)
b1 <- b01/n
l2 <- (2*b1)-11
jb2 <- c(3:n)
z2 <- c(((jb2-1)*(jb2-2))/((n-1)*(n-2)))
y2 <- c(x[jb2])
q2 <- c(z2*y2)
b02 <- sum(q2)
b2 <- b02/n
l3 <- (6*b2)-(6*b1)+11
jb3 <- c(4:n)
z3 <- c(((jb3-1)*(jb3-2)*(jb3-3))/((n-1)*(n-2)*(n-3)))
y3 <- c(x[jb3])
q3 <- c(z3*y3)
b03 <- sum(q3)
b3 <- b03/n
l4 <- (20*b3)-(30*b2)+(12*b1)-11
LCV <- l2/l1
ISKew <- l3/l2
IKurt <- l4/l2
LM <- c(l1, l2, l3, l4, LCV, ISKEW, IKurt)
names(LM) <- c("L1", "L2", "L3", "L4", "L-CV", "L-Skewness", "L-kurtosis")
c1 <- (((2*b1)-11)/((3*b2)-11))-(log(2) / log(3))
kappa <- (7.8590*c1)+(2.9554*(c1^2))
shape <- kappa
a <- (((2*(b1))-11)*kappa)/(gamma(1+kappa)*(1-(2^-kappa)))
Scale <- a
mu <- (11+(a*((gamma((1+kappa))-1)/kappa))

```

```
Loc <- mu
```

```
a<-z$mle
```

```
mat<-z$cov
```

```
dat<-z$data
```

```
eps <- 1e-06
```

```
a1 <- a
```

```
a2 <- a
```

```
a3 <- a
```

```
a1[1] <- a[1] + eps
```

```
a2[2] <- a[2] + eps
```

```
a3[3] <- a[3] + eps
```

```
f <- c(seq(0.01, 0.09, by = 0.01), 0.1, 0.2, 0.3, 0.4, 0.5,
```

```
0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 0.995, 0.999, 0.9999)
```

```
qfocl <- Loc + (Scale * (1-(-log(f))^(shape)))/shape
```

```
q <- gevq(a, 1 - f)
```

```
d1 <- (gevq(a1, 1 - f) - q)/eps
```

```
d2 <- (gevq(a2, 1 - f) - q)/eps
```

```
d3 <- (gevq(a3, 1 - f) - q)/eps
```

```
d <- cbind(d1, d2, d3)
```

```
v <- apply(d, 1, q.form, m = mat)
```

```
plot(-1/log(f), q, log = "x", type = "n", xlim = c(0.1, 10000),
```

```
ylim = c(min(dat, q), max(dat, q)), xlab = "Return Period (Years)",
```

```
ylab = "Standardised Rainfall Depth")
```

```
title("Focused Growth Curve Plot")
```

```
lines(-1/log(f), q, lty=1)
```

```
lines(-1/log(f), qfocl, col=2, lty=2)
```

```
lines(-1/log(f), q + 1.96 * sqrt(v), col = 4, lty=3) #Confidence interval (upper)
```

```
lines(-1/log(f), q - 1.96 * sqrt(v), col = 4, lty=3) #Confidence interval (lower)
```

```

points(-1/log((1:length(dat))/(length(dat) + 1)), sort(dat))

leg1x <- c(0.4,0.45)          #Legend Location, x
leg1y <- c(0.9*max(dat, q),0.9*max(dat, q)) #Legend Location, y

text((-log(-log(0.5))), 0.95*max(dat, q), 'rmed(mm) =')
#Legend for MLE parameter results
text((-log(-log(0.875))), 0.95*max(dat, q), av[focus])
lines(-log(-log(leg1x)), leg1y,lty=1)
text((-log(-log(0.9))), 0.9*max(dat, q), 'MLE / R.P. (mm)')
text((-log(-log(0.42))), 0.85*max(dat, q), 'u =')
text((-log(-log(0.5))), 0.85*max(dat, q), round(z$mle[1],3))
text((-log(-log(0.42))), 0.8*max(dat, q), 'a =')
text((-log(-log(0.5))), 0.8*max(dat, q), round(z$mle[2],3))
text((-log(-log(0.42))), 0.75*max(dat, q), 'k =')
text((-log(-log(0.5))), 0.75*max(dat, q), round(-1*z$mle[3],3))

leg2x <- c(0.4,0.45)
leg2y <- c(0.7*max(dat, q),0.7*max(dat, q))

#Legend for L-Moment parameter results
lines(-log(-log(leg2x)), leg2y,col=2,lty=2)
text((-log(-log(0.95))), 0.7*max(dat, q), 'L-Moments / R.P.(mm)')
text((-log(-log(0.42))), 0.65*max(dat, q), 'u =')
text((-log(-log(0.5))), 0.65*max(dat, q), round(Loc,3))
text((-log(-log(0.42))), 0.6*max(dat, q), 'a =')
text((-log(-log(0.5))), 0.6*max(dat, q), round(Scale,3))
text((-log(-log(0.42))), 0.55*max(dat, q), 'k =')
text((-log(-log(0.5))), 0.55*max(dat, q), round(shape,3))

# Quatile estimates, 10, 50 and 100-years, in mm ##

```

```

forge10L <- av[focus]*(Loc + (Scale * (1-(-log(0.9))^(shape)))/shape)
forge50L <- av[focus]*(Loc + (Scale * (1-(-log(0.98))^(shape)))/shape)
forge100L <- av[focus]*(Loc + (Scale * (1-(-log(0.99))^(shape)))/shape)

forge10ML <- av[focus]*(z$mle[1] + (z$mle[2] * (1-(-log(0.9))^(1*z$mle[3])))/(-
1*z$mle[3]))
forge50ML <- av[focus]*(z$mle[1] + (z$mle[2] * (1-(-log(0.98))^(1*z$mle[3])))/(-
1*z$mle[3]))
forge100ML <- av[focus]*(z$mle[1] + (z$mle[2] * (1-(-log(0.99))^(1*z$mle[3])))/(-
1*z$mle[3]))

text((-log(-log(0.85))), 0.65*max(dat, q), '10 Yr=')
text((-log(-log(0.99))), 0.65*max(dat, q), round(forge10L))
text((-log(-log(0.85))), 0.6*max(dat, q), '50 Yr=')
text((-log(-log(0.99))), 0.6*max(dat, q), round(forge50L))
text((-log(-log(0.85))), 0.55*max(dat, q), '100-year=')
text((-log(-log(0.99))), 0.55*max(dat, q), round(forge100L))
text((-log(-log(0.85))), 0.85*max(dat, q), '10 Yr=')
text((-log(-log(0.99))), 0.85*max(dat, q), round(forge10ML))
text((-log(-log(0.85))), 0.8*max(dat, q), '50 Yr=')
text((-log(-log(0.99))), 0.8*max(dat, q), round(forge50ML))
text((-log(-log(0.85))), 0.75*max(dat, q), '100-year=')
text((-log(-log(0.99))), 0.75*max(dat, q), round(forge100ML))

gev1 <- matrix(nrow = 1, ncol = 4)
gev1[1] <- z$mle[1]
gev1[2] <- z$mle[2]
gev1[3] <- -1 * z$mle[3]
gev1[4] <- -1 * (z$nlh)
gev2[(bb+1),4] <- sum(gev2[zz:bb,4])
gev2[(bb+2),4] <- 2*(gev2[(bb+1),4]-gev1[4])

```

```
gev2[(bb+3),4] <- qchisq(conf,df=((bb*3)-3))
```

```
chisq <- gev2[(bb+3),4]
```

```
tstat <- gev2[(bb+2),4]
```

```
ifelse(tstat < chisq, statement<-'Homogeneity', statement<-'Heterogeneity')
```

```
capture.output(gev1, file = "gevforge.txt")
```

```
capture.output(gev2, file = "regsites.txt")
```

```
text((-1/log(0.999)), 1, statement)
```

```
text((-1/log(0.995)),0.75, 'Number of sites in Pool = ')
```

```
text((-1/log(0.99975)), 0.75, len)
```

```
}
```