Australian Government
Bureau of Meteorology

HYDROLOGY REPORT SERIES

HRS Report No. 10

A PILOT STUDY TO EXPLORE METHODS FOR DERIVING DESIGN RAINFALLS FOR AUSTRALIA - PART 1

Hydrology Unit
Melbourne
June 2005

Australian Government

Bureau of Meteorology

HYDROLOGY REPORT SERIES

HRS Report No. 10

A PILOT STUDY TO EXPLORE METHODS FOR DERIVING DESIGN RAINFALLS FOR AUSTRALIA - PART 1

Dörte Jakob, Brian Taylor, Karin Xuereb

Hydrometeorological Advisory Service
www.bom.gov.au/hydro/has
Hydrology Unit
Melbourne
June 2005

Contents

List of Figures

List of Figures (continued)

Page

List of Tables

1 Introduction

A project to review the methods for the estimation of design rainfall commenced in November 2003 in the Hydrometeorology Advisory Service of the Bureau of Meteorology. This report will cover work done during the first year of this pilot study (up to and including December 2004).

1.1 Acknowledgements

This study greatly benefited from the support of colleagues within and outside the Bureau of Meteorology, only some of them can be named here.

Data were made available by Ray Maynard and Rod Dew (DNRM Queensland), and Darren Thompson and Andrew Brissett (DIPNR NSW). Mike Denham (University of Reading, UK) generously allowed us to use a library of functions he had written to perform univariate and multivariate PLSR. Rory Nathan (SKM) and Craig Thompson (NIWA, NZ) supplied us with copies of their software free of charge: GetDat and HIRDS respectively. In particular, we would like to thank Erwin Weinmann (Monash University), Geoff Bonnin (HDSC, US), Ray Canterford, Sri Srikanthan, Jim Elliott and Jeanette Meighen (Bureau of Meteorology) for useful discussions and suggestions. We would also like to thank Garry Moore and Cathy Beasley (Bureau of Meteorology) for their assistance in data analysis.

1.2 Background

Australian Rainfall and Runoff (ARR), published by the Institution of Engineers, Australia in 1987, (Institution of Engineers, 1987), includes recommended procedures for estimating intensity-frequency-duration (IFD) information for any location in Australia (Pilgrim 1987). These procedures were developed by the Bureau of Meteorology, which provides this information to clients using a system known as CDIRS (Computerised Design IFD Rainfall System) as a cost recovery service.

This development was undertaken in the 1980s and, since then, there has been a significant amount of additional data collected, as well as advances made in statistical analysis and computing capabilities. As a precursor to a possible review of ARR, a pilot study has been undertaken.

1.3 Scope

Ultimately, IFD relationships will be derived for any location within Australia (at a resolution of 0.025 degrees). In the pilot study, average recurrence intervals (ARI) from 1 to 100 years will be investigated. Higher ARI (up to 200 years) will be investigated to assess the robustness of the methods developed. Methods to derive design rainfall estimates for extreme frequencies are presented elsewhere.

Durations considered in this study range from 1 hour to 3 days. Durations below one hour are important in urban hydrology applications but these will have to be studied at a later time.

Revising the methods for design rainfall estimation is a major effort and might only be feasible as a state-by-state approach. The main objective of this pilot study is to develop valid methods for design rainfall estimation.

2 Outline

Figure 1 aims to give an overview of the approach to design rainfall estimation used in the pilot study. The second part of the section outlines the structure of the report. Terminology will be introduced in the relevant sections. Refer to the index on page 51 for help in locating keywords.

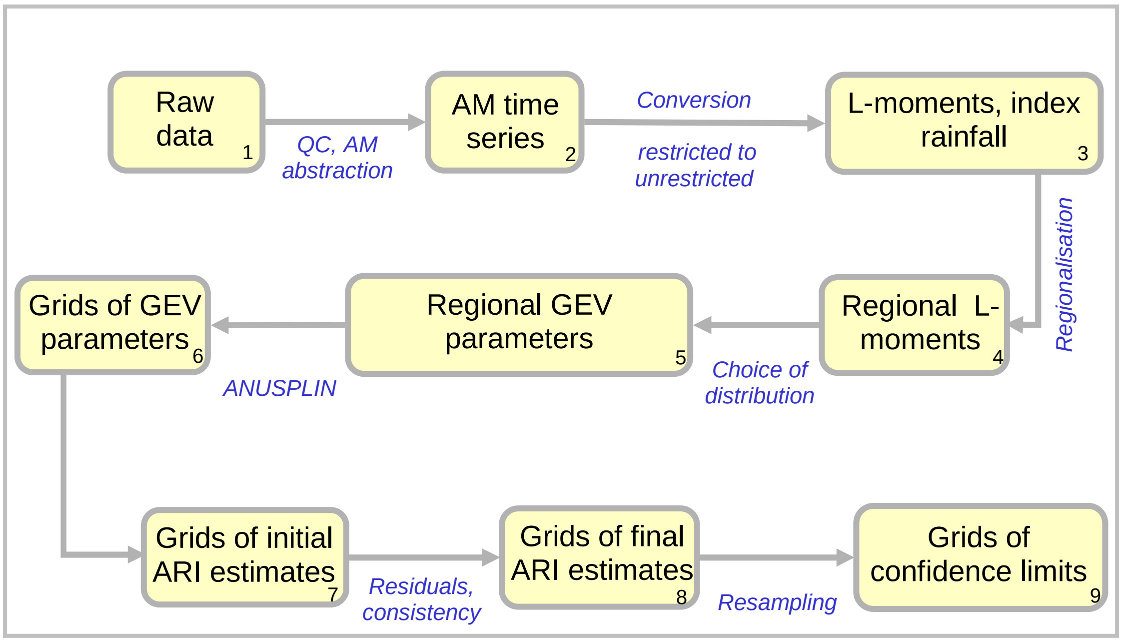

Figure 1 Approach to design rainfall estimation used in the pilot study

Section 3 describes the data used for the pilot study, including a description of the pilot area, data sources and approaches to quality controlling the data.

Section 4 deals with the revision of two adjustment factors - to convert from restricted to unrestricted durations; and to convert from an annual maximum to a partial duration scale.

Section 5 gives background on L-moments and LH-moments.

Section 6 describes the mapping of the index rainfall which is required as part of the regionalisation approach.

Section 7 discusses the choice of frequency distribution to be used for site and regional growth curves.

Section 8 contains information on methods for inferring statistics of sub-daily rainfall data from daily data.

Section 9 focuses on regionalisation. Two different approaches are introduced and their respective performance evaluated.

Section 10 summarises the results and recommendations.

Section 11 concludes the report with an extensive list of references.

Section 12 contains a glossary of technical terms used in this report along with definitions of the terms.

Section 13 contains a list of acronyms used in this report.

Section 14 is an appendix containing a listing of gauges used in this study.

3 Data

The pilot study area (Figure 2) extends from north of Brisbane to south of Byron Bay and 250 km inland. The area chosen for the pilot study is relatively small but interesting from a design rainfall point of view. The lighter shading in Figure 2 indicates a buffer zone, which is required to minimise boundary problems when deriving gridded surfaces (refer Section 6) and for regionalisation (refer Section 9).

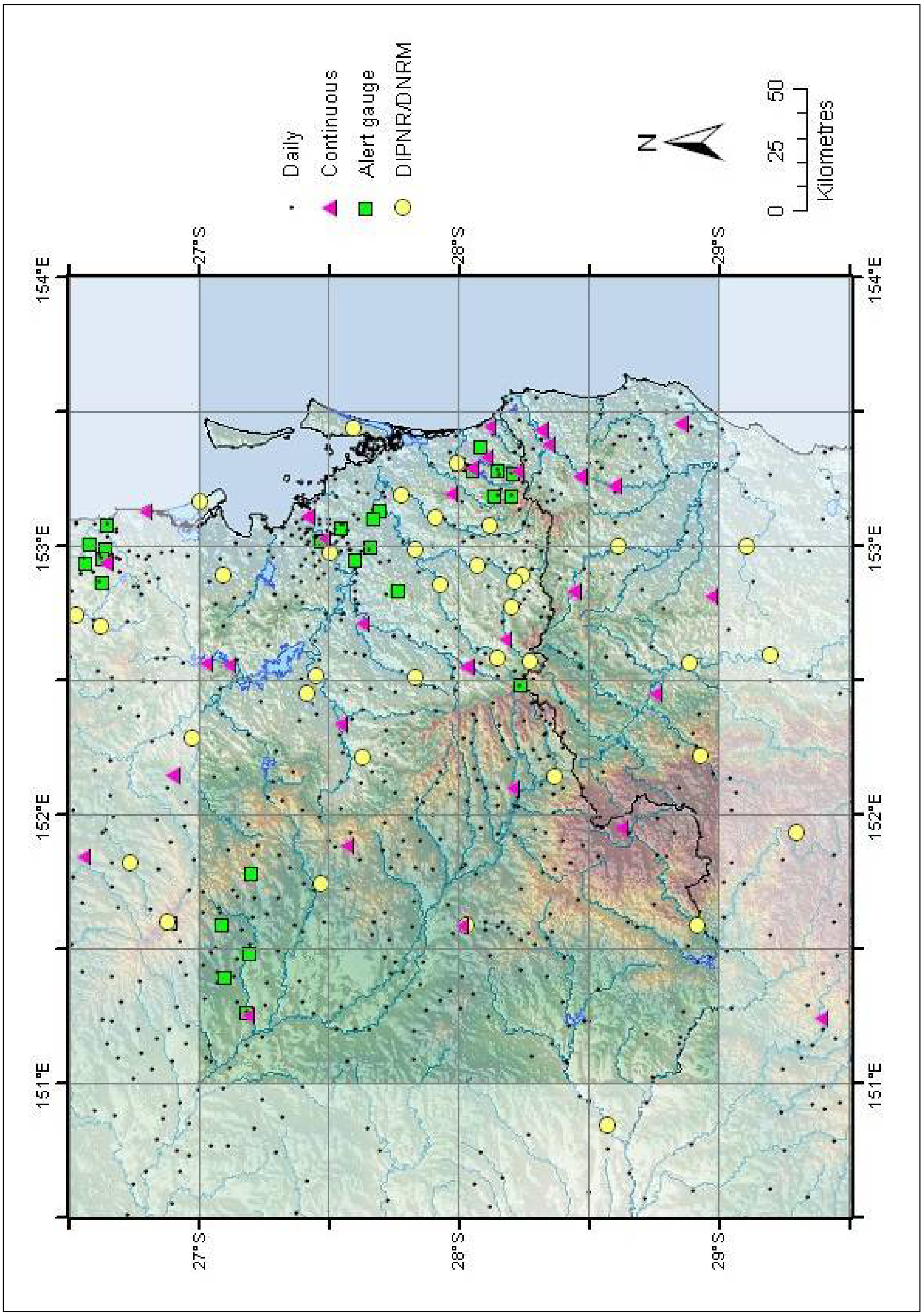

Figure 2 Pilot area showing the buffer zone (lighter shading) and gauges used in the study.

The network of daily gauges is quite dense. Overall 748 daily gauges (9:00am to 9:00am) with a minimum record length of 20 years are available; the typical record length is about 45 years. Far fewer gauges record rainfall for sub-daily intervals (see Figure 3). Continuous recording raingauges from the Bureau of Meteorology standard network sited within the pilot study area have a typical record length of about 25 years. ALERT gauges (Automated Local Evaluation in Real-Time) are part of the Bureau of Meteorology's flood warning network. The typical record length for these is about 10 years. Additional data came from the Department of Infrastructure, Planning and Natural Resources NSW (DIPNR) and the Department of Natural Resources and Mines Queensland (DNRM). Typical record lengths are again about 10 years. Considerable effort was invested into quality controlling data. As part of this process, 'discordancy' tests were applied to identify sites which are unusual with respect to the statistical properties of their annual maximum series. These sites were then inspected more closely.

For brevity, equipment-collecting data at sub-daily intervals will collectively be referred to as 'sub-daily gauges'. Continuous recording raingauges that are part of the standard network of Bureau of Meteorology stations will be referred to as 'pluviographs'.

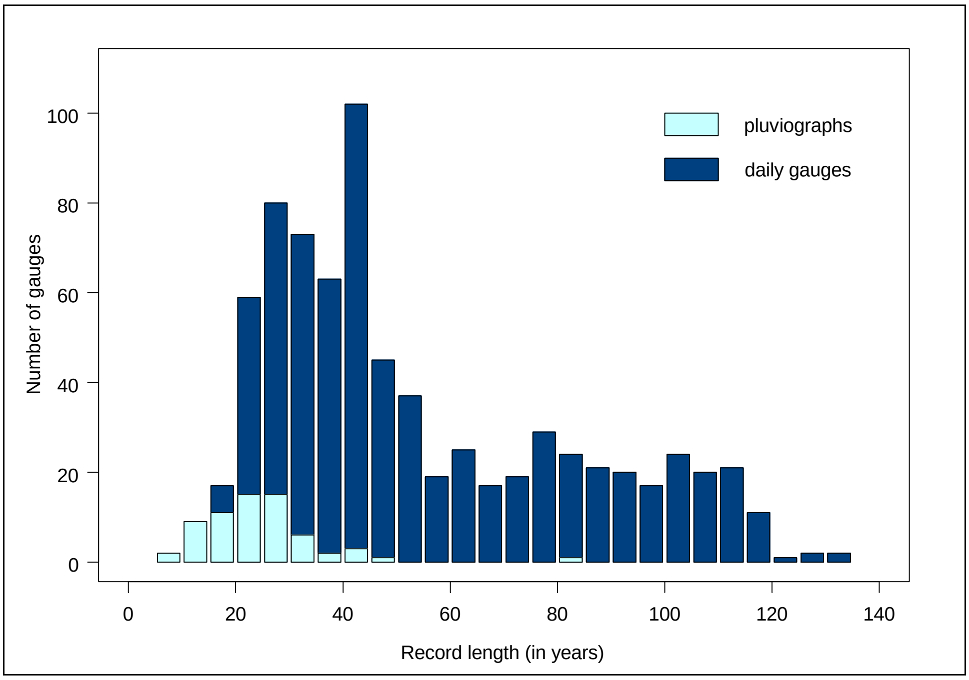

Figure 3 Record lengths for pluviographs and daily gauges used in the pilot study

4 Adjustment factors

4.1 Restricted to unrestricted durations

4.1.1 Introduction

Data at a daily resolution are abundant, giving an adequate coverage of the pilot area and with good record lengths. There are far fewer and shorter records available for durations below one day. Methods were explored to infer information at shorter durations from daily data. This required a conversion from 1-day rainfalls ('fixed' or 'restricted', usually from 9 am to 9 am) to 24 hour rainfalls ('sliding' or 'unrestricted', no pre-defined start time).

A conversion factor of 1.13 was used in ARR87 (restricted rainfall is multiplied by 1.13 to convert to unrestricted rainfall). This factor had originally been derived in the 1960's based on gauges in the US (Miller *et al*. 1973). Using an approach suggested by Dwyer and Reed (1995) this factor was reviewed using seven Australian capital cities. There is considerable variation between gauges but also from one period (just under two years) to the next. Dwyer and Reed (1995) suggested that climate might influence the value of the conversion factor. Based on the current data, no strong effects could be found. It is suggested that the previously-used factor of 1.13 might be somewhat low and a value of 1.15 would be more appropriate. However, this factor is an average and might be completely inappropriate for converting specific events.

4.1.2 Background

The term 'sliding' (or 'unrestricted') in the following refers to rainfall as measured by pluviographs whereas 'fixed' (or 'restricted') refers to rainfall starting at a certain time of the day, e.g. 9 am. ARR 87 states that: 'The restricted-to-unrestricted corrections used are the same as in Miller *et al.* (1973).'

Miller *et al*. (1973) use the terminology of 'fixed- versus true-interval precipitation'. The relations had been investigated in studies dating back to the fifties and sixties. It is suggested that 1.13 times the value for a particular return period based on observation day ('fixed' interval) is equivalent to the same return period value obtained from a series of 1440 min ('true' interval) values. It is pointed out that this factor is not based on a causal relationship, but is an average of index ratios. Weiss (1964) confirmed the empirical values using probability theory. Dwyer and Reed (1995) come back to Weiss' theory and show that his argument is flawed.

Pierrehumbert (1972) compares 24-hour unrestricted falls to restricted falls with the observation day starting either at 9am or at midnight at four Australian sites (Adelaide, Brisbane, Melbourne and Sydney). Correction factors vary from 1.10 (Adelaide) to 1.19 (Sydney and Melbourne).

Flood Estimation Handbook (FEH) Volume 2 (Institute of Hydrology, 1999) uses factors taken from Dwyer and Reed (1995) to convert 'fixed' to 'sliding' durations. The 1-day (fixed-duration) rainfall is multiplied by 1.16 to yield the sliding duration rainfall. See Table 1 for other durations/resolutions.

Rainfall measured daily		Rainfall measured hourly	
Duration (days)	**Multiply by**	**Duration (hours)**	**Multiply by**
1	1.16	**1**	1.16
2	1.11	**2**	1.08
4	1.05	**4**	1.03
8	1.01	**8**	1.01
		≥ 12	1.00

Table 1 Correction factors used in FEH Volume 2.

Dwyer and Reed (1995) found differences between correction factors for the five sites investigated (three Australian and two UK sites): Brisbane, Leeming (North Yorkshire), Melbourne, Ringway (near Manchester), Sydney and Eskdalemuir (Southern Scotland). However, only in a few cases were these differences statistically significant.

Eskdalemuir (prone to long-duration frontal events) requires high correction factors (1.167) while Brisbane (short-lived monsoonal storms) requires lower correction factors (1.147). Moderate correction factors apply for other sites (average 1.158 or 1.16 as used in FEH). Generally it is found that conversion factors are higher than those usually applied (in the range 1.15 to 1.17 rather than 1.13 to 1.14).

For recent work on this subject, see Young and McEnroe (2003) and Michel (2005).

4.1.3 Data

The aim of this study was to derive correction factors specifically tailored to conditions in Australia. In addition to the three Australian sites investigated by Dwyer and Reed, correction factors were derived for another four capital cities (Darwin, Perth, Adelaide and Hobart), ensuring coverage of different climate conditions. Refer to Table 2 for the start of selected periods. As will be explained in 4.1.4, the maximum length for a period is 16384 hours, about 683 days or a month and a half short of two years. For Sydney two periods (G and H) have been selected in addition to the six periods listed in Dwyer and Reed. The earlier of these two periods starts in October 1996, after the publication of the report. For neither period are hours missing or has rainfall been accumulated over a number of hours.

Gauge Name	Gauge #	Period	Start
Melbourne (Regional Office)	086071	A	21/12/1953
		B	24/11/1962
		C	07/05/1944
		D	09/09/1952
		E	10/07/1957
Sydney (Observatory Hill)	060662	A	20/06/1943
		B	02/09/1947
		C	30/05/1952
		D	03/05/1961
		E	11/06/1963
		F	20/04/1965
		G	29/10/1996
		H	30/09/1998
Brisbane (Regional Office)	040214	A	12/02/1916
		B	04/04/1934
		C	19/03/1940
		D	02/08/1947
		E	06/11/1953

Gauge Name	Gauge #	Period	Start
Darwin (Airport)	014015	A	13/02/1957
		B	13/01/1959
		C	16/12/1996
		D	17/09/1999
Perth (Regional Office)	009034	A	29/12/1950
		B	30/08/1969
		C	17/08/1977
Hobart (Airport)	094008	A	23/12/1978
		B	01/01/1997
		C	02/12/1998
		D	07/11/2000
Adelaide (Kent Town)	023090	A	19/02/1985
		B	10/08/1988
		C	01/02/1997
		D	01/02/1999

Table 2 Gauges and periods selected

4.1.4 Missing data

Hourly data for each of the sites were abstracted from the Bureau of Meteorology's rainfall archive. Unlike for the UK stations, missing data were of some concern. Completeness of records ranged between 91 and 96% (for Sydney and Adelaide respectively). Periods were chosen mainly to minimise the number of missing hours. Dwyer and Reed do not discuss how missing values were treated in their study. However, for our purposes it is assumed that the length and depth of an event are not correlated with the likelihood of the event being missing. Under this assumption, missing values were set to zero and are ignored in the following analyses. To check the assumption, the frequency of event lengths of accumulated rainfalls could be compared to the frequency of durations of rainfall for which data are available. Complete events could randomly be set to zero (i.e. ignored) and the correction factors rederived to assess the impact of missing data. Alternatively, or in addition, the accumulated rainfall depths could be used (together with information from neighbouring sites) to infill for the missing hours.

4.1.5 Selecting periods

Dwyer and Reed list the periods they used in the appendix to their report. For Melbourne, Sydney and Brisbane the cumulative sum of missing hours was plotted (against the dates) and these periods were indicated (as in Figure 4 for Brisbane). The periods 'fit' remarkably well and are also a means of confirming that this particular site was used in the study by Dwyer and Reed, since neither gauge number nor grid references are listed in their report. Although the results presented here do not in detail agree with those in the report by Dwyer and Reed, the overall picture is the same. It is therefore suggested that the basic ideas from the report were implemented correctly but subtle differences (for instance in the treatment of missing data) led to slight differences in the results.

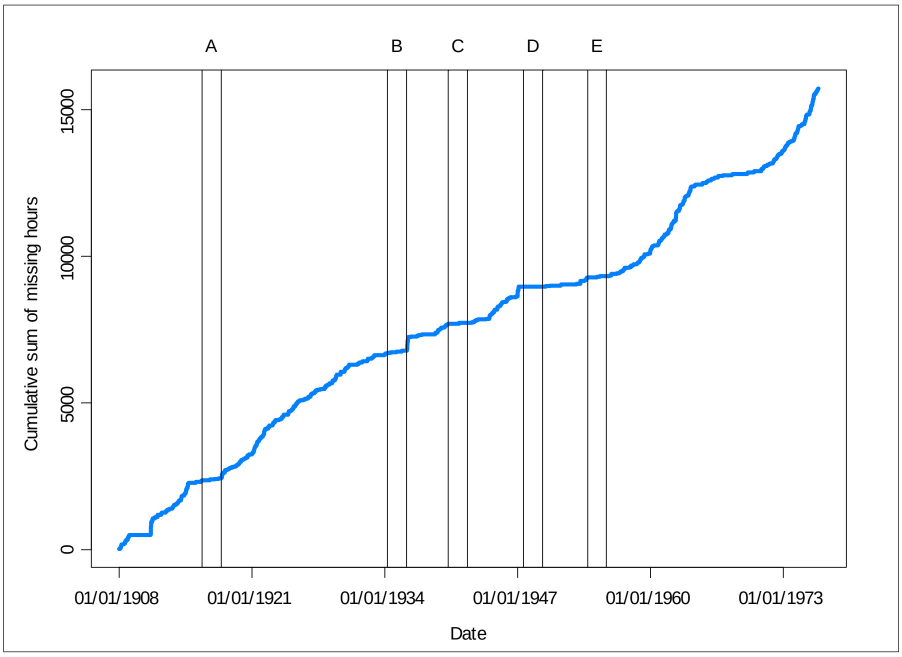

Figure 4 Cumulative sum of missing hours for Brisbane and the five periods used in Dwyer and Reed (1995).

4.1.6 Methods

Rules for abstracting fixed and sliding durations

The rules for abstracting the fixed and sliding durations are described in some detail in Dwyer and Reed (1995). The main idea behind their approach is to use relatively short periods of record (less than two years) to represent much longer records (of say 45 years). While the basic concept is easily understood, understanding the exact methodology may prove a little more challenging. An attempt is made to describe the approach nonetheless, refer to Dwyer and Reed (1995) for a more detailed explanation.

Periods represent records of 32 'years' and the length of a 'day' is equal to the duration, e.g. 10 hours. Although there is an upper limit to the length of a period (16384 hourly values) the actual length will vary with the duration. From each period 32 'annual maxima' will be abstracted. The length of a period is therefore specifically chosen to accommodate the maximum number n of complete intervals of duration m, where n is a multiple of 32. So for the 10-hour duration, the period length is 16320 (hours), yielding 32 'years' consisting of 51 'days' each and each 'day' has the duration of 10 hours. For the 64-hour duration the period length is 16384, but there are only 8 'days' in a 'year', making the 'annual maximum' less representative. The sliding durations however will not always fall neatly into one 'year' or another but may extend from one 'year' into the next. Dwyer and Reed introduce rules to avoid counting sliding maxima twice or losing them because they extend 'across the border'. Thereby, sliding intervals belong to the 'year' in which they mainly occur. If they are split into even lengths they are counted in the first of the two 'years'.

In rare cases the maximum rainfall depth for a period for the sliding duration may be lower than the rainfall depth for the fixed duration. Assume that for a duration of say seven hours the maximum sliding interval happens to straddle the 'border' with four hours in the first 'year' and three hours in the second. The maximum will be counted for the first 'year' and to avoid counting the same values twice the first three hours of the next 'year' will not be 'accessible' for the sliding intervals. For the fixed interval though the first three hours are of course available and if some extreme rainfall occurred in these hours, the maximum fixed duration rainfall for this period may exceed the maximum sliding duration rainfall depth and this effect

has been observed in the current study. It is difficult however to devise an alternative approach to remedy this problem without introducing new problems.

4.1.7 The model function

The following is a short summary of some of the theoretical background presented in Dwyer and Reed (1995). To convert individual maxima, the sample correction factor could be calculated as the mean of the individual ratios $R_i(D) = V_i(D)/F_i(D)$ where D is the duration, $V_i(D)$ are the *sliding maxima* and $F_i(D)$ are the *fixed maxima*. These estimates however can be seriously biased. We should therefore prefer $R(D)$, which is equivalent to the weighted sum of the individual ratios with greater weight being given to larger events. (m is the number of periods.)

$$R(D) = \frac{\sum_{i=1}^{m} V_i(D)}{\sum_{i=1}^{m} F_i(D)} \qquad \text{Equation 1}$$

One can calculate the *standard error s(R)* for *R(D)* and calculate confidence intervals using a t-distribution with (m-1) degrees of freedom:

$$s^2(R) = \frac{1}{m(m-1)\overline{F}^2} \sum (V_i - RF_i)^2 \qquad \text{Equation 2}$$

where $\overline{F} = \frac{1}{m}\sum_{i=1}^{m} F_i$

The *model function* $\rho(D)$

$$\rho(D) = 1 + a[1 - \exp(-b(D-1))] \qquad \text{Equation 3}$$

with $R(1) = 1$ and the *limiting value* $\rho^* = \lim \rho(D) = 1 + a$ (for D $\rightarrow \infty$)
can be fitted to the data by non-linear least squares regression. *a* and *b* are parameters found through non-linear regression.

4.1.8 Nonlinear least squares regression

A nonlinear model has a functional part that is not linear. The method of least squares is used to estimate the values of the unknown parameters. Processes in nature are sometimes best described using nonlinear models, for example in cases where a value approaches an asymptote. The biggest advantage of nonlinear least squares regression is that a broad range of functions can be fitted. A good estimate of the unknown parameters can be produced from a relatively small set of data. One of the major costs of moving from linear to nonlinear least squares regression is the need to use iterative optimisation procedures. The user needs to provide starting values. Bad starting values can cause convergence to a local minimum rather than to the global minimum. Like linear least squares regression the method is strongly sensitive to outliers and fewer model validation tools are available.
(Source: NIST/SEMATEC e-Handbook of Statistical Methods, www.itl.nist.gov/div898/handbook/pmd/section1/pmd142.htm, accessed April 2005)

4.1.9 Fitting the model

Dwyer and Reed suggest fitting a non-linear regression using least squares. In the current study this was done using the statistical software package S-PLUS (function 'nls'). Arguments that have to be supplied, are the form of the model to be fitted and starting values for the parameters *a* and *b*. In some cases S-PLUS aborted with an error message ('Singular gradient matrix'). Tracing the estimates for *a* and *b* through the iterations shows that *b* in these cases tends to take on unrealistically high values. This is a known Splus problem which in some cases can be avoided by supplying more appropriate starting values for a and b or by supplying the derivatives directly (Venables and Ripley, 2002). Both approaches have been tried but to no avail. There are hints that other software packages/languages might be more 'stable', e.g. Fortran, SAS. Dwyer and Reed mention 'records … which have an undefined value for b', possibly for the same reasons. For our purposes we excluded cases where the model could not be fitted from the following analyses.

Figure 5 shows an example for the model being fitted to hourly rainfalls at Sydney. The blue rather spiky line denotes the ratios *R(D)* for durations from 1 up to 32 hours. For a duration of 1 hour the fixed and sliding rainfall depths are the same and the ratio for this duration is therefore always one. The thick black line shows the fitted model and the red line indicates the confidence limit for *R*(24).

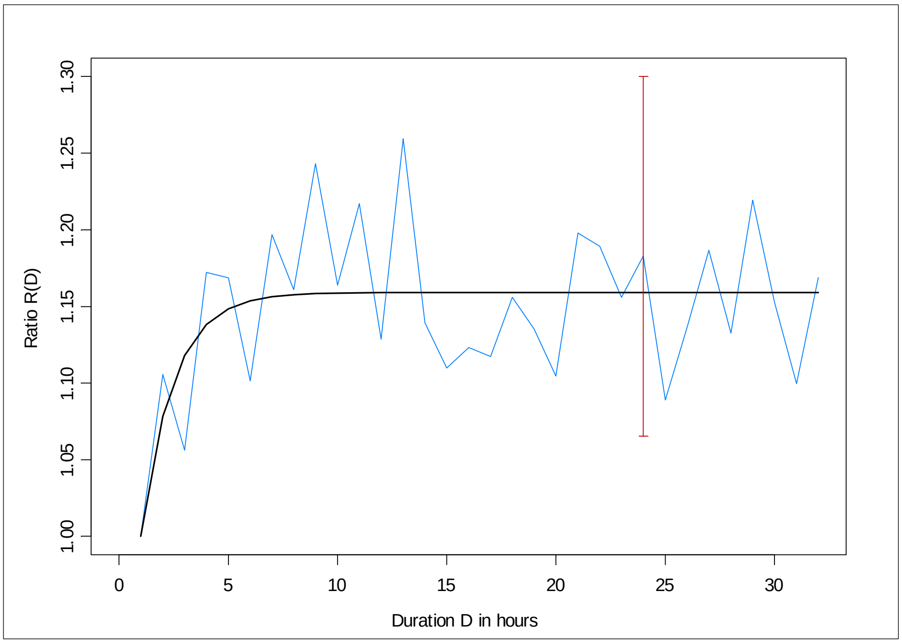

Figure 5 Non-linear least squares regression model fitted to hourly rainfall, Sydney period E and the 95% confidence limits for the 24-hour duration.

4.1.10 Results

First results and comparison to Dwyer and Reed

After fitting the model to derive estimates of parameters *a* and *b*, the correction factors were derived using equation 3. The averages agree well with those published by Dwyer and Reed. Considering only the three Australian sites used in their report, the highest correction factor was derived for Sydney (1.16), the lowest for Brisbane (1.15) and the correction factor for Melbourne lies just between these two values (1.155). According to this, correction factors between 1.15 and 1.16 would be sensible choices. As argued by Dwyer and Reed, the relatively low correction factor for Brisbane could be due to climatic conditions.

Four more sites

Based on the plots of cumulative sums of missing hours, periods were selected for the four additional sites. This task proved particularly difficult for Darwin. Four periods were selected but for only two of those could the model successfully be fitted. The average correction factors for converting daily (restricted) rainfall to 24-hour (unrestricted) for the seven sites are shown in Figure 6. The crosses denote estimates from single periods; the red dots indicate the average over the periods for one particular site and the dashed line indicates the average over all seven sites (1.151).

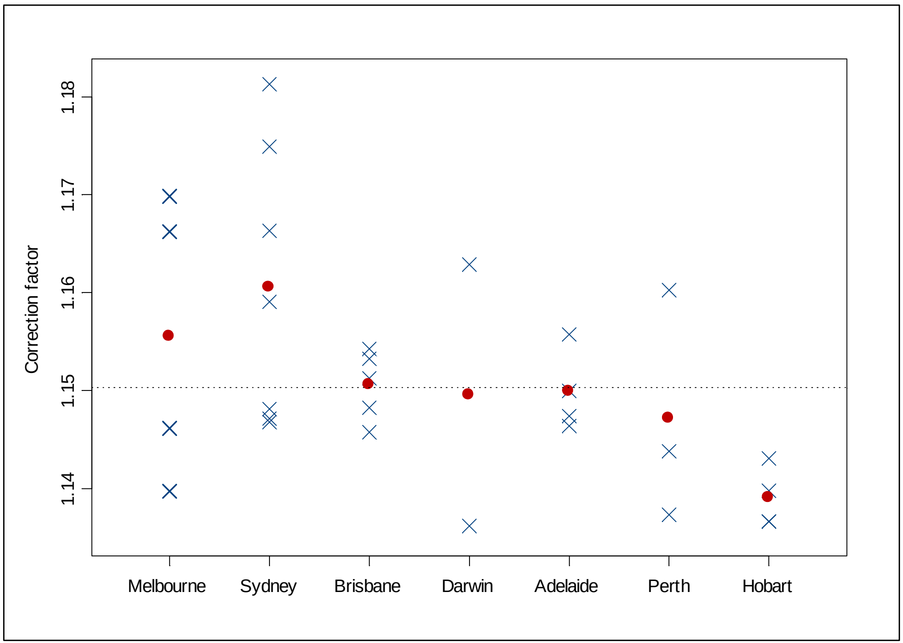

Figure 6 Correction factors for 24-hour duration rainfall for seven capital cities; dashed line indicates the overall mean

Variation over durations

Initially an 'average' ratio *R(D)* is derived as the ratio of the sums of the rainfall depths for sliding and fixed duration for a given duration (equation 1). This approach (rather than using the average of the ratios) gives more weight to large rainfall depths, which is sensible since we want to correct annual maxima. The derived ratios can be plotted against duration (blue line in Figure 5).

Figure 5 shows just how much variation there is from one duration to the next, necessitating some sort of 'smoothing' procedure. For the 24-hour duration, the 95% confidence limits are indicated (equation 2 and Haan (2002) for values of the *t*-distribution). While $R(24)$ is 1.183, the upper confidence limit is 1.300 and the lower limit is 1.065. For reliable estimates of the correction factor (even if only for one particular duration) it appears therefore appropriate to fit a model rather than just using the ratio for the 24-hour duration directly. Following the method developed by Dwyer and Reed, a non-linear least squares model can be fitted to estimate two parameters, *a* and *b*.

Variation between sites

The spread in estimates derived from a number of periods at one site is considerable compared to the spread between the averages at the seven sites. The differences between the average correction factors for Sydney and Hobart are significant at the 99% level. For Melbourne and Hobart the differences are significant at the 90% level. These values are based on variants of the two-sided t-test since the assumption of equal variance was tested

and accepted for Melbourne and Sydney, but rejected for Sydney and Hobart (not surprisingly, see Figure 6). The differences between Sydney and Brisbane are not significant on a 90% level. It is therefore suggested that it is not appropriate to derive correction factors depending on location.

Variation over periods

There is quite a wide spread between the correction factors derived for a given gauge. This is particularly apparent for Melbourne and Sydney. It was investigated what exactly causes the wide spread of estimates derived from the periods. It is suggested that this effect is caused by the relatively short time span covered by a 'period', just under two years. With the high variability of climate in Australia, two periods for one site may represent quite different 'climate states'. Frequency, duration and depths of extreme rainfall events may differ greatly from one period to the next. For Sydney in period F an event starting on the 16th October 1965 with a length of 57 hours occurred, totalling well over 200% of the average for the month of October. Period F also happened to be one of the periods with a rather high correction factor (1.181). One could assume that the wide spread of correction factors derived for Sydney is a sampling effect and that the relatively high correction factor might be due to some very long events.

4.1.11 Summary and conclusions

Previously attempts had been made to derive correction factors based purely on theoretical considerations. However, reasonable assumptions need to be made about certain rainfall characteristics, e.g. the rainfall profile. From these considerations it appears that correction factors should be lower than 1.20. The current study follows the approach taken by Dwyer and Reed.

There is a high degree of uncertainty in estimating the correction factor for a given location and duration. The derived correction factors should therefore preferably be used for 'average' conditions, e.g. for converting *series* of annual maxima. They may be inappropriate for single events.

There is some evidence that the previously used correction factor of 1.13 is too low. Correction factors for the sites in this study vary between 1.139 (Hobart) and 1.160 (Sydney). Although there is considerable variation between sites, it is suggested that some of it is due to the data available (missing data, period of record) and the method (treatment of sliding duration, number of 'days' for longer duration). It appears that the estimates of correction factors may be quite sensitive to these factors. Differences between average correction factors for the seven sites can only in some cases be considered to be statistically significant. The average over all site-specific correction factors is 1.151. The recommended adjustment factor for sites in Australia is therefore 1.15.

4.2 Annual maximum series and partial duration series

Ideally, estimates of rainfall depths (or intensities) for low ARI would be derived directly from partial duration series (Madsen *et al.*, 1997). If this is not feasible, a conversion factor has to be used instead.

4.2.1 Values used in ARR87

ARR87 refers to Miller et al. (1973) for conversion factors. Miller recommends using reciprocals of the empirically derived factors listed in Table 3. These factors were derived by Hershfield in 1961 (Hershfield, 1961). Prior to comparing between 2-year ARI maps of rainfall depths derived from annual maxima and the respective ARR maps, a correction factor of 1.13 should be applied.

ARI (in years)	2	5	10
Factor in Miller (1973)	0.88	0.96	0.99
Reciprocals	1.14	1.04	1.01
Values used in ARR87	1.13	1.04	1.00

Table 3 Previously used conversion factors

Langbein's equation

$$\frac{1}{T_{AMS}} = 1 - \exp\left(-\frac{1}{T_{PDS}}\right)$$

This approximate relationship (Langbein, 1949) allows conversion between ARIs on the two scales (T_{AMS} and T_{PDS} stand for the return period on the annual maximum scale and on the partial duration series scale respectively). A 2-year ARI on a PDS scale for instance corresponds to about a 2.54 year return period on an AMS scale. Similarly, probabilities could be converted. Since T = 1 / (1-F), Langbein's equation could also be written as:

$$F_{AMS} = \exp(F_{PDS} - 1)$$

F_{AMS} and F_{PDS} are non-exceedance probabilities on an annual maximum scale and on a partial duration series scale respectively. For $F_{PDS} = 0.5$ the corresponding F_{AMS} would be estimated as about 0.606. Assuming a Generalised Extreme Value distribution (GEV) is appropriate, this relationship could then be used to derive conversion factors to apply to quantile estimates rather than return periods or probabilities.

4.2.2 Deriving conversion factors

Data from the pilot area were used to derive appropriate factors to convert from AMS to PDS scale. For 413 (out of 547) regions (refer glossary), for the 24-hour duration in the pilot area, the GEV was judged to give an acceptable fit. For these regions, quantile estimates were derived for a range of ARI (2 to 1000 years) on a PDS scale on one hand and on an annual maximum scale on the other hand. The correction factor can then be defined as either the median or the mean of the ratios found (see Figure 7). Differences between mean and median of the ratios are minute and use of the average is suggested (as chosen in other studies).

For the revision of NOAA Atlas 14 (Bonnin *et al.* 2003) both AMS and PDS were abstracted and ratios of estimated rainfall depths on both scales were derived (see last row in Table 4). Between ARIs of 2 and 50 years these ratios are used as conversion factors. For higher ARI (up to and including ARI of 1000 years) a conversion factor of 1.004 is used.

ARI (in years)	2	5	10	25	50	100	200	500	1000
Median	1.111	1.0295	1.0125	1.0043	1.002	1.0009	1.0004	1.0002	1.0001
Mean	1.1106	1.0292	1.0124	1.0043	1.002	1.0009	1.0004	1.0002	1.0001
NOAA	1.113	1.029	1.013	1.006	1.004	1.005	1.006	1.008	1.009

Table 4 Ratios of quantile estimates for 413 regions, 24-hour duration (PDS/AMS)

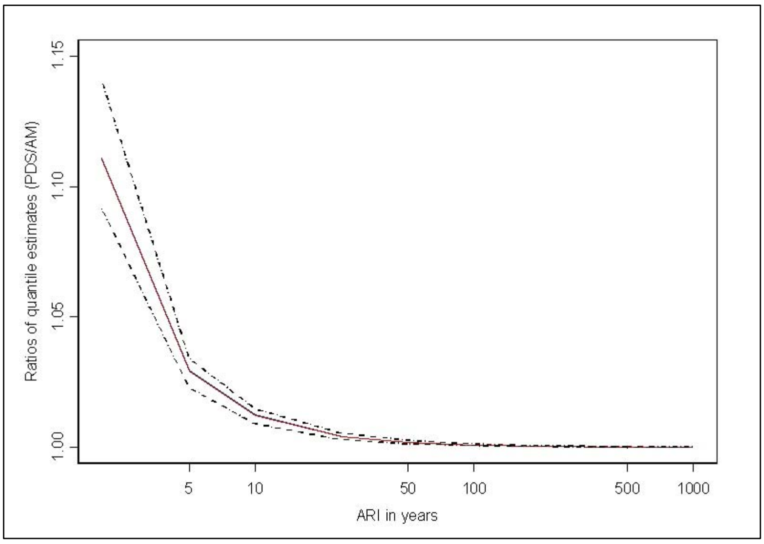

Figure 7 Ratios of quantile estimates (PDS/AMS) for 413 regions, 24-hour duration. (Dashed lines indicate the range of ratios found. Solid red and blue lines represent mean and median of ratios. Note: the lines are virtually indistinguishable)

4.2.3 Conclusion and Recommendations

The findings of this chapter are as follows:

- The conversion factors found in the revision of NOAA Atlas 14 are very similar to the ratios derived in this study. For NOAA Atlas 14, 59 homogeneous regions were considered. A Generalised Pareto distribution was used with the PDS. It appears that for the AMS at least two different distributions were used (depending on region) - a generalised extreme value (GEV) and a generalised log-normal (GNO, also known as 3-parameter log-normal, LN3) distribution. Durations between 5 minutes and 60 days were covered.
- The approach used here, makes two implicit assumptions: that Langbein's equation is exact (rather than just an approximation) and that the GEV is the true underlying distribution.
- It appears that the conversion factor used in ARR to convert rainfall depths (intensities) at ARI of 2 years is somewhat high.
- Given the way the conversion factors were derived and taking into account the errors introduced elsewhere in the estimation of rainfall depths (e.g. choice of distribution, regionalisation and sampling errors) it appears sensible to use only two decimal places for the conversion factors. The recommendation is therefore to use the set of conversion factors in Table 5. No conversion is required for ARI above 10 years.

ARI	**2 years**	**5 years**	**10 years**	**> 10 years**
	1.11	1.03	1.01	1.00

Table 5 Recommended conversion factors. Multiply rainfall depths derived on an AMS scale to convert to rainfall depths on a PDS scale.

5 Choice of moments

5.1 L-moments

The statistical properties of AMS (and PDS) can be described using L-moments. The ‘L’ in L-moments stands for ‘linear combination of order statistics’. These statistics have a number of useful properties and are now commonly used in rainfall and flood frequency estimation (Hosking and Wallis, 1997). Even for small sample sizes, L-moments (unlike conventional moments) should not become severely biased. L-moments can be used to identify the underlying distribution more efficiently than, for instance, Probability Weighted Moments.

For the purposes of this study L-CV and L-skewness will be used. L-CV stands for ‘coefficient of L variation’ and is a measure of spread. ‘L-skewness’ is calculated as the ratio of the third and second L-moment and is therefore independent of scale. Sample L-moments were derived using the ‘direct’ approach suggested by Wang (1996).

5.2 LH-moments

L-moments (and L-moment ratios) are relatively robust against outliers. For the purposes of estimating extreme rainfalls this might be considered a weakness rather than an advantage. Wang (1997) suggested a modified version called LH-moments. The ‘H’ in LH-moments stands for ‘higher’. In hydrology, one sometimes wants to put more emphasis on extreme events, rainfalls or floods. An LH-moment approach gives more weight to the largest events and accordingly lowers the influence of small values. (Similarly an LL-moment approach gives a high weight to the smallest values.) The following definitions of LH-moments are taken from Wang (1997):

$$\lambda_1^{\eta} = E\,[X_{(\eta+1):(\eta+1)}]$$

$$\lambda_2^{\eta} = \frac{1}{2}E[X_{(\eta+2):(\eta+2)} - X_{(\eta+1):(\eta+2)}]$$

$$\lambda_3^{\eta} = \frac{1}{3}E\,[X_{(\eta+3):(\eta+3)} - 2X_{(\eta+2):(\eta+3)} + X_{(\eta+1):(\eta+3)}]$$

$$\lambda_4^{\eta} = \frac{1}{4}E\,[X_{(\eta+4):(\eta+4)} - 3X_{(\eta+3):(\eta+4)} + 3X_{(\eta+2):(\eta+4)} - X_{(\eta+1):(\eta+4)}]$$

λ_1^{η} is the expectation of the largest variable in a sample of size η + 1 (location) where η is a shift (refer Wang, 1997),
λ_2^{η} is the expectation of the difference between the largest and the second-largest variables in a sample of size η + 2 (spread in upper part of distribution),
λ_3^{η} is a measure of symmetry based expectations for the three largest variable in a sample of size η + 3 (skewness) and,
λ_4^{η} is a measure of ‘peakedness’ derived from expectations of the largest four variables in a sample of size η + 4 (kurtosis).

For η = 0, LH-moments revert to L-moments. As for L-moments, LH-moment ratios are usually used in preference to LH-moments

With increasing shifts (η), LH-moments reflect more and more the larger events and the influence of smaller events is diminished. This is an approach that to the authors' knowledge is not currently widely studied or used. However, there are suggestions that a shift η of between two and four might be appropriate. To ensure consistency, one particular shift should be selected for an area/duration. An attempt was therefore made to find an 'optimal' shift η based on judging the goodness of fit of a GEV distribution (Wang 1998) to site data. These 'optimal' shifts varied considerably and in no systematic fashion.

It was found that use of LH-moments resulted in generally lower estimates of rainfall depths (and therefore intensities) than when L-moments are used (Figure 8). While the average difference at an ARI of 50 years is relatively small, in extreme cases the 50-year ARI estimate derived from a GEV based on LH-moments with a shift of 2 can be by more than 10% lower than an estimate derived based on L-moments. Site-based design estimates of ARI much higher than 50 years cannot be considered reliable. However, it is likely that a difference apparent at moderate ARI would be even stronger for high ARI. A decision was therefore made to use L-moments.

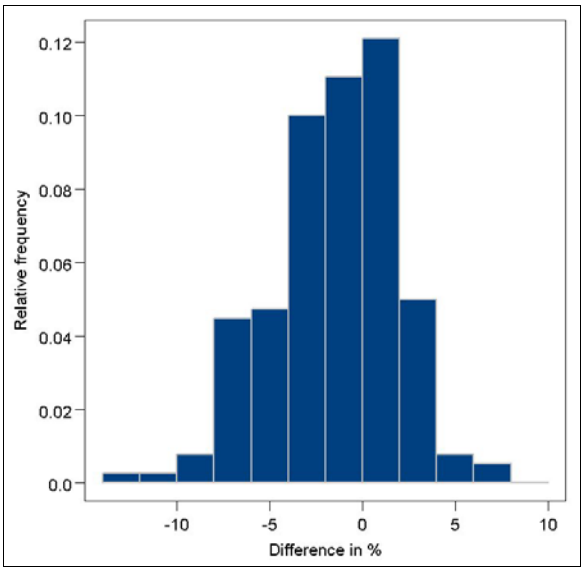

Figure 8 Relative frequencies of percentage differences in design rainfall estimates based on L-moments compared with LH-moments (shift η=2) for the 1-hour duration at an ARI of 50 years derived from the GEV fitted to site data.

6 Mapping index rainfall

6.1 Introduction

Based on annual maximum series (AMS), frequency curves can be constructed. In the pilot study a regionalisation approach is used that requires 'splitting' the frequency curve into two components: the index rainfall (median of the annual maxima) and the growth curve (frequency curve scaled by the index rainfall). At gauged locations the index rainfall can be derived from at-site data. A mapping approach was used to derive estimates of index rainfall elsewhere. (Refer to section 7 for background on frequency curves and section 9 for details of the regionalisation approach used in the pilot study.)

Regular grids were derived using the software package ANUSPLIN, which employs the technique of thin-plate spline smoothing (Hutchinson, 2004). This section describes which choices were made when mapping the index rainfall for three durations: 24, 48 and 72 hours. Figure 9 shows the surfaces fitted for the index rainfall (for each of the three durations).

6.2 Independent variables

The following characteristics were considered:
- latitude (in decimal degrees), LAT
- longitude (in decimal degrees), LON
- elevation (in m), ELEV
- mean annual rainfall (in mm), MAR
- distance from coast (in decimal degrees)
- characteristics derived from slope and aspect (dimensionless)

- Latitude and longitude come from a Digital Elevation Model (DEM, available online from Geoscience Australia at http://www.ga.gov.au/) with a resolution of 0.025 degrees.
- The elevation was derived based on a smoothed DEM. The idea behind this is that for daily rainfall (and above) very small-scale features will be of lesser importance compared to local/ regional scales.
- The mean annual rainfall comes from a grid produced in the Bureau of Meteorology's National Climate Centre (covers all of Australia at a resolution of 0.025 degrees).
- Two additional sets of characteristics (derived from slope and aspect) were prepared but were not selected in the final setup because they were not found to be useful.

The following characteristics were used: latitude, longitude, elevation and mean annual rainfall. Judging from the statistics supplied in the ANUSPLIN output files, models based on just LAT, LON and ELEV performed similarly well to those that were using MAR in addition to these three characteristics. However, in areas with particularly high index rainfall, residuals showed that the observed values were in places strongly underestimated if MAR was not used. MAR was included in the final model; this will lead to significantly higher estimates than ARR87 in some areas of the pilot area and somewhat lower estimates than ARR87 in other areas since the newly derived surfaces are far less smooth than those in ARR87 (see Figure 10).

6.3 Transformation (dependent and independent variables)

For many statistical procedures one has to make certain assumptions about the statistical properties of the data under investigation. There are three basic assumptions commonly made in regression analysis:
1. independence (an often used term is *iid* - independent identically distributed),

2. normality, and
3. constant variance (heteroscedascity, is often not satisfied as variance tends to increase with increasing values).

Validity of these assumptions can be assessed using tests (e.g. *Chi square test* to test against a Normal distribution) or plots (e.g. scatter plots to assess heteroscedascity). These assumptions are usually not strictly valid. The analysis can still be undertaken but there will be a degree of uncertainty associated with the conclusions drawn from these analyses.

One reason to apply transformations is to improve the normality of data. Common transformations are logarithmic transformation (for brevity often referred to as 'log transformation') and power transformations (e.g. square root transformation).

Transformations were made as follows:
- The dependent variables (rainfall depth for the three durations) were transformed using a square-root transformation.
- Some of the independent variables were transformed too. For the mean annual rainfall, a logarithmic transformation was applied.
- Elevation was scaled (divided) by a factor of 1000 as suggested by Hutchinson 1997. (Note the different units used on the horizontal surface and the vertical axis, i.e. degree for latitude and longitude, and meter for elevation.) It may be possible to achieve minor improvements by refining this factor (using generalised cross-validation as a guide).

6.4 Data problems

Locations with incorrect latitude and longitude or data problems are likely to appear in the list of highest ranking residuals (from fitting the surfaces). We found this to be true for two sites. For one site (Gatton QDPI Research Station, gauge number 40436) we corrected the location, the other site (Upper Crystal Creek, 58150) was rejected because of apparent data problems.

6.5 Investigating fitted surfaces

It is important to ensure consistency between estimates derived for the three durations. Thompson (2002) used a fixed signal (same for all durations). This option was used in the testing phase but was not required for the final model. Differences between two grids (say the 48 hour and 24 hour index rainfall, see Figure 11) can be derived and displayed.

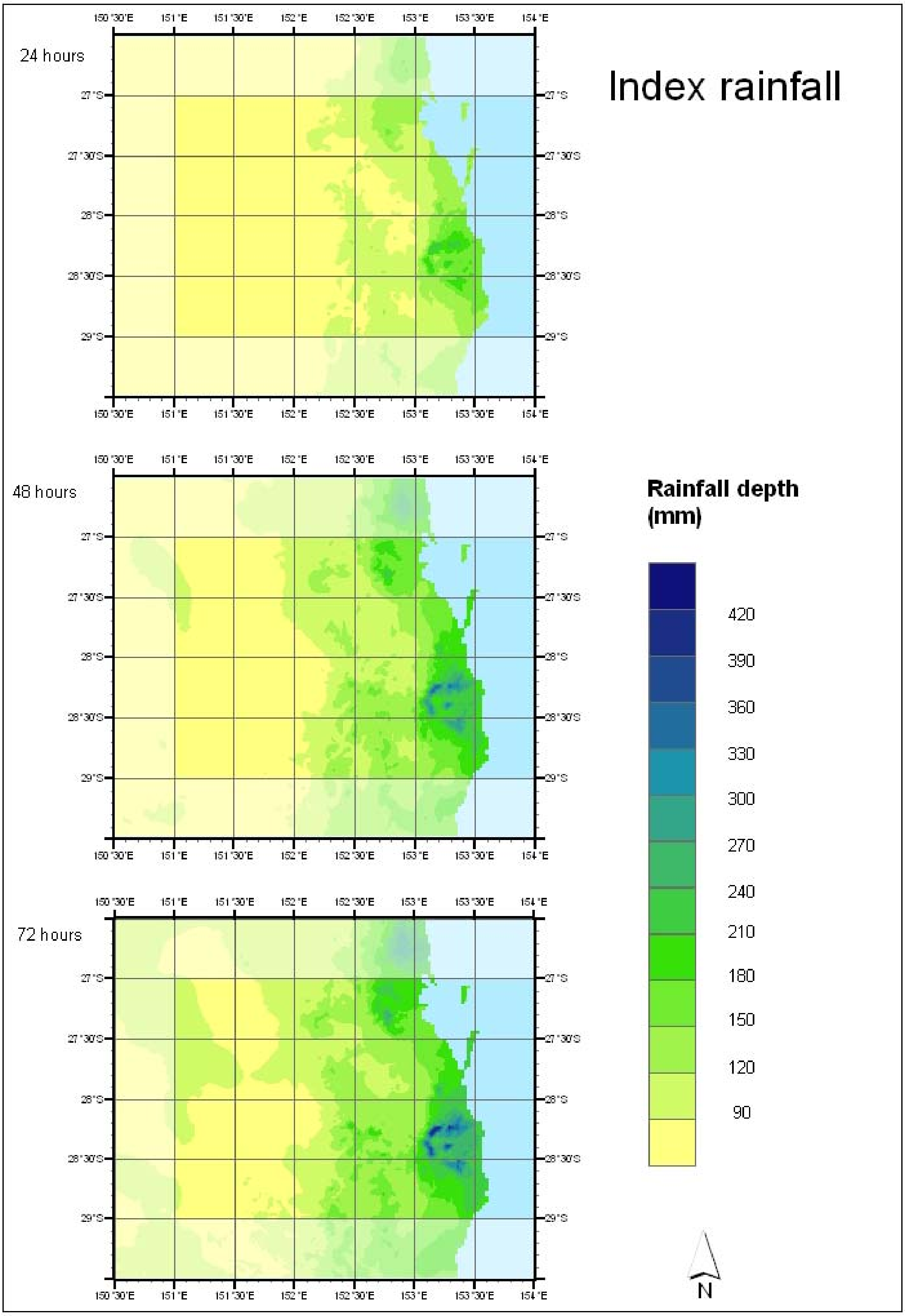

Figure 9 Index rainfall for 3 durations (top 24 hours, centre 48 hours and bottom 72 hours)

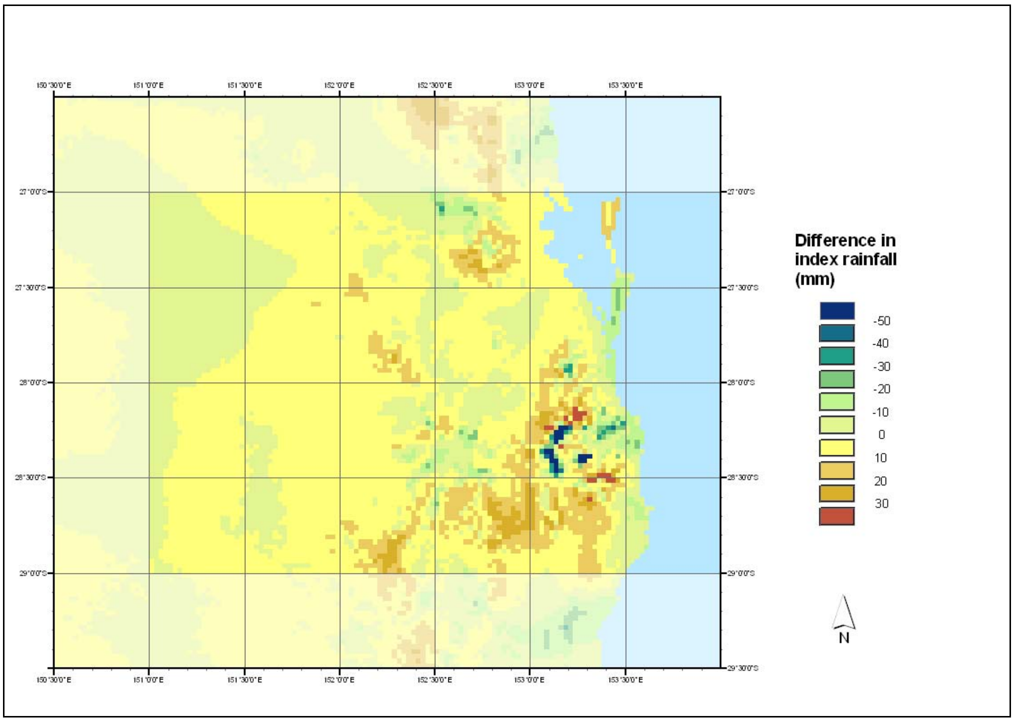

Figure 10 Differences in index rainfall estimates for the 24-hour duration. Green and blue shades denote locations where the new index rainfall estimates exceed the ARR87 estimates, whereas yellow and brown shades indicate locations where the new estimates are lower than the ARR87 estimates.

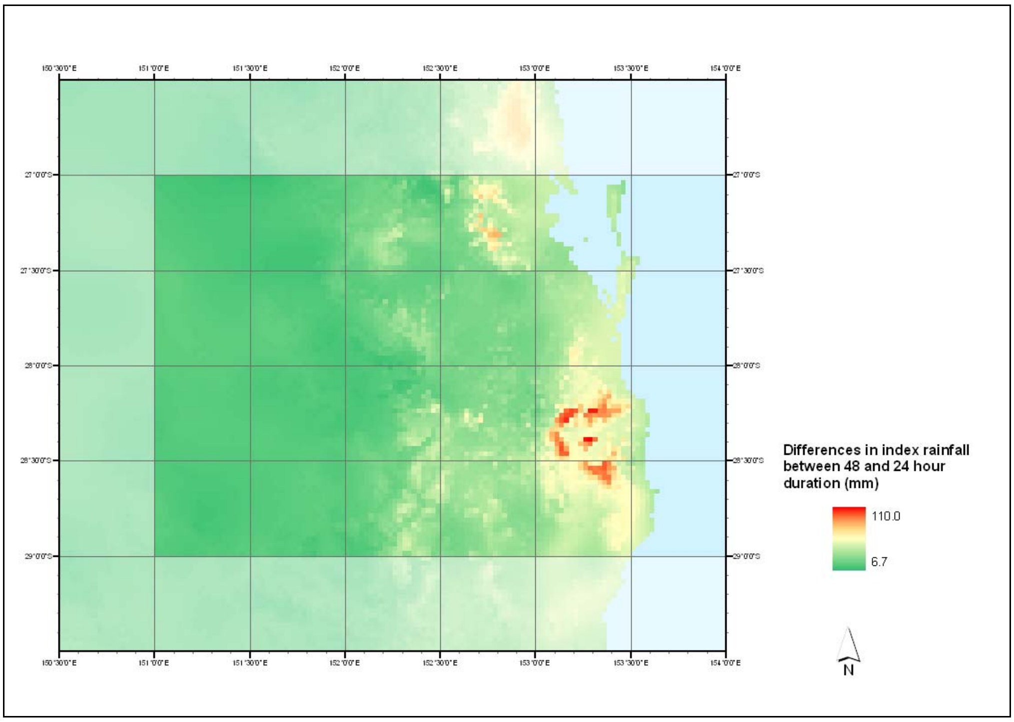

Figure 11 Example for a check of consistency between durations. Shown is the grid of differences of index rainfall between two durations: 48 hours and 24 hours. Large differences are shown in red, smaller differences in green. There are no negative values.

7 Choice of frequency distribution

In order to derive rainfall estimates, one typically fits frequency distributions to the annual maximum series (or partial duration series). Occasionally, non-parametric approaches are used (Faulkner 1999). In the past, two-parameter distributions would have been fitted, e.g. Gumbel or Logistic. ARR87 used the three-parameter log Pearson Type III.

Currently valid methods for design rainfall estimation apply mostly three-parameter distributions, in some cases even five parameter distributions, e.g. Wakeby. A three-parameter distribution can be defined using location ξ, scale α and shape k. *Frequency curves* can be scaled using either the mean or the median of the annual maxima. This scaling factor is called the *index rainfall*. The scaled version of a frequency curve is often referred to as a *growth curve*.

In the index rainfall approach (as used in the pilot study) index rainfalls are calculated from at-site data. Growth curves on the other hand (particularly if estimates for higher ARI are to be derived) should not be constructed from at-site data alone. Instead *regions* of gauges which are *similar* (with respect to characteristics relevant to design rainfall depths) - either to each other or a site of interest - are defined. For rainfall frequency estimation these regions are usually geographically contiguous. Two different approaches to defining such regions were explored in the pilot study: the *cluster* approach and the *circle* approach. The results presented in section 7.2 are based on an approach which in this report will be referred to as the *circles* approach (a region-of-influence approach) - all sites within a radius of *n* km from the site of interest are considered similar to the site of interest and are therefore included in the region. For more background on regionalisation refer to section 9.

Goodness-of-fit measures allow identification of the distribution that gives the best fit to a particular series and judge if this fit is acceptable. In this study, a goodness-of-fit measure based on L-moments, as suggested by Hosking and Wallis (1997), is used. Goodness-of-fit is assessed using the more general four-parameter Kappa distribution in combination with a resampling approach. The L-kurtosis of the fitted three-parameter candidate distribution is compared to the regional L-kurtosis, taking into account the bias and standard deviation for the regional L-kurtosis (derived using a resampling approach). Hosking and Wallis recommend a goodness-of-fit measure Z^{Dist} with a threshold $|Z^{Dist}| \leq 1.64$. Refer to Hosking and Wallis (1997) for an in-depth explanation. Lin *et al.* (2004) suggest two alternative goodness-of-fit measures.

7.1 Site growth curves

Five three-parameter distributions were tested: Generalised Logistic (GLO), Generalised Extreme Value (GEV), Generalised Normal (GNO), Pearson Type III (PT3) and Generalised Pareto (GPA). In the vast majority of cases the GEV gave the best fit and at the same time gives an acceptable fit to site data in at least 90% of cases.

7.2 Regional growth curves

Unlike for the site data, there is a significant number of regions for which a GEV will not give an acceptable fit. These results are supported by findings from other studies (Nandakumar *et al*, 2004). The exact numbers depend on the duration studied and the size of the region. For the 24-hour duration (based on daily data) and a radius of 20 km, the GEV is not acceptable in 25% of all cases. Using two additional distributions allows finding an acceptable distribution in about 98% of cases. Choosing a distribution solely on the grounds that it gives the best fit may well lead to inconsistencies between 'neighbouring' regions and durations.

7.2.1 Distributions for regional growth curves

First tentative regions have been defined using circles with radii of 60, 30 and 20 km. For these the following characteristics were studied: the number of gauges within a region, the number of station years, the heterogeneity and the goodness-of-fit for a set of five distributions.

For acceptably homogeneous regions, the goodness-of-fit measure can be used to judge the fit of a particular distribution. A distribution is then judged to give an acceptable fit if *Z* is less than 1.64 in absolute value. The distribution with the smallest absolute value of *Z* is said to give the best fit.

7.2.2 Best fit

Ignoring any possible effects of heterogeneity on the goodness-of-fit measure, the picture is the following: for the 60 km radius the GEV gives the best fit in the majority of cases, followed by the GNO and in relatively few cases the GLO will give the best fit. The pattern shifts as we move to smaller regions. The preference for the GEV becomes less pronounced.

The number of heterogeneous regions and the degree of heterogeneity increases with increasing region size. One would therefore expect any bias in the goodness-of-fit measure to be particularly strong for the 60 km radius. The analyses above were repeated for only those regions that were judged ‘possibly homogeneous’ (Figure 12).

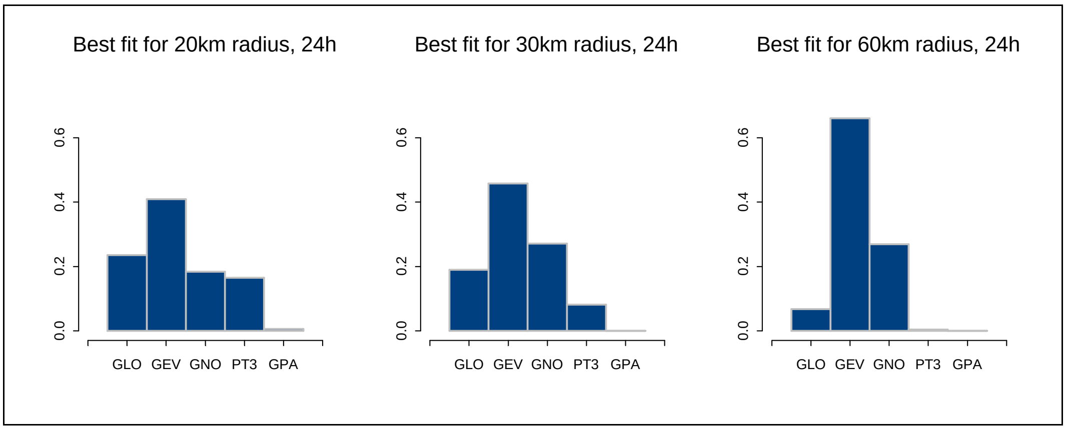

Figure 12 Distributions giving the best fit (relative frequency)

7.2.3 Acceptable fit

The GEV gives an acceptable fit in 75%, 70% and 57% of cases (for the 20, 30 and 60 km radius respectively). For the 20 km regions, an alternative distribution can be found that gives an acceptable fit for nearly 24 of the remaining 25% - so that for nearly 99% of all regions, distributions giving an acceptable fit can be found. For the largest radius (60 km) this percentage drops somewhat, but we can still find an acceptable distribution for nearly 87% of regions.

7.2.4 Choice of distribution for regions

Fortran code written by Hosking (2000) for use with the method of L-moments allows testing for five 3-parameter distributions: Generalised Logistic (GLO), Generalised Extreme Value

(GEV), Generalised Normal (GNO, sometimes referred to as 3 parameter Lognormal, LN3), Pearson Type 3 (PT3) and Generalised Pareto (GPA).

Goodness of fit was tested at two stages, initially for site data and later on for the regional distributions. For site data and across durations, the GEV gave an acceptable fit for well over 90% of cases. However, for regional data this picture is changed and in about 1 out of 4 cases the GEV did (according to the goodness-of-fit-measure) not give an acceptable fit.

Using just one distribution brings distinct advantages, compared to a set of say 3 or 4 different distributions. Firstly, this approach helps avoid overcomplicating the methods but (possibly more importantly) the use of different distributions for different regions may lead to inconsistencies as we ‘go’ from one region to another. It was therefore investigated how the choice of distribution affects the ARI of estimates. The investigations were undertaken for one duration (24 hours) and one region size (20 km). Two distributions were considered in addition to the GEV: GLO and GNO.

The GEV was found not to give an acceptable fit for 134 of the 547 regions (about 25% of cases). In 42 of these cases the GLO gave an acceptable fit, in another 58 cases the GNO gave an acceptable fit. Using these two distributions in addition to the GEV, it would be possible to find an acceptable distribution for nearly 95% of the regions.

Assuming the effect of choice of distribution would be most pronounced at the upper tail, growth factors at ARI of 100 years were derived. Figure 13 (left panel) shows the ratios of growth factors derived using a GEV and a GLO for the 42 cases in which the GEV did not give an acceptable fit but the GLO did. Figure 13 (right panel) shows a similar plot for the 58 cases in which the GEV was not acceptable but the GNO was.

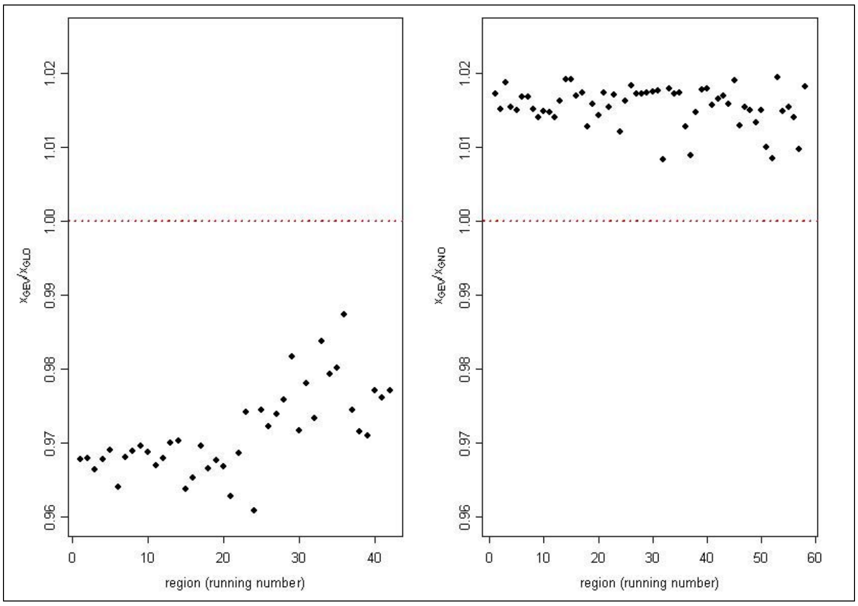

Figure 13 Ratios of growth factors at ARI 100 years, on the left for GEV v GLO and on the right for GEV v GNO

For the cases studied here, use of a GEV instead of a GLO leads to ARI 100 estimates which are between 1 and 4 % lower (than those derived using a GLO, which gives an acceptable fit in these cases.) Using a GEV where it is not acceptable but a GNO would be acceptable, leads to overestimation of the growth factor at ARI 100 years by between 1 and 2%.

Although an error of 4% is not negligible it is suggested for the reasons mentioned above (consistency, simplicity) that it may be defensible to use only one distribution. It is assumed that for the ARI under consideration (up to ARI 100 years) this error could be considered small when compared to other sources of uncertainty.

7.2.5 Fitting distributions for a median approach

The decision was made to use a median approach rather than the standard mean approach taken by Hosking and Wallis (1997). For site growth curves it is possible to use the 'standard' equations and divide by the (distribution) median. This approach cannot be taken for deriving the regional growth curves (ℓ_1, the first L-moment, is undefined). Below are the equations needed to derive growth curves for three distributions based on the median approach. These have been taken from Flood Estimation Handbook (FEH) Volume 3 (Institute of Hydrology 1999). For the GNO two typographical errors had to be corrected: a minus has been added in one equation (approximation for k) and a minus has been removed in another equation (estimating β, last term should be $\exp(k^2/2)$ not $\exp(-k^2/2)$).

In the following *x(F)* is the growth curve where *F* is the non-exceedance probability parameter and ξ is the location parameter. k is the shape parameter and c is used in calculating the shape parameter k for the GEV. t_2 denotes L-CV, t_3 denotes L-skewness and Γ the Gamma function.

Generalised extreme value distribution (GEV)

$$x(F) = 1 + \frac{\beta}{k}\left\{(\ln 2)^k - (-\ln F)^k\right\}$$

where $k \approx 7.8590c + 2.9554c^2$, $c = \frac{2}{3+t_3} - \frac{\ln 2}{\ln 3}$,

and $$\beta = \frac{kt_2}{t_2\left\{\Gamma(1+k) - (\ln 2)^k\right\} + \Gamma(1+k)(1-2^{-k})}$$

Generalised Logistic distribution (GLO)

$$x(F) = 1 + \frac{\beta}{k}\left\{1 - \left(\frac{1-F}{F}\right)^k\right\}$$

$k = -t_3$ and $\beta = \frac{\alpha}{\xi} = \frac{t_2 k \sin \pi k}{k\pi(k+t_2) - t_2 \sin \pi k}$,

where α is the scale parameter and ξ is the location parameter.

Generalised Normal distribution (GNO, often referred to as 3-parameter Log-Normal LN3)

$$x(F) = 1 + \frac{\beta}{k}\left[1 - \exp\left\{-k\Phi^{-1}(F)\right\}\right]$$

where

$$k \approx -t_3\left[\frac{E_0 + E_1 t_3^2 + E_2 t_3^4 + E_3 t_3^6}{1 + F_1 t_3^2 + F_2 t_3^4 + F_3 t_3^6}\right]$$

E_0 = 2.0466534
E_1 = -3.6544371
E_2 = 1.8396733
E_3 = -0.20360244

F_1 = -2.0182173
F_2 = 1.2420401
F_3 = -0.21741801

$$\beta = \frac{\alpha}{\xi} = \frac{t_2 k \exp(-k^2/2)}{1 - 2\Phi(-k\sqrt{2}) - t_2 \exp(-k^2/2)\left\{1 - \exp(k^2/2)\right\}}$$

Φ is the normal cumulative distribution function. Φ^{-1} is the inverse of the normal cumulative distribution function. A good approximation for Φ^{-1} can be found at http://home.online.no/~pjacklam/notes/invnorm/.

8 Inferring information about statistics of sub-daily data from those of daily data

There are far fewer pluviographs located in the pilot area than daily gauges (Figure 3). To improve on the coverage for sub-daily durations, a technique was needed that would allow transferring information from the daily to shorter durations.

A Partial Least Squares Regression (PLSR) approach was used to infer statistics for sub-daily durations from statistics at the daily duration based on 190 records with a minimum record length of 30 years.

A PLSR approach can be seen as two linked Principal Component Analyses (PCA), one on the set of independent variables, the other on the set of dependent variables. Generalised cross-validation was used to choose the appropriate number of factors (Geladi and Kowalski, 1986).

In the development of ARR87 methods, PCA (followed by regression) was used to derive equations for predicting the average recurrence intervals (ARI) at durations below 24 hours from the ARI for the 24, 48 and 72-hour durations. Maps were subsequently derived for three key durations and two key ARI.

The approach taken in the pilot study attempts to infer the index rainfall, the L-CV and the L-skewness at sub-daily time steps from these statistics at longer durations (together with additional independent variables). The main advantage of this approach is that it is far more general. Once the three key parameters are known, a frequency distribution can be fitted and rainfall estimates for any ARI can be derived without the need for interpolation. PLSR models have been developed to predict index rainfall, L-CV and L-skewness.

8.1 L-skewness

8.1.1 Estimating L-Skewness

L-moments for sub-daily durations had to be calculated from relatively short records (see Figure 3). The following investigations try to answer the question: How long a record should be used to derive representative L-skewness estimates?

90% of pluviographs (in the pilot area) have data for the period from 1991 to 2003. Data from 308 daily gauges with data covering this period as well as a 30-year standard period were used to derive estimates of L-skewness for the 24-hour duration based on 4 different periods (see Figure 14).

The blue line in each of these plots shows the *x=y* line, the red lines indicate the respective medians and the range in L-moment estimates based on the complete record is indicated by dotted lines. Figure 14a compares estimates based on the complete record for the daily gauges to estimates derived for the much shorter period from 1991 to 2003 (typical for pluviographs). The median of L-skewness estimates is very similar for the two periods (as indicated by the red lines). The range of estimates derived for the shorter period however is far wider than that for estimates from the complete records. Comparing the location of the cloud of points to the blue line shows clearly that it is not possible to infer the long-term L-skewness estimate from a 15-year record.

The experiment was repeated for a 30-year period (see Figure 14b). The standard period 1961-90 was chosen because it is generally accepted that the meteorological conditions (including extreme rainfalls) can be assumed as indicative of average conditions. Compared

to panel a), the points sit quite close to the *x=y* line and the range of L-skewness estimates from the 30 year period is only slightly wider than that for estimates from the complete records (typically 45 years).

The remaining two panels (Figure 14c and d) compare estimates from the complete records against estimates derived for the period 1961-75 (panel c) and 1976-1990 (panel d). Both panels show close resemblance to panel a. In none of these three plots is there an indication of a strong period-of-record effect.

Overall it is found that the length of currently available records is generally not sufficient for deriving reliable L-skewness estimates. Ideally such estimates would be based on records with a minimum length of 30 years.

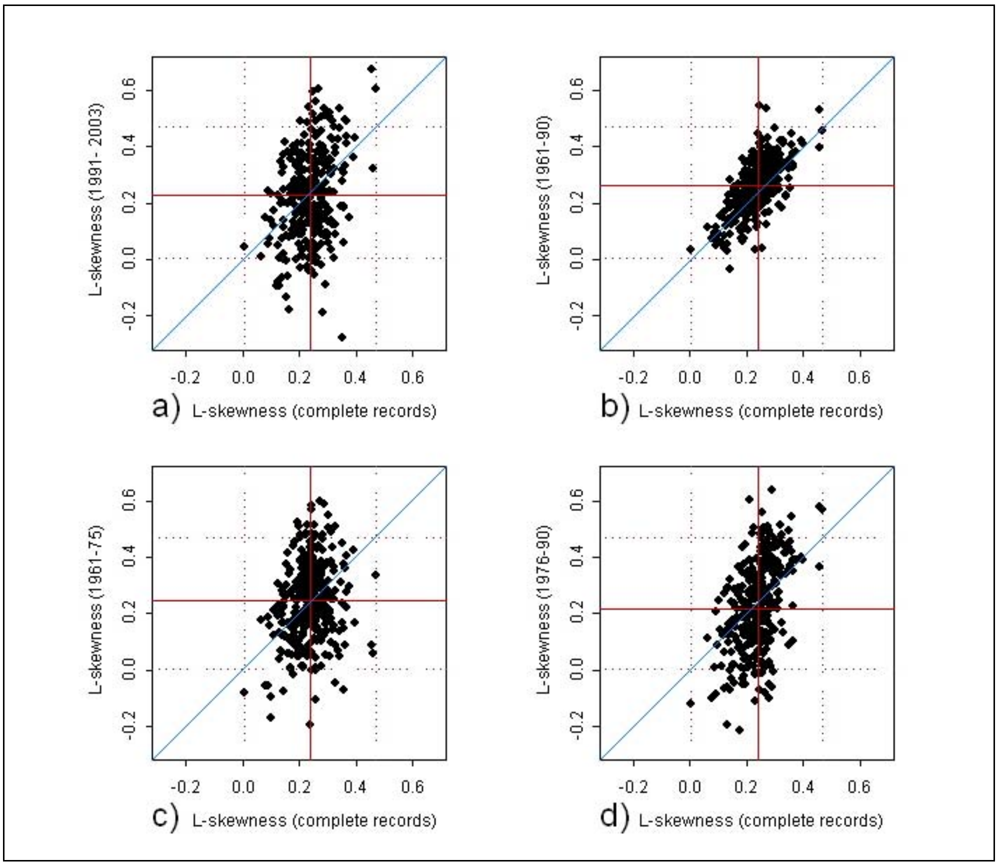

Figure 14 L-skewness estimates derived for the 24-hour duration based on a dataset of 308 daily gauges. Estimates based on the complete records are compared to estimates for 4 shorter periods. - Panel a) for the 15 years typically spanned by pluviographs in the pilot area, b) for the 30-year standard period (1961-90), for two 15-year periods within the standard period: c) from 1961-75 and d) from 1976-90.

8.1.2 Data and method

The set of dependent variables is: The L-skewness for durations 1, 2, 3, 6 and 12 hours. The set of independent variables is: L-skewness at 24 hours, latitude, longitude, elevation, mean annual rainfall and distance from coast. There is some degree of correlation within the sets of dependent and independent variables. An appropriate method for 'regressing' between these sets is Partial Least Squares Regression (PLSR).

Four slightly different datasets were used to derive PLSR models: based on minimum record lengths of 30, 40 and 45 years. Previous investigations (see Figure 14) had shown that records less than 30 years cannot be expected to give good approximations of the 'true' L-

skewness. A compromise should therefore be found to maximise the minimum record length used while on the other hand avoiding basing the model on a very limited number of gauges. Figure 15 shows the location of gauges with minimum record lengths of 45 and 50 years respectively. Each variable was standardised to mean 0 and standard deviation 1. A log-transformation for mean annual rainfall was attempted but rejected.

A problem frequently encountered in multivariate regression is the selection of an optimum number and set of independent variables. Different algorithms are used, for instance ‘stepwise forward’ and ‘stepwise backward’. It was found that for the problem at hand an exhaustive search was required.

The Predictive Residual Error Sum of Squares (PRESS) values give a good indication how well a particular model performs. The maximum number of factors is defined by the number of independent variables. The optimum number of factors was found as the one with the lowest sum of PRESS values for all five durations. This minimum sum of PRESS values was used to compare models with each other.

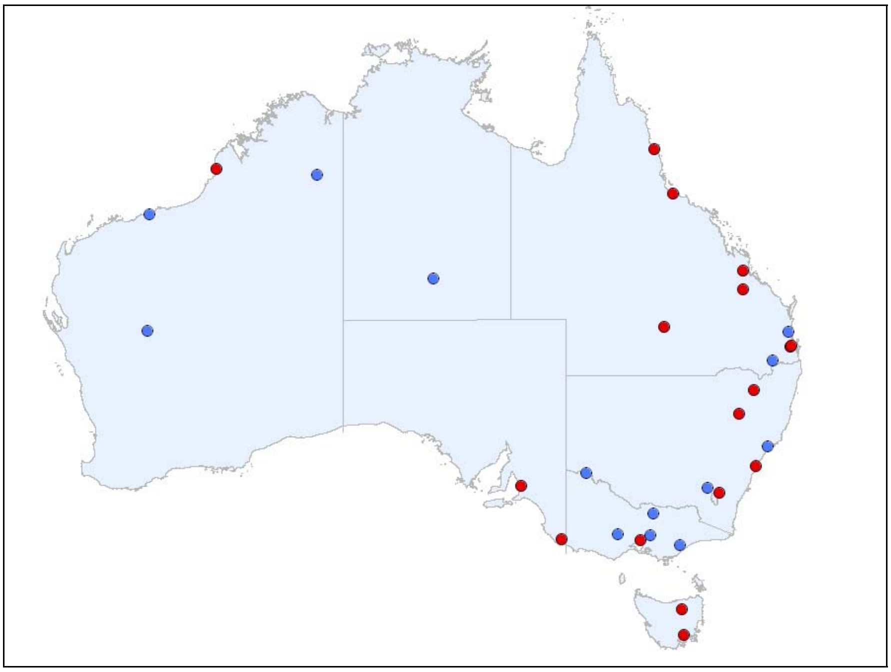

Figure 15 Location of pluviographs. Gauges with at least 50 years of data are shown in red, gauges with less than 50 years of data but at least 45 years are shown in blue

Instead of using a PLSR model one could assume that L-skewness is constant for a given duration. A PLSR model is only useful if it performs better than this assumption. The two approaches can be compared using the following measure: For each duration calculate the residuals between the ‘predicted’ L-skewness (either assuming constant or PLSR model). The smaller the sum the better the performance of the model. Figure 17 shows that the PLSR model always outperforms the average L-skewness.

Looking at the sum over PRESS values derived for the four datasets (Figure 16), there are some similarities. The optimum combination of 3 variables is either L-skewness (24 hours), latitude and distance from coast (30 and 45 years) or L-skewness (24 hours), latitude and mean annual rainfall. The optimum combination found from the ’50 year set’ is never in the top 4, but still among the better sets to use, while the alternative model (latitude, mean annual rainfall and distance from coast) performs rather poorly. Four-parameter models were tested for the ’45 year set’ but these did not perform better than the 3 parameter model.

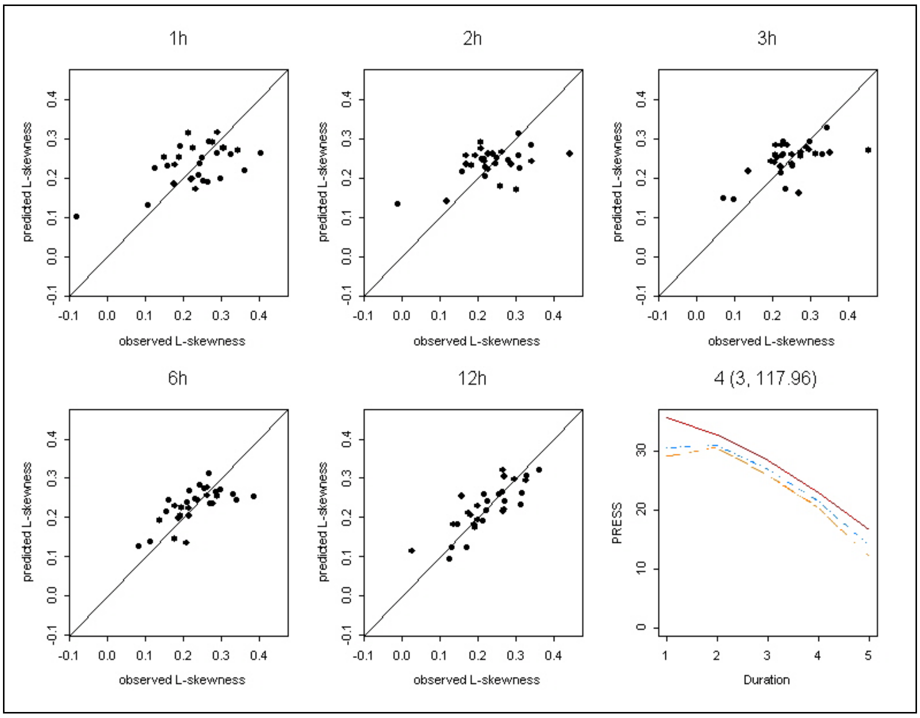

Figure 16 Scatterplots of predicted versus observed L-skewness for five durations and PRESS values (bottom right corner) for the 1 (red), 2 (blue) and 3 (orange) factor models based on '45 year set'. The optimum model uses L-skewness at 24 hours, latitude and longitude.

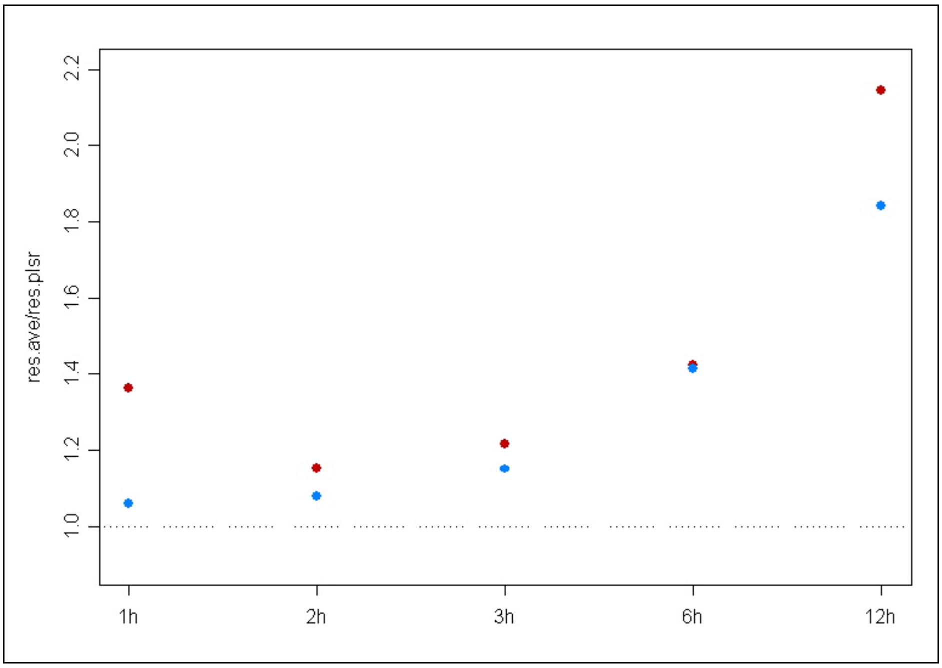

Figure 17 Comparing performance of PLSR model against using average L-skewness based on '50 year set' (17 gauges, red dots) and '45 year set' (60 gauges, blue dots), 3 variable 3 factor models.

For Figure 17, residuals were calculated for two models, once for a constant L-skewness and then again for the PLSR model. The y-axis shows the ratios between the sums of the (absolute values of) residuals. If the ratio is 1.0, then the PLSR model performs no better than

the assumption of constant L-skewness. The higher the ratio, the better does the PLSR model perform (compared to a model based on assuming spatially constant L-skewness).

Duration (in hours)	**1**	**2**	**3**	**6**	**12**
R^2	0.31	0.24	0.34	0.46	0.67
R	0.56	0.49	0.58	0.67	0.82

Table 6 R^2 and R for the '45 year set'

Table 6 gives R^2 and *R* values. R^2 gives an indication of the predictive power of a model. The closer this value is to one, the better does the model perform. R^2 for a perfect model would be 1.0. (All the points in Figure 16 would line up exactly on the $x=y$ line.) It is difficult to define what a 'good' R^2 value is but 0.24 (for the 2-hour duration) has to be considered low while 0.67 (for the 12-hour duration) is more satisfactory.

8.1.3 Recommendation

The recommendation is to adopt the 3 variable (3 factor) model based on L-skewness at 24 hours, latitude and distance from coast (based on the '45 year set'). This choice will also aid in avoiding large inconsistencies between L-skewness derived for 12 hours and 24 hour observed values.

This model performs better than assuming L-skewness is constant for a given duration. For the shortest durations (1 and 2 hours) the PLSR model performs only marginally better than if an average L-skewness value were used. It is suggested that this is not due to the choice of method or model but rather stems from the fact the L-skewness at short durations varies in a somewhat random fashion.

In the regionalisation approach, record length is used as weight to give more weight to estimates derived from longer records. Predicted values should be given less weight than estimates derived directly from observed data. The R^2 value gives a good indication of the certainty in the predicted value. It is therefore suggested that the record length at daily sites be multiplied by the R^2 value for this duration when combining estimates derived from pluviograph data with data predicted using the PLSR model.

8.2 Index rainfall and L-CV

A PLSR model was developed using a dataset of gauges with a minimum record length of 30 years (to ensure reliable estimates of the median). The number of gauges fulfilling these criteria (minimum record length and sub-daily durations) in the pilot area is insufficient, therefore gauges from all over Australia are used (118 gauges, see Figure 18).

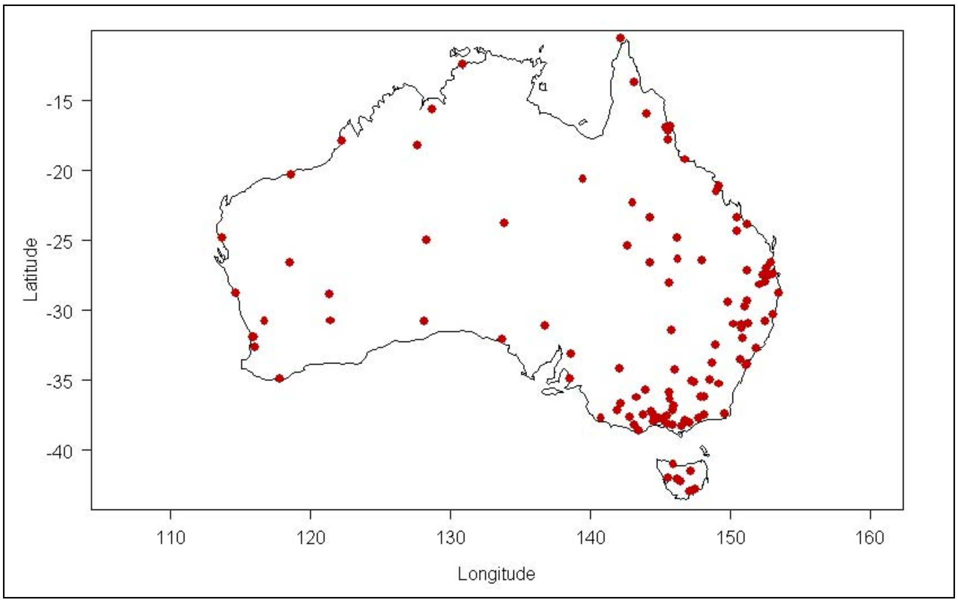

Figure 18 Location of pluviographs with a minimum record length of 30 years

The independent variables for 'predicting' the index rainfall for sub-daily durations are index rainfall at 24 hours, latitude, longitude and mean annual rainfall (for standard period 1961-90). The use of all three of the longer durations (48 and 72 hours in addition to 24 hours) results in only marginal improvements. The index rainfall (for both dependent and independent variables) is transformed using a square-root transformation.

The model performs very well judging from the R^2 values (see Table 7) and the residuals (comparing predicted values for the 118 gauges to the original values, Figure 19). For the index rainfall the PLSR model performs clearly better than an assumption of a spatially averaged index rainfall.

Duration	**1 hour**	**2 hours**	**3 hours**	**6 hours**	**12 hours**
R^2	0.94	0.96	0.96	0.97	0.98

Table 7 R^2 values for PLSR model

The PLSR model for predicting L-CV at sub-daily durations was developed in a similar fashion but based on records with a minimum record length of 45 years. The optimum model is a 3-factor model with the independent variables: L-CV (at 24 hours), latitude and distance from coast. The R^2 values are much higher than for predicting sub-daily L-skewness. Values range from 0.65 (for the 2-hour duration) to 0.94 (for 12 hours).

Figure 20 shows a comparison of the performance of the PLSR model against the assumption of spatially constant L-skewness.

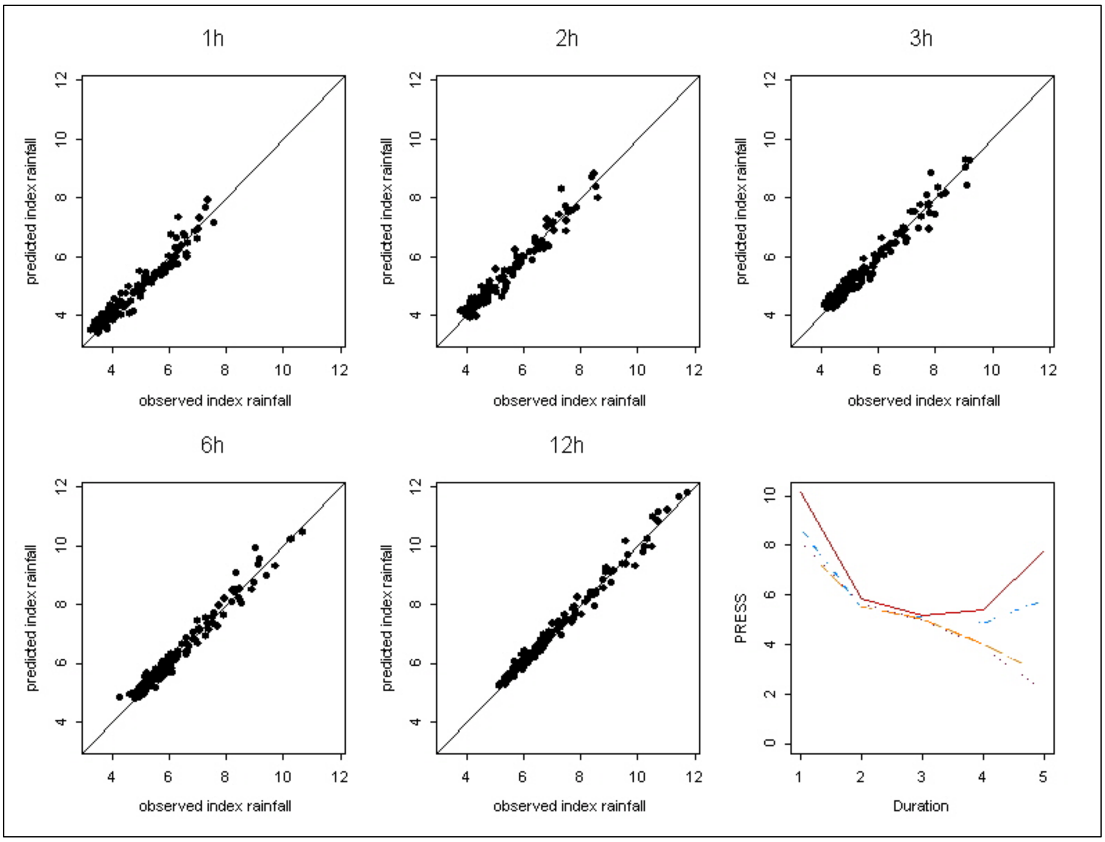

Figure 19 Observed and predicted index rainfall for 118 locations used in building the PLSR model for durations from 1 hour to 12 hours. The panel in the bottom-right corner shows the PRESS (Predicted Residual Error Sum of Squares) values for the five durations for 1 (red), 2 (blue), 3 (orange) and 4 (purple) factor models.

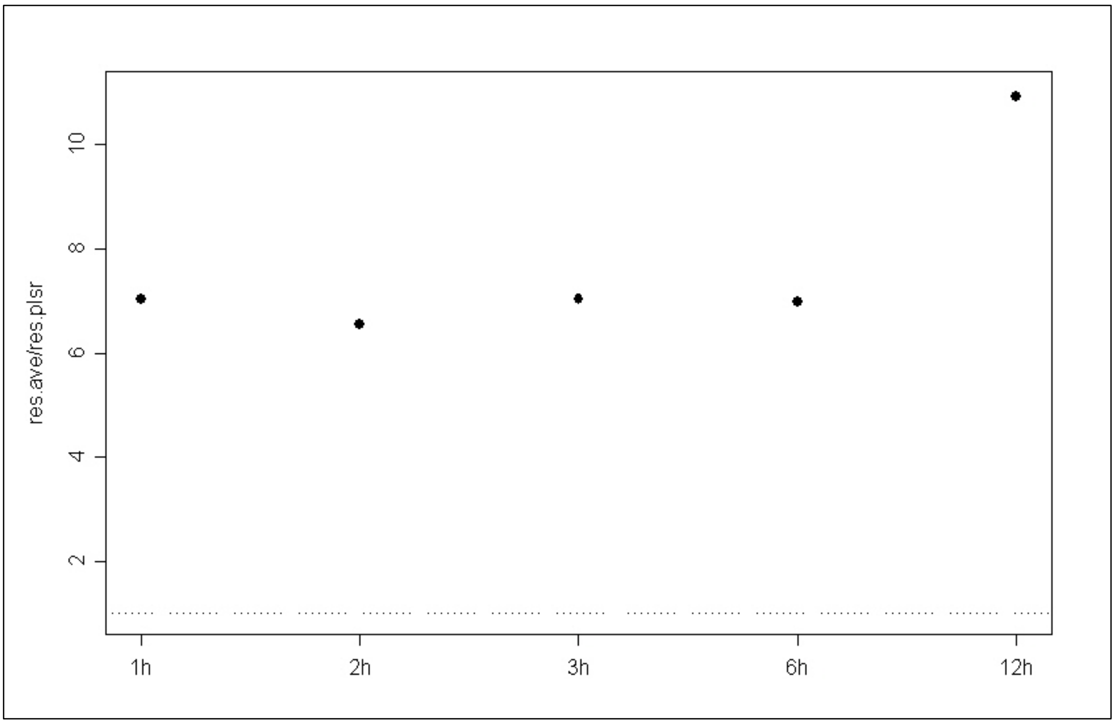

Figure 20 Comparing sum over residuals for assumption of constant index rainfall against sum over residuals from PLSR model. The dotted line indicates a ratio of 1.

9 Regionalisation

One of the fundamental problems in frequency estimation is the need to extrapolate to return periods significantly longer than the available records. This issue is usually solved using regionalisation, an approach pioneered by Dalrymple (1960) for flood frequency estimation. The basic idea of this approach is to define regions containing gauges that are similar - apart from a scaling factor - with respect to the characteristics of extreme floods. This scaling factor is these days often referred to as 'index flood'. In this report the expression 'index rainfall' will be used instead.

Dalrymple suggests the use of the mean annual flood as index flood and recommends using 'the graphical mean determined by the intersection of the visually best-fitting frequency line with the line corresponding to the 2.33-year recurrence interval' (Dalrymple, 1960, p. 37). For this study the median annual rainfall was chosen over the mean annual rainfall, mainly because the median is deemed more stable and robust against severe outliers. The median annual rainfall is derived from a fitted frequency curve (rather than using order statistics on the data directly). It is the rainfall intensity that would be equalled or exceeded in 50% of cases and would therefore have an ARI of 2 years. Other recently revised methods have preferred the median over the mean (Thompson, 2002, Faulkner, 1999).

Some of the major steps in a regionalisation approach are:

- Estimate the index rainfall for the location in question, i.e. median annual maximum.
- Define regions.
- Derive regional L-moments (as the weighted average of L-moments for sites in the region).
- Fit a regional growth curve.
- Multiply the regional growth factors by the best estimate for the index rainfall to derive the rainfall estimates.

At gauged locations the index rainfall can be derived directly. A mapping approach was used to derive estimates of index rainfall at ungauged locations. Maps were prepared using the software package ANUSPLIN, which employs the technique of thin-plate spline smoothing (Hutchinson, 2004). See section 6 'Mapping index rainfall'.

Regions defined for design rainfall estimation are usually geographically contiguous. The most appropriate size of a region can be judged using spatial extent, the number of sites in the region or the sum over the record length of all sites within the region ('station-years'). A statistical measure used to judge similarity between gauges (see 'Regionalisation' above) is the heterogeneity measure. This measure is another useful guide in defining the appropriate maximum size of regions. The minimum size will depend on the degree of extrapolation required, that is the ARI for which design rainfall estimates are required. Hosking and Wallis suggest that 20 gauges are sufficient for most applications.

9.1 Region-of-influence approach

One of the approaches to be investigated in this study is the 'region-of-influence' approach (Burn, 1990). Regions are defined for a 'site of interest' (which could be either gauged or ungauged), taking into account for which ARI (typical or maximum) estimates are to be derived. One of the major advantages of this approach is that there are no sharp boundaries (and therefore no inconsistencies) between regions. A relatively simple way of defining these

‘regions of interest’ was the ‘circles’ approach (see Chapter 7) successfully used by Thompson (2002).

Regions were defined for the 24-hour duration (from daily data) using a 20 km radius. Focal points of a region can only be located within the pilot area (excluding the buffer zone). Sites located within the buffer zone may be included in the region.

Hosking and Wallis (1997) define a heterogeneity measure H:

$H = \frac{(V - \mu_V)}{\sigma_V}$ where V is the weighted (according to record length) standard deviation of at-site sample L-CVs. A large number N_{SIM} (say 500) of realisations of a region with N sites is simulated. These simulated regions are homogeneous and sites have the same record length as the region for which the heterogeneity measure is to be derived. From simulations, the mean and standard deviation of the N_{SIM} values of V, μ_V and σ_V, are calculated. Hosking and Wallis (1997) suggest the following classification of regions according to heterogeneity, while pointing out that these values ‘are somewhat arbitrary’ and that they should be understood as ‘useful guidelines’.

$H < 1$	‘acceptably homogeneous’
$1 \leq H < 2$	‘possibly heterogeneous’
$H \geq 2$	‘definitely heterogeneous’

Hosking and Wallis (1997) define alternative heterogeneity measures based on L-CV and L-skewness (H_2), and L-skewness and L-kurtosis (H_3). The heterogeneity measure used in this study is based on L-CV only and the notation ‘H_1’ will be used.

Most of the regions in the pilot area have a heterogeneity measure H_1 of less than or equal to one. Apart from the size of the region and the selection of gauges to include, the estimate of regional growth factors depends on the weighting used in averaging the site L-moments to derive the regional L-moments. In addition to record length (which is typically used) other characteristics could be utilised - such as elevation, aspect and slope (Durrans and Kirby, 2004).

9.1.1 Hybrid approach

Durrans and Kirby (2004) suggested what they called a ‘hybrid approach’. In this approach the (final) regional estimate is derived as the weighted average of two initial estimates - one derived as the regional estimate using record length as the weight, the other using distance (or more precisely the inverse squared distance).

Durrans and Kirby (2004) found optimal weights to combine the initial estimates for a number of durations (from 1 hour to 2 days) by minimising the mean square error and the bias respectively. They found that ‘little is lost by preferring one to the other’ and that the hybrid estimator is superior to both the ‘index flood’ estimator and the ‘reciprocal distance squared’ estimator. Durrans and Kirby (2004) suggest that their approach could be extended to include other characteristics like elevation and ‘slope aspect’.

A set of characteristics was compiled for the sites in the pilot area, including latitude, longitude (in degrees), record length (in years), elevation (in metres), mean annual rainfall (in millimetres), eastern and northern components of the unit normal vector (dimensionless) as suggested by Hutchinson (1998).

Five characteristics were used in the following investigations (the northern component was excluded from further investigations). A Fortran routine was used to produce a file containing all 1001 possible combinations of 5 weights adding up to 1 (considering steps of 0.1). For each of these sets the ‘final’ regional estimates for L-CV and L-skewness were derived as the weighted average of the ‘initial’ estimates.

As a next step, RMSE and bias were calculated for each of these ‘final’ estimates (comparing with site estimates taken to represent the ‘true’ value). The best combinations (see equations 4 and 5) only very narrowly outperform the original index-flood approach.

$\text{L-CV}_{best} = 0.7\ \text{L-CV}_{reclen} + 0.1\ \text{L-CV}_{distance} + 0.2\ \text{L-CV}_{mar}$ Equation 4

$\text{L-skewness}_{best} = 0.7\ \text{L-skewness}_{reclen} + 0.1\ \text{L-skewness}_{elevation} + 0.2\ \text{L-skewness}_{mar}$ Equation 5

It is therefore recommended to use the original index rainfall approach. The marginal improvement in the model does not justify using a far more complicated approach.

9.1.2 Cross-correlation

There is a requirement to specify the uncertainty associated with the final estimates. When constructing confidence intervals, it is usually assumed that data are independent (in time and space). While the assumption of temporal independence is defensible, annual maxima of gauges located within 20 km of each other cannot be assumed independent. For a region with *m* gauges, the number of station years can be calculated as $\sum_{i=n}^{m} n_i$, where n_i is the record length for gauge *i*. The ‘effective’ number of station years however will be lower than the actual number of station years. Therefore spatial dependence will have to be taken into account when constructing confidence intervals. This will be an issue particularly for longer durations (larger-scale events).

The average correlation for annual maximum 1, 2 and 3-day rainfalls between each station and all other stations within a specified distance of it was calculated. The mean correlations were derived as weighted averages, using the number of overlapping years for the two stations as weights. In order to be included, stations were required to have an overlap of at least 5 years. Obviously station-years could be included only if data were present for that year at both stations.

At 24 hours there is a clear distinction between the higher correlations to the east of the ranges and the lower correlations about, and west of, the ranges (see Figure 21). This distinction decreases markedly from 24 to 48 hours - as correlations increase in the west - and less markedly - but still noticeably - from 48 to 72 hours. However, the pattern persists even at 72 hours.

Some stations within the same region were negatively correlated, usually where the overlapping period was relatively short. Some stations, not in the same region but within the pilot area, and with significant years of overlap, had high negative correlation (about -0.8). In the stand-out case that was followed up, there was a significant spatial separation (about 180 km) and the stations were on opposite sides of the divide.

These results could be used to assess the impact of intersite dependence on regional growth factors (Lin *et al*. 2004).

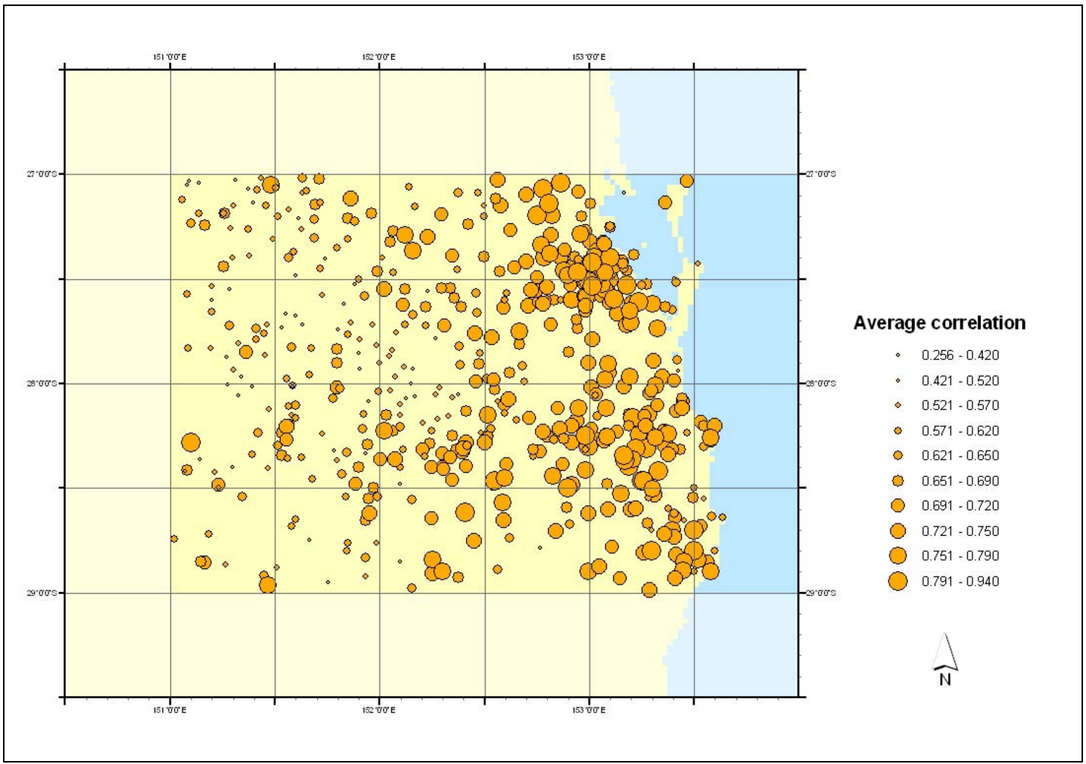

Figure 21 Average correlation between annual maxima of daily rainfalls for sites within regions (20 km 'circles')

9.2 Clustering approach

For this approach, regions were defined using a clustering approach. For shortness this approach will in the following also be referred to as the ‘cluster’ approach. Two different techniques were employed: the hierarchical Ward’s method and a k-means approach (Hartigan and Wong, 1979).

The Ward method is a hierarchical agglomerative approach to clustering. In each step, objects (or clusters) with a minimum distance (from each other) are grouped together. This process is repeated until all objects are allocated to just one class (or cluster). A whole range of agglomerative clustering techniques has been developed (e.g. Single Linkage, Complete Linkage). These methods differ only in the way the distance between objects/clusters is calculated (e.g. minimum distance between members of two clusters, maximum distance or average distance). In the Ward approach clusters are joined so that the increase in variance is minimised. Dendrograms are used to show in which order objects and clusters are combined. Dendrograms can be useful tools in deciding on an optimum number of clusters to use. One major drawback of hierarchical approaches is the fact that once an object has been assigned to a cluster it will stay in this cluster and cannot be reassigned.

For the k-means approach on the other hand, the final number of clusters needs to be specified at the outset and seeds for these clusters need to be supplied (for instance these could be randomly chosen observations). In an iterative process, each object is assigned to the nearest seed (usually using Euclidean distance). In the next step, cluster centroids are calculated and objects reassigned if they are found to be closer to another centroid. This process usually takes only very few iterations (less than ten), alternatively an epsilon value can be specified as cut-off value. Ideally these two approaches should be used in tandem, with the hierarchical approach (Ward’s method) to decide on the number of clusters and k-

means to do the actual classification. To guarantee that clusters are geographically contiguous latitude and longitude were given relatively high weights.

The characteristics used in the pilot study are latitude (LAT), longitude (LON), mean annual rainfall (MAR), elevation (ELEV), distance from coast as well as the eastern and northern components of the unit normal vector. The first four of these characteristics (LAT, LON, MAR and ELEV) were given a weight of 3, while a weight of 1 was assigned to the remaining three characteristics.

Clusters were explored for a range of cluster numbers. The optimum number of clusters for use in the classification was found by inspecting the homogeneity values for the resulting sets of clusters. The final set contains 16 clusters (see Figure 22). The number of gauges within a cluster ranges from 12 to 101, on average there are about 50 gauges in a cluster. Smaller clusters will be found close to the coast, in particular towards the north-east corner of the pilot area. Larger clusters tend to be located west of the mountain ranges. The initial clusters were 'manually' adjusted (with the aim of improving homogeneity within clusters), introducing some degree of subjectiveness.

For comparison: for circles the number of gauges varies between 2 and 66 gauges, on average there are about 20 gauges in a region. Regions with relatively few gauges are located in the most southerly parts of the pilot area (towards the buffer zone) and in the northeast (islands).

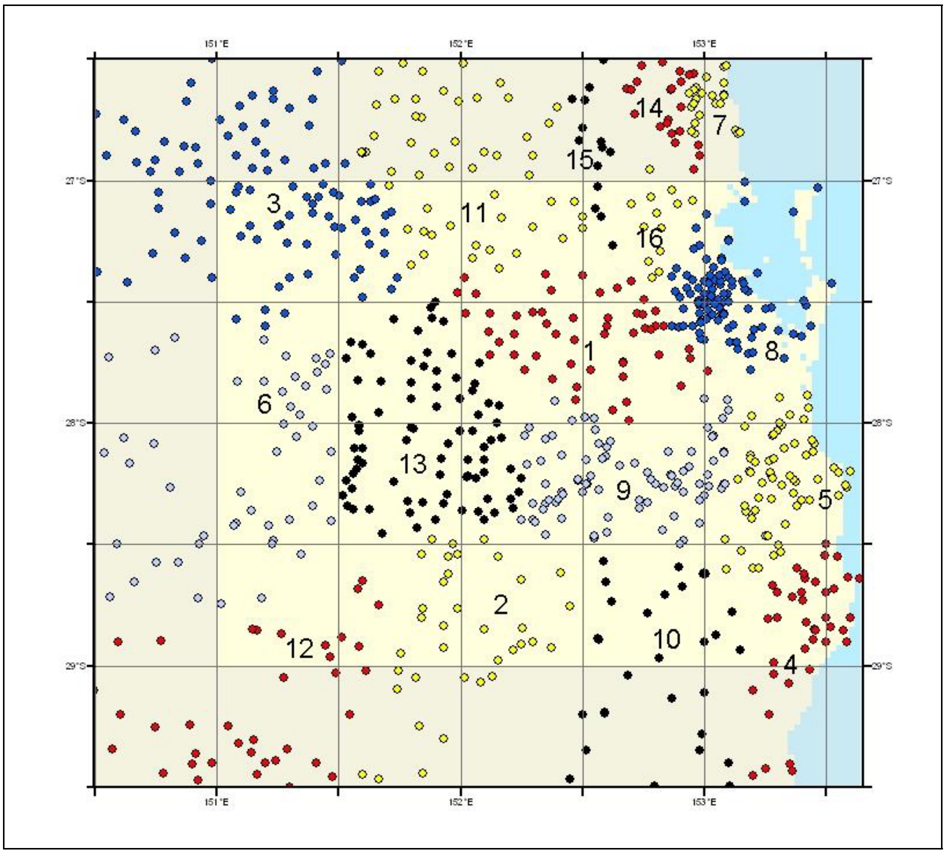

Figure 22 Clusters defined for the 24-hour duration (Neighbouring gauges belong to the same cluster if the symbols denoting their location are plotted in the same colour.)

9.3 Performance of regionalisation approaches

Two regionalisation approaches are under consideration: clusters and circles. Direct comparison will show where and by how much the derived growth factor estimates vary. To

judge however which of the two approaches performs better, regional estimates were compared to site estimates. Comparisons were undertaken for ARI 20 and 100 years. The results were in both cases very similar. Only the results for an ARI of 20 years will be presented however, since site estimates derived for ARI of 100 years may not be reliable.

9.3.1 Direct comparisons

The growth factors derived using the cluster and circle approaches were plotted by cluster (excluding clusters 7 and 14 which are located completely within the buffer zone, see Figure 22 for the location of these clusters and Figure 23 for the growth factors). The regional growth factor derived for a cluster is a single value denoted by a red line in Figure 23, regional growth factors derived using the circles approach are shown as blue dots. While there are some clusters where the estimates derived using the circles approach scatter relatively evenly around the line representing the cluster estimate (see for instance clusters 9 and 13) there are other cases where the circle estimates are almost exclusively lower than the cluster-based estimates (e.g. cluster 16) or higher than the cluster-based estimate (e.g. cluster 12).

The non-parametric Wilcoxon test was used to test if the 'means' over growth factors are significantly different. (A t-test may not be appropriate, since a standard deviation cannot be calculated for a single value.) The differences between the average regional 20-year ARI estimates derived via the two different approaches are statistically significant, apart from cluster 15. (Here only the four sites within the pilot area were used to derive the regional estimate, see Figure 22).

Differences in 20-year ARI growth factor estimates between the two approaches are relatively small (between -0.59 and 0.66), in particular when compared to differences in estimates between regional and site estimates (see below).

9.3.2 Overall performance

Two measures characterising the overall performance of the two approaches were used, mean absolute error *(MAE)* and *bias*.

$$MAE = \frac{1}{n}\sum_i \left| x_i^{reg} - x_i^{site} \right| \quad and \quad bias = \frac{1}{n}\sum_i (x_i^{reg} - x_i^{site})$$

where *n* is the number of sites. x_i^{reg} is the regional growth factor derived for site *i* and x_i^{site} is the growth factor derived from at-site information.

The *bias* is in both cases negligible, considering typical 20-year ARI growth factors take a value of 2. The *MAE* for the circles approach is lower than that for the clusters approach (as was found based on 100-year ARI estimates).

$MAE_{circles}$ = 0.136 $MAE_{clusters}$ = 0.148
$bias_{circles}$ = 0.0037 $bias_{clusters}$ = -0.0029

Figure 24 shows histograms of differences between regional and at-site estimates for the two approaches. The figure shows that the performance of the two approaches is very similar. If anything, clusters show a slightly higher tendency of severely underestimating site estimates.

9.3.3 More detailed performance assessment

Figure 25 shows how the difference between regional and site estimates varies with geographical location (left for circles approach, right for cluster approach). There is a high degree of similarity between the resulting plots. Overall it appears the circles approach performs better. This rather subjective visual assessment is confirmed by the *MAE* and bias

statistics above. One cannot identify a region where the cluster approach clearly outperforms the circle approach.

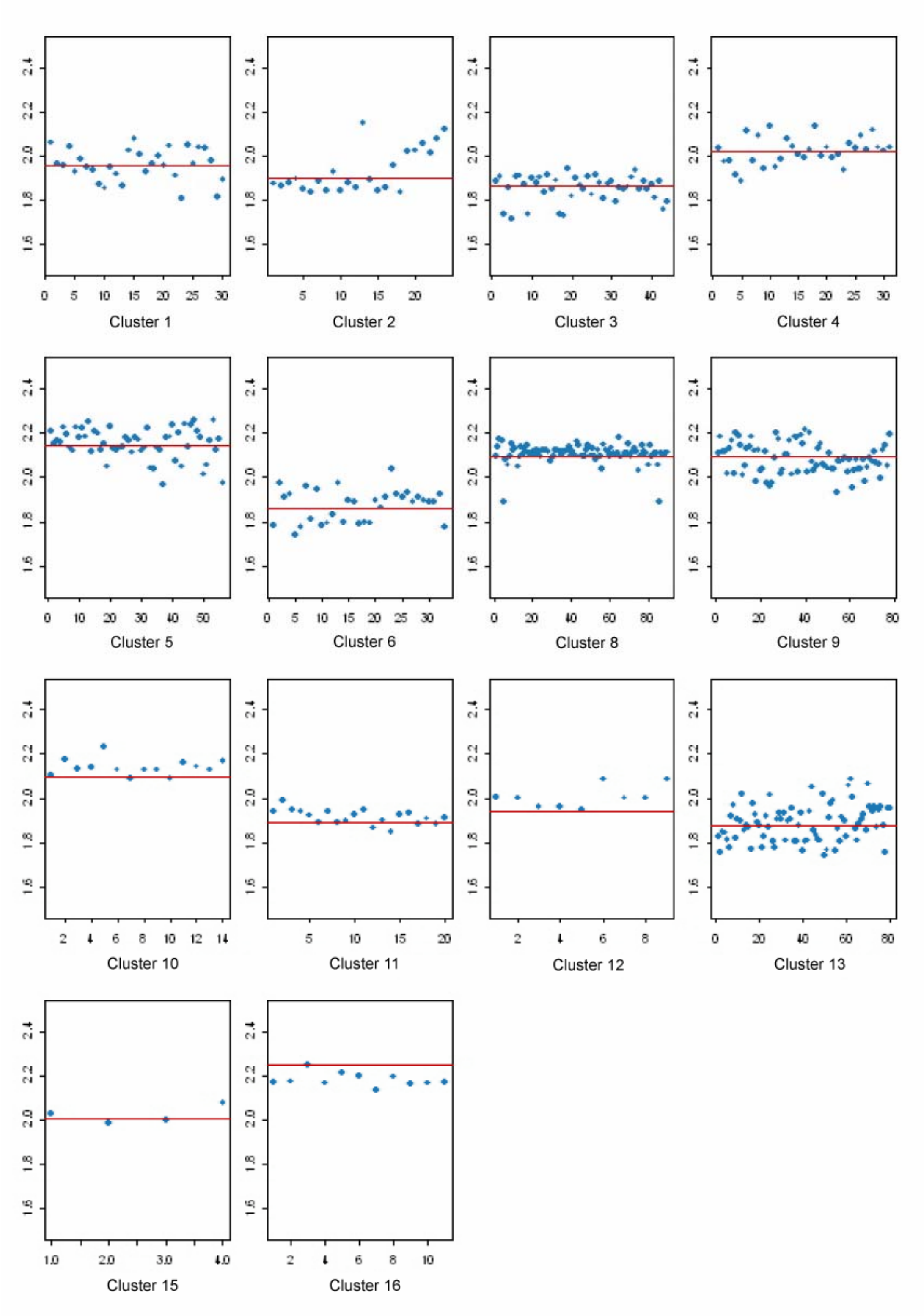

Figure 23 ARI 20 years growth factor estimates, red line - cluster, blue dots - circles

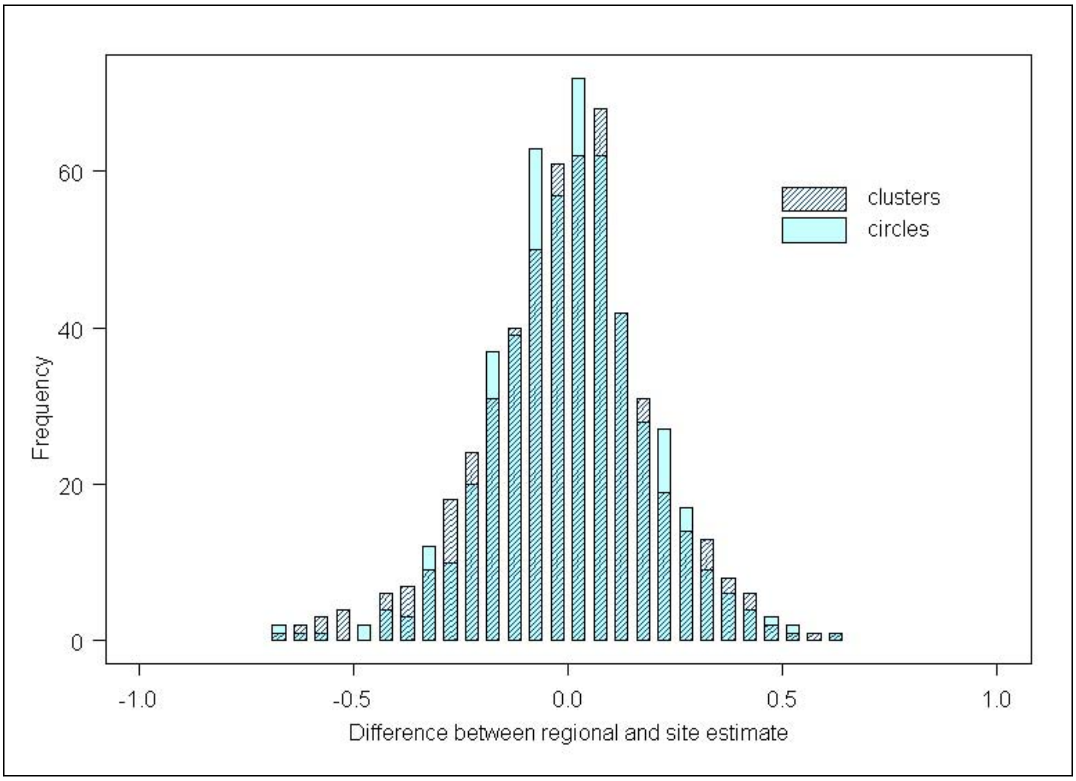

Figure 24 Histograms of differences between regional and site estimates

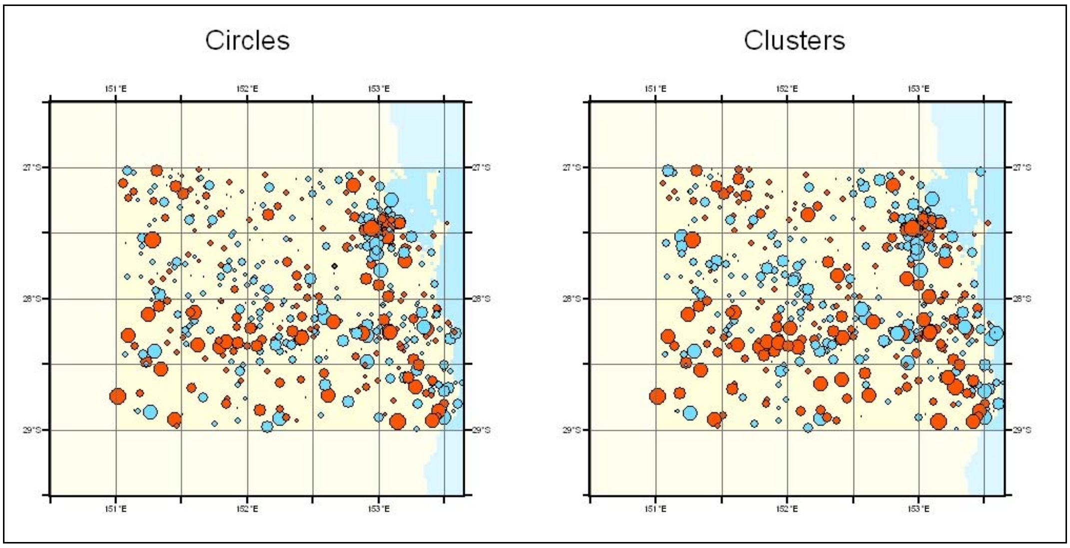

Figure 25 Differences between regional and site estimates of ARI 20 years growth factors (blue - regional estimate higher than site estimate, red - regional estimate lower than site estimate). Circle size indicates magnitude of difference.

9.3.4 Recommendation

The preferred approach is the circles approach. There are three arguments for choosing this approach over the cluster approach:

1. The circles approach is easy and quick to implement, no ‘manual’ adjustment is required.
2. Clusters defined may vary with duration, which in turn may lead to inconsistencies.
3. Statistics together with visual assessment against site estimates suggest that the circles approach (marginally) outperforms the cluster approach.

10 Recommendations

1. Annual maximum series (AMS) were abstracted for daily and recording gauges in the pilot area. The original data and the abstracted AMS were quality controlled. Estimates of index rainfalls and L-moments may be unreliable (see 8.1 'Estimating L-skewness'). A discordancy measure is a useful tool in identifying unusual sites.

2. The factor for conversion between 1-day rainfalls and 24 hour rainfalls was reviewed. It is recommended that a factor of 1.15 be used (see 4.1 'Restricted to unrestricted durations').

3. For design rainfall estimates at ARI 10 years and higher, analyses based on AMS are appropriate. For ARI below 10 years these estimates could be derived based on partial duration series (PDS). A conversion to transform from the AMS to a PDS scale has been reviewed. See Table 5 for the recommended values.

4. For durations and average recurrence intervals (ARI) considered in this study, no clear advantage could be found in using LH-moments (shifts η between one and four) over L-moments (see 5.2 LH-moments). Based on ARI estimates derived (assuming a Generalised Extreme Value distribution, GEV), the use of LH-moments instead of L-moments would lead to on average slightly lower rainfall estimates. In extreme cases, estimates based on LH-moments may be over 10% lower than those derived from a GEV based on L-moments. It is therefore recommended that L-moments be preferred over LH-moments.

5. Index rainfall for 24, 48 and 72 hours was mapped using thin-plate spline smoothing. The independent variables used in mapping are latitude, longitude, elevation and mean annual rainfall (see 6 'Mapping index rainfall').

6. For site frequency curves a GEV was assumed, since it gave an acceptable fit for over 90% of sites and for all of the durations (see 7.1 Site growth curves). For regional growth curves a GEV will be acceptable only in about three out of four cases. Investigations show however that it is justified to use a GEV only, provided estimates are required for moderate ARI (up to about 100 years) only (see 7.2 'Regional growth curves').

7. Partial Least Squares Regression (PLSR) models to infer statistics at durations below 24 hours from statistics of 24-hour AMS were derived. These models can be used to infer sub-daily index rainfall, L-CV and L-skewness from daily data.

8. Regions were defined using two different approaches and the performance of these approaches was assessed. A circles approach is recommended for daily and sub-daily durations (see 9.3 'Performance of regionalisation approaches').

11 References

Burn, D H (1990). Evaluation of regional flood frequency analysis with a region of influence approach', Water Resources Research, Vol. 26, pp. 2257-2265.

Bonnin, G M, Todd D , Lin B, Parzybok T, Yekta M and Riley D (2003). NOAA Atlas 14, 'Precipitation-Frequency Atlas of the United States, Vol. 1: Semiarid Southwest (Arizona, Southeast California, Nevada, New Mexico, Utah)', U.S. Department of Commerce, NOAA/NWS, Silver Spring, Maryland. (available online at: http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_docs.html)

Dalrymple, T (1960). 'Flood Frequency Analyses, Manual of Hydrology: Part 3, Flood Flow Techniques', USGS Water Supply Paper 1543-A.

Durrans, S R and Kirby, J T (2004). 'Regionalisation of extreme precipitation estimates for the Alabama rainfall atlas', Journal of Hydrology, Vol. 295, pp. 101-107.

Dwyer, I J and Reed D W (1995). 'Allowance for discretization in hydrological and environmental risk estimation (ADHERE)', Wallingford, Institute of Hydrology.

Faulkner, D (1999). 'Flood Estimation Handbook Vol 2: Rainfall frequency estimation', Institute of Hydrology, Wallingford.

Geladi, P and Kowalski, B R (1986). 'Partial Least Squares Regression: A Tutorial', Analytica Chimica Acta vol. 185, no. 12, pp. 1-5.

Haan C T (2002). 'Statistical Methods in Hydrology', Second ed., Iowa State Press.

Hartigan, J A and Wong M A (1979). 'Algorithm AS 136: A K-means clustering algorithm'. Applied Statistics, 28, 100-108

Hershfield, D M (1961). 'Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years'. Weather Bureau Technical Paper No. 40, U.S. Weather Bureau Washington, D.C., 115 pp.

Hosking, J R M (2000). 'Fortran routines for use with the method of L-moments, Version 3.03'. Research report. (available online at: http://www.research.ibm.com/people/h/hosking/lmoments.html)

Hosking, J R M and Wallis, J R (1997). 'Regional frequency analysis: An approach based on L-moments', Cambridge University Press.

Hutchinson, M F (1998). 'Interpolation of Rainfall data with thin plate smoothing splines - part 1: two dimensional smoothing of data with short range correlation, Journal of Geographic Information and Decision Analysis 2 (30), 139-151.

Hutchinson, M F (2004), 'Anusplin version 4.3, User Guide', The Australian National University, Centre for Resources and Environmental Studies, Canberra.

Institute of Hydrology (1999). 'Flood Estimation Handbook'. Institute of Hydrology, Wallingford.

Institution of Engineers, Australia (1987). Australian Rainfall and Runoff - A Guide to Flood Estimation, Vol. 2, R P Canterford (ed.), Institution of Engineers, Australia, Barton, ACT, 1987.

Langbein, W B (1949). 'Annual floods and the partial duration method'. Trans. Amer. Geophys. Union 30, 879-881.

Lin, B, Bonnin, B, Todd, D, Parzybok, T, Yekta, T and Riley, D (2004). 'Regional Frequency Studies of Annual Extreme Precipitation in the United States using Regional L-moments Analysis', International Ocean-Atmosphere Conference, Chinese-American Oceanic and Atmospheric Association (COAA), June 2004, Beijing, China.

Madsen, H , Rasmussen P F and Rosbjerg D (1997). 'Comparison of annual maximum series and partial duration series methods for modelling extreme hydrologic events: 1. At-site modeling', Water Resources Research, Vol. 33, No. 4, pp. 747-757.

Michel, C (2005). "Discussion of 'Sampling Adjustment Factors for Rainfall Recorded at Fixed Time Intervals' ", Journal of Hydrologic Engineering, vol. 10 , no. 1, pp 89-90.

Miller J F, Frederick R H and Tracey R J (1973). 'Precipitation-Frequency Atlas of the Western United States', U.S. Department of Commerce.

Nandakumar, N, Schöpf, J M, Weinmann, P E and Ruprecht, J K (2004). 'Estimation of Spatial dependence for application of the CRC-FORGE method in Western Australia', Proceedings International Conference on Storms, 5-9 July 2004, Brisbane, Australia.

Pierrehumbert C L (1972). 'Short Period Rainfall Intensity Analysis'. Working Paper 156, Bureau of Meteorology, Melbourne, Australia.

Pilgrim, D H, (ed). 1987. Australian Rainfall and Runoff - A guide to Flood Estimation, Institution of Engineers, Australia, Barton, ACT.

Thompson, C S (2002). 'The High Intensity Rainfall Design System: HIRDS', Paper presented to International Conference on Flood Estimation, 6-8 March 2002, Bern, Switzerland.

Venables, W N and Ripley, B D (2002). 'Modern Applied Statistics with S', Springer-Verlag, New York.

Wang, Q J (1996). 'Direct sample estimators of L-moments', Water Resources Research, Vol. 32, No. 11, pp. 3617-3619.

Wang, Q J (1997), 'LH moments for statistical analysis of extreme events', Water Resources Research, Vol. 33, No. 9, pp. 2841-2848.

Wang, Q J (1998). 'Approximate goodness-of-fit tests of fitted generalized extreme value distributions using LH moments', Water Resources Research, vol. 34, no. 10, pp. 3497-3502.

Ward, J (1963) 'Hierarchical grouping to optimize an objective function', Journal of the American Statistical Association 58 236-244.

Weiss L L (1964). 'Ratio of true to fixed-interval maximum rainfall', J. Hydraul. Div. Proc. ASCE 90, 77-82.

Young, C B and McEnroe, B M (2003). 'Sampling Adjustment Factors for Rainfall Recorded at Fixed Time Intervals', Journal of Hydrologic Engineering, Vol. 8, No. 5, pp 294-296.

12 Glossary

Annual Exceedence Probability (AEP) - The probability that a given rainfall total for a given duration will be exceeded in any one year.

Annual Maximum Series (AMS) - a series that contains only the event with the largest magnitude that occurred in each year.

ArcMap - a commercial geographical information system.

Average Recurrence Interval (ARI) - Generally, and when applied to the PDS: the average, or expected, value of the periods between exceedences of a given rainfall total for a given duration. When specifically applied to the AMS: the average, or expected, number of years (integer) between years in which there are one or more exceedences of a given rainfall total for a given duration.

Confidence limits - The upper and lower boundaries between which the value of a particular parameter is estimated to lie with a certain level of confidence (e.g. 95%).

Conversion factors - Factors to convert parameters calculated from an annual maximum series to (estimates of) those calculated from a partial duration series; factors to convert parameters calculated from a restricted (fixed) time series to (estimates of) those calculated from an unrestricted (continuous) time series.

Discordancy - A measure of how different or 'discordant' a site is, in terms of L-moments compared to a whole group of sites. Refer to Hosking and Wallis, 1997 for a formal definition.

Distribution - See frequency distribution.

Fixed (of values in a rainfall time series) – Rainfall observed over a significant fixed time interval (e.g. every three hours on the hour, 9 a.m. to 9 a.m. the next day).

Frequency curve - A graphical depiction of a *frequency distribution*; in this text, the ordinate is the magnitude and the abscissa the frequency.

Frequency distribution – The pattern of variation of a variable. The distribution records the numerical values of the variable and how often each value occurs.

General Extreme Value distribution (GEV) – A theoretical distribution often found to fit frequency distributions of rainfall and flood events.

Growth curve - A *frequency curve* scaled by index rainfall. The variate is therefore dimensionless.

Growth factor - The ratio of the size of an event at a given frequency to the index variable.

Index, index variable (rainfall or flood) - The scale parameter of an individual site in a region - usually the mean or median of the variable (e.g. mean annual maximum 24-hour rainfall) at that site, used in regionalization.

L-CV - The L-moment ratio "analogous to the ordinary coefficient of variation, C_v; but not an abbreviation of 'L-coefficient of variation': in words it would be more properly described as Coefficient of L-variation.')". (Quote from Hosking and Wallis, 1997)

LH-Moments - A generalisation of L-moments, introduced for characterising the upper part of distributions and larger events in data.

L-Moment ratios - Frequently, and often in this text, referred to as *L-moments*. Dimensionless versions of L-moments achieved by dividing the higher-order L-moments by the scale measure λ_2 . (Based on Hosking and Wallis, 1997).

L-Moments - A term often, and in this text, used for *L-moment ratios*. The expected value of certain *linear* combinations of the elements of an ordered sample and multiplied, for numerical convenience by scalar constants. They contain information about the location, scale and shape of the distribution from which the sample was drawn. The "L" in L-moments emphasizes the construction of L-moments from linear combinations of order statistics. (Based on Hosking and Wallis., 1997).

L-Skewness - The L-moment ratio analogous to skewness.

Moments - The mean value of the power of a variate. In a univariate distribution, the first moment is the arithmetic mean of that distribution, the second moment is the mean of its squares and so on.

Parameter(s) (of a frequency distribution) **-** Numbers, often identified with location, scale and shape, which form part of the equation defining the distribution of a population. Parameters may be estimated from moments or L-moments of a sample.

Partial Duration Series (PDS) – A series of independent events of magnitude above a pre-selected base.

Partial Least Squares Regression (PLSR) – A regression method which utilises the correlation amongst the predictands as well as amongst the predictors.

Pluviograph - In the context of this report, a continuous recording raingauge that is part of the standard network of Bureau of Meteorology stations.

R^2 Value - The square of the correlation coefficient. The fraction of the variation in the values of the predictand that is explained by the regression on the predictors.

Region - A set of sites whose *frequency distributions* are (after appropriate scaling) considered to be approximately the same. (After Hosking and Wallis, 1997)

Regionalisation - A method of using data from several neighbouring sites (each of limited observation period) to provide estimates of the frequency of rare events with greater confidence than would be available for a single site. The area of regionalisation must be reasonably homogeneous with respect to the variable studied.

Restricted (of values in a rainfall time series) **-** Rainfall observed over a fixed time interval (e.g. 9 a.m. to 9 a.m. the next day).

Sliding (of values in a rainfall time series) - Rainfall recorded essentially continuously in time (e.g. 1 minute, 6 minutes intervals).

Unrestricted (of values in a rainfall time series) **-** Rainfall recorded essentially continuously in time (e.g. 1 minute, 6 minutes intervals).

13 Acronyms

AEP - Annual Exceedence Probability

ALERT – **A**utomated **L**ocal **E**valuation in **R**eal-**T**ime

ANUSPLIN - Software developed at the Australian National University (ANU), Canberra, employing the technique of thin-plate SPLINe smoothing.

ARR, ARR87 - Australian Rainfall and Runoff, 1987.

CDIRS - Computerised Design IFD Rainfall System

DEM - Digital Elevation Model

DIPNR - Department of Infrastructure, Planning and Natural Resources, NSW

DNRM - Department of Natural Resources and Mines, Queensland (now DNRME)

DNRME - Department of Natural Resources, Mines and Energy, Queensland (formerly DNRM)

GEV - General Extreme Value (distribution)

HDSC - Hydrometeorological Design Studies Center, U.S.A.

HIRDS - High Intensity Rainfall Design System (NIWA)

IFD - Intensity-Frequency-Duration (graph or table)

MAE - Mean Absolute Error

MAR - Mean Annual Rainfall

NCC - National Climate Centre (Australian Bureau of Meteorology)

NIWA - National Institute of Water and Atmospheric Research (New Zealand)

PDS - Partial Duration Series. Series of maxima

PLSR - Partial Least Squares Regression

PRESS - Predictive Residual Error Sum of Squares

QC - Quality control

RMSE - Root Mean Square Error

SKM - Sinclair Knight Mertz Pty Ltd (consulting engineers)

14 Appendix

This appendix contains listings of sub-daily gauges used in the pilot study. Listed are gauge numbers, gauge types (or source of data), the station names, latitude and longitude (in decimal degree) and the record length for gauges recording at sub-daily durations.

Type and Source

'Pluvio' stands for pluviograph (run by the Bureau of Meteorology). 'ALERT' denotes gauges from the ALERT network (Queensland). Gauges marked 'DIPNR/DNRM' are run by the Department of Infrastructure, Planning and Natural Resources NSW (DIPNR) and the Department of Natural Resources and Mines Queensland (DNRM).

Record length

This is the number of annual maxima that were actually used. For some gauges, this number is significantly lower than suggested by the start and end year of a record. Annual maxima were abstracted for years which had at least 10 months of data (considering only those months where no more than 25% of the data were missing) and which were overall at least 60% complete.

Gauge number	Type/ Source	Site	Latitude	Longitude	Record length
40004	Pluvio	AMBERLEY AMO	-27.63	152.711	41
40019	Pluvio	BENARKIN FOREST STATION	-26.9	152.15	18
40082	Pluvio	UNIVERSITY OF QUEENSLAND GATTON	-27.544	152.338	38
40112	Pluvio	KINGAROY PRINCE STREET	-26.554	151.846	22
40135	Pluvio	MOOGERAH DAM	-28.03	152.553	35
40189	Pluvio	SOMERSET DAM	-27.117	152.555	22
40192	Pluvio	SPRINGBROOK FORESTRY	-28.226	153.279	19
40197	Pluvio	MT TAMBORINE FERN ST	-27.969	153.195	18
40214	Pluvio	BRISBANE REGIONAL OFFICE	-27.478	153.031	84
40223	Pluvio	BRISBANE AERO	-27.418	153.114	50
40282	Pluvio	NAMBOUR DPI	-26.643	152.939	41
40318	Pluvio	KIRKLEAGH	-27.026	152.564	30
40496	Pluvio	CALOUNDRA WATER TREAT	-26.793	153.129	27
40584	Pluvio	HINZE DAM	-28.048	153.288	22
40606	Pluvio	UPPER MUDGEERABA WATER	-28.106	153.329	22
40609	Pluvio	ELANORA WATER TREAT	-28.118	153.446	23
40677	Pluvio	MAROON DAM	-28.175	152.655	24
40726	DIPNR/DNRM	Dandabah Standalone	-26.882	151.598	8
40785	ALERT	CAROLE PARK ALERT	-27.602	152.951	10
40788	ALERT	FORESTDALE(JOHNSON RD)AL	-27.658	153	11
40790	ALERT	MT GRAVATT ALERT	-27.547	153.072	11
40793	ALERT	LYONS ALERT	-27.763	152.837	12
40839	ALERT	BRISBANE (BCC) ALERT	-27.467	153.022	9
40844	ALERT	BEECHMONT ALERT	-28.135	153.188	11
40845	ALERT	BINNA BURRA ALERT	-28.201	153.187	11

Gauge number	Type/ Source	Site	Latitude	Longitude	Record length
40846	ALERT	CLEARVIEW ALERT	-28.002	153.308	11
40847	ALERT	HINZE DAM ALERT	-28.052	153.282	10
40848	ALERT	SPRINGBROOK LOWER ALERT	-28.208	153.27	10
40876	ALERT	WILSONS PEAK ALERT	-28.239	152.486	9
40878	ALERT	WATERFORD ALERT	-27.695	153.136	10
41044	Pluvio	HERMITAGE	-28.206	152.1	46
41060	Pluvio	LEYBURN	-28.009	151.586	30
41140	Pluvio	WAMBO SHIRE COUNCIL	-27.186	151.255	33
41175	Pluvio	STANTHORPE (GRANITE BELT HRS)	-28.621	151.953	30
41467	Pluvio	TOOWOOMBA CITY COUNCIL	-27.567	151.885	24
54104	Pluvio	PINDARI DAM	-29.39	151.245	31
57095	Pluvio	TABULAM (MUIRNE)	-28.755	152.45	31
58013	Pluvio	CONDONG SUGAR MILL	-28.317	153.433	15
58026	Pluvio	GREVILLIA (SUMMERLAND WAY)	-28.441	152.83	21
58044	Pluvio	NIMBIN POST OFFICE	-28.597	153.223	16
58081	Pluvio	UPPER MONGOGARIE (KIMBERLEY)	-28.967	152.817	11
58129	Pluvio	KUNGHUR (THE JUNCTION)	-28.466	153.263	23
58131	Pluvio	ALSTONVILLE TROPICAL FRUIT RESEARCH STA	-28.852	153.456	38
58158	Pluvio	MURWILLUMBAH (BRAY PARK)	-28.341	153.378	30
203030	DIPNR/DNRM	Myrtle Ck @Rappville	-29.112	152.998	8
203900	DIPNR/DNRM	Richmond R @Kyogle	-28.617	153	8
204002	DIPNR/DNRM	Clarence R @Tabulam	-28.887	152.565	11
204036	DIPNR/DNRM	Cataract Ck @Sandy Hill(below Snake Ck)	-28.933	152.217	26
204900	DIPNR/DNRM	Clarence R @Baryulgil	-29.198	152.592	9
416023	DIPNR/DNRM	Deepwater R @Bolivia	-29.3	151.933	9
416310	DIPNR/DNRM	Dumaresq River at Farnbro	-28.919	151.584	10
422338	DIPNR/DNRM	Canal Creek at Leyburn	-28.032	151.586	10
422394	DIPNR/DNRM	Condamine River at Elbow Valley	-28.372	152.141	8
540000	DIPNR/DNRM	Atkinson Dam Standalone	-27.421	152.45	10
540002	DIPNR/DNRM	Yarrahappini GS145014A	-27.833	152.986	10
540004	DIPNR/DNRM	Wolfdene GS145196A	-27.783	153.19	10
540005	DIPNR/DNRM	Glenhurst GS146002B	-28	153.31	11
540006	DIPNR/DNRM	Burnett Ck. Standalone	-28.276	152.571	8
540008	DIPNR/DNRM	Boat Mt. GS143010B	-26.979	152.285	9
540009	DIPNR/DNRM	Upper Caboolture GS142001A	-27.098	152.891	12
540010	DIPNR/DNRM	Brooklands GS 136203A	-26.738	151.819	10
540013	DIPNR/DNRM	Bellbird Ck GS 138110A	-26.629	152.704	11
540023	DIPNR/DNRM	Cainbable Ck. Standalone	-28.123	153.079	26
540030	DIPNR/DNRM	Moy Pocket GS 138111A	-26.528	152.743	9
540035	DIPNR/DNRM	Adams Br. GS143110A	-27.833	152.509	9
540038	DIPNR/DNRM	Indooroopilly Standalone	-27.508	152.977	9
540045	DIPNR/DNRM	Forest Home GS145003B	-28.202	152.77	9

Gauge number	Type/ Source	Site	Latitude	Longitude	Record length
540046	DIPNR/DNRM	Dieckmann's Br. GS145010A	-28.248	152.891	8
540047	DIPNR/DNRM	Croftby GS145011A	-28.149	152.582	10
540051	DIPNR/DNRM	North. Stradbroke Island Bore 14400009A	-27.6	153.438	8
540054	ALERT	LITTLE NERANG DAM ALERT	-28.147	153.285	11
540055	DIPNR/DNRM	Bribie Island Bore 14100090A	-27.008	153.169	8
540078	ALERT	MARSDEN (FIRST AVE) AL	-27.669	153.103	8
540083	ALERT	NELSONS ALERT	-26.642	152.993	8
540085	ALERT	MAPLETON ALERT	-26.626	152.865	9
540092	ALERT	YANDINA ALERT	-26.564	152.938	9
540094	ALERT	PICNIC POINT ALERT	-26.646	153.081	9
540095	ALERT	DUNETHIN ROCK ALERT	-26.578	153.008	9
540137	ALERT	NAMBOUR ALERT	-26.63	152.958	9
540254	ALERT	MUDGEERABA ALERT	-28.085	153.372	10
541041	ALERT	DALBY ALERT	-27.184	151.264	9
541042	ALERT	MOFFATT ALERT	-27.097	151.398	11
541043	ALERT	CLYDESDALE ALERT	-27.193	151.486	11
541044	ALERT	COORINGA ALERT	-27.09	151.596	11
541045	ALERT	MT BRIGALOW ALERT	-27.202	151.786	10
541046	ALERT	MT MOWBULLAN ALERT	-26.889	151.6	10

Table A 1 List of sub-daily gauges used in the pilot study

15 Index

HYDROLOGY REPORT SERIES

HRS1	Temporal Distributions Within Rainfall Bursts	September 1991
HRS2	Analysis of Australian Rainfall and Rainday Data with Respect to Climatic Variability and Change	February 1992
HRS3	Catalogue of Significant Rainfall Occurrences Over Southeast Australia	October 1995
HRS4	Development of the Generalised Southeast Australia Method for Estimating Probable Maximum Precipitation	August 1996
HRS5	Temporal Distributions of Large and Extreme Design Rainfall Bursts over Southeast Australia	December 1998
HRS6	Rainfall Antecedent to Large and Extreme Rainfall Bursts over Southeast Australia	December 1999
HRS7	Development of the Method of Storm Transposition and Maximisation for the West Coast of Tasmania	January 2001
HRS8	Revision of the Generalised Tropical Storm Method for Estimating Probable Maximum Precipitation	August 2003
HRS9	Catalogue of Significant Rainfall Occurrences of Tropical Origin Over Australia	February 2004
HRS10	A Pilot Study to Explore Methods for Deriving Design Rainfalls for Australia - Part 1	June 2005