Australian Government
Bureau of Meteorology

HYDROLOGY REPORT SERIES

HRS Report No. 11

A PILOT STUDY TO EXPLORE METHODS FOR DERIVING DESIGN RAINFALLS FOR AUSTRALIA - PART 2

Water Division
Melbourne
January 2009

Australian Government
Bureau of Meteorology

HYDROLOGY REPORT SERIES

HRS Report No. 11

A PILOT STUDY TO EXPLORE METHODS FOR DERIVING DESIGN RAINFALLS FOR AUSTRALIA - PART 2

Dörte Jakob, Karin Xuereb, Jeanette Meighen, Brian Taylor

Hydrometeorological Advisory Service
www.bom.gov.au/hydro/has/
Water Division
Melbourne
January 2009

Contents

Page

List of Figures

List of Tables

BACKGROUND

Use of design rainfall information

Revised design rainfall estimates will be the key building block for a large number of sections in a revision of 'Australian Rainfall and Runoff' (ARR87) (IEAust, 1987). Design rainfall information is required for estimating design floods, which in turn are required for the design of hydraulic structures and in floodplain management. Another ten or so projects scheduled to be undertaken as part of the ARR87 revision will be building on results from the IFD revision to be undertaken at the Bureau of Meteorology.

Following a design event approach, there are two major steps required to derive design flood estimates: firstly, computation of the depth of rainfall excess and secondly, conversion of the rainfall excess into a flood hydrograph. To derive estimates of rainfall excess, information is required about the design rainfall intensity (together with spatial and temporal variability), initial and continuing losses, as well as correlations between these variables.

Some of the limitations of the current methods are well outside the scope of the proposed projects (such as loss models for instance). For most input variables, typical values are assumed (such as areal reduction factors and temporal patterns) and these input variables are assumed to be independent of each other. It is expected that significant improvement in modelling design floods could be gained by considering these variables as stochastic variables and by accounting for correlation between these variables.

Anticipated improvements

As a precursor to a potential review of ARR87, a project to review the methods for estimation of design rainfall commenced in November 2003 in the Hydrometeorological Advisory Service of the Bureau of Meteorology.

Apart from providing updated and more reliable information, it would be possible to

- increase the range for which estimates are supplied (for instance it would be desirable to include estimates at durations below 5 minutes and seasonal estimates)
- supply uncertainty estimates (only a very basic assessment of uncertainty has been undertaken for the current version of ARR87)
- supply data in user-friendly formats (for instance shape files for use with a geographic information system (GIS)).

More importantly, a shift in methodology would allow addressing the following two issues: Firstly, include detailed information on spatial and temporal variability of design rainfall events and secondly, allow assessing potential effects of climate change.

Three projects for which Bureau has been invited to submit proposals are focussed on the revision of design rainfall and related spatial and temporal patterns. The relevant parts setting out methods for design rainfall estimation in 'Australian Rainfall and Runoff' (ARR) have last been updated in 1987. Clearly, there have been major developments over the last 20 years that mean the data availability has dramatically increased (in particular at durations below 24 hours) and it is only now possible to consider including remotely sensed data. New statistical techniques have emerged in the field of extreme value analysis, Bayesian approaches are starting to be used for applications in hydrology. Advances in hydrological modelling and computing have been made so that users are now increasingly in a position to use more complex input for models.

The needs of users requiring design rainfall information have also changed over the last 20 years. Increasingly, estimates are required for short durations (5 minutes and less) for urban design and demand for seasonal estimates (currently not available) has increased.

A pilot study to review methods for design rainfall estimation was undertaken at the Bureau of Meteorology. Comparison of preliminary estimates from this pilot study with ARR87 estimates showed differences of up to 30% at the 24-hour duration. Assessments show that these differences are mainly due to different methods rather than changes in data availability or record length.

To current best knowledge, there are perceived flaws in the method used to derive design rainfall estimates. For example, due to limited data availability particularly at short durations, meteorological interpretation was used to draw isohyets. Judging from data available now, it appears that orographic enhancement might have been overestimated in some regions.

The approach employed for the pilot study was based on regional frequency analysis. Frequency distributions were fitted using the technique of L-moments (linear combination of order statistics). A partial least squares regression approach was used to predict L-moments and index rainfall for sub-daily durations by making use of information at longer durations.

One of the basic assumptions behind the classic frequency approach is a stationary climate. Results from a study undertaken to assess potential effects of climate change on estimates of Probable Maximum Precipitation, (Jakob et al, 2008b) showed that while for most of Australia the frequency of heavy rainfall events has increased. In some parts, such events have become less intense and less frequent, and these changes are accompanied by a decline in annual rainfall totals.

The Australian public has over the last couple of years become increasingly aware of effects climate change has on our environment. Practitioners are questioning the validity of design rainfall estimates derived on the basis of short records which may not be representative of our current climate. Before assessing whether and to which extent climate change might affect design rainfall estimates, one needs to assess what representative estimates for a current climate are. Documentation on data and methods used to develop the currently valid estimates is insufficient to answer this question.

Purpose of this document

This document should be read in conjunction with HRS Report No.10 'A pilot study to explore methods for deriving design rainfalls for Australia - Part 1' (Jakob et al 2005). Work has continued since the publication of Part 1 in June 2005. Part 2 (this report) will expand on information given in Part 1, with particular emphasis on subdaily durations. Additional topics such as smoothing estimates across durations and fixing inconsistencies will also be covered. Not all parts of the methodology have been finalised (for instance construction of confidence intervals). In these cases possible approaches will be discussed. An overview of the procedure to deriving design rainfall estimates developed in the pilot study is given in Figure 1.

This document was written as a reference for the work undertaken in the second part of the pilot study. It is anticipated, that versions of this document will serve to relate relevant information to a varied readership. The purpose of this document is threefold:
firstly, to provide information to members of the ARR Revision Technical Committee,
secondly, to serve as a point of reference for the team undertaking the IFD revision and
thirdly, as a source of information to those that are not involved in the IFD revision but would like to get an overview over methods developed in and findings from the pilot study. Certain parts of this report may therefore be relevant to a particular audience only.

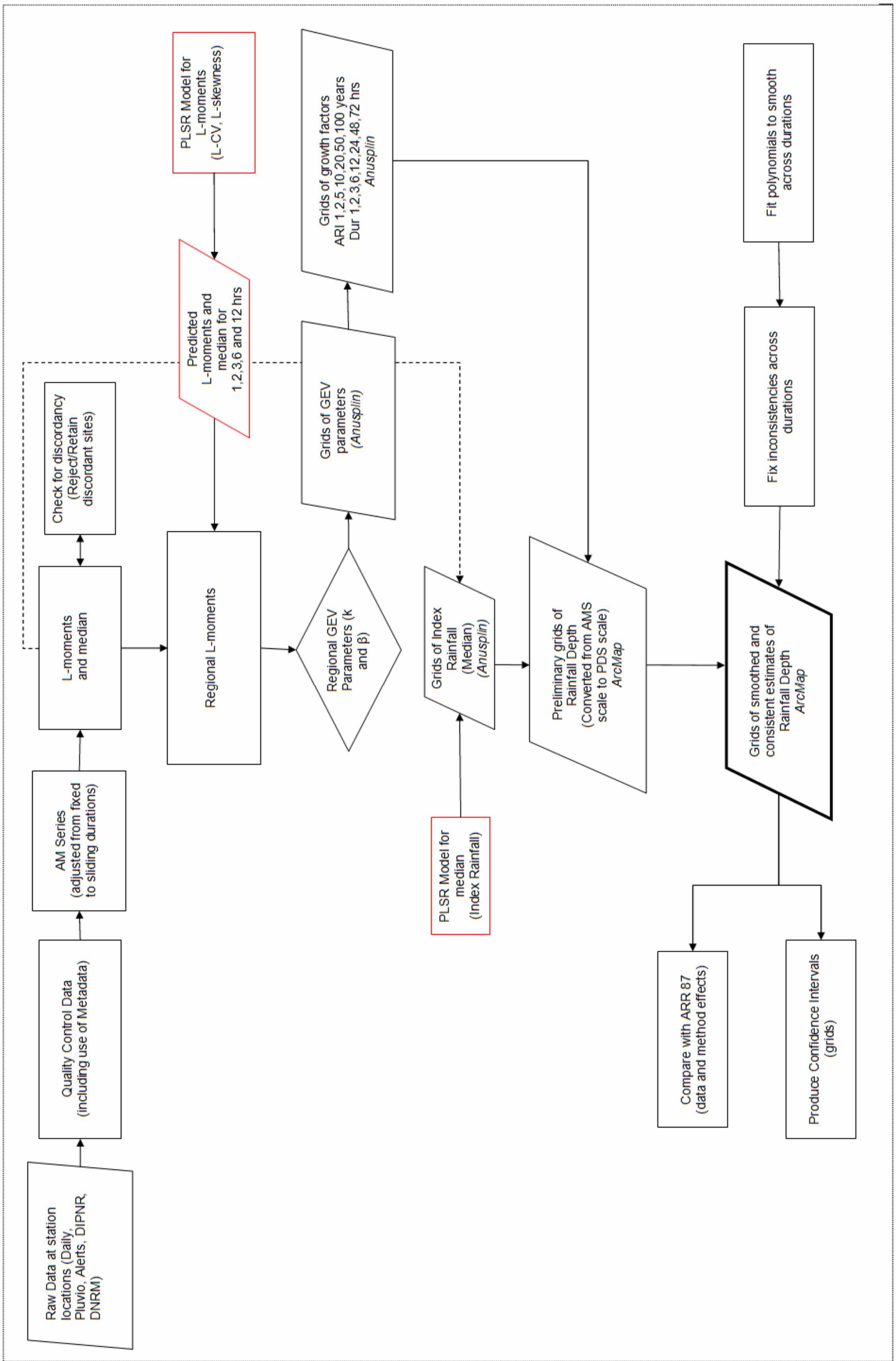


Figure 1 Flowchart for deriving design rainfall estimates at daily and subdaily durations. Boxes outlined in red refer to steps required for subdaily durations only.

1 UPDATING RAINFALL DATA AND METADATA

1.1 Additional Rainfall Data at Daily and Subdaily Durations

Minimum record lengths were set for daily and subdaily durations. As records at daily durations are typically long and stations provide good spatial coverage, a high threshold of 20 years or longer was chosen. Records of this length should allow the derivation of reliable estimates of L-skewness.

Record lengths at subdaily durations are typically much shorter and therefore it was necessary to choose a lower threshold. However, it was considered that records of less than 8 years were unsuitable for calculating meaningful L-moments (based on annual maximum series). Therefore the threshold for subdaily stations was set at 8 years.

In early 2007, a decision was made to include data which had become available since the start of the pilot study. This decision was taken on the basis that record lengths at subdaily durations would have increased and that additional sites with at least 8 years would be available. While additional stations were added in the process for subdaily durations, at daily durations there was no perceived need to increase the number of sites used in the analyses. However, where additional data had become available at daily durations, it was added to the record. Figure 2 and Figure 3 give an overview of the record lengths and spatial coverage at both daily and subdaily durations for the updated dataset.

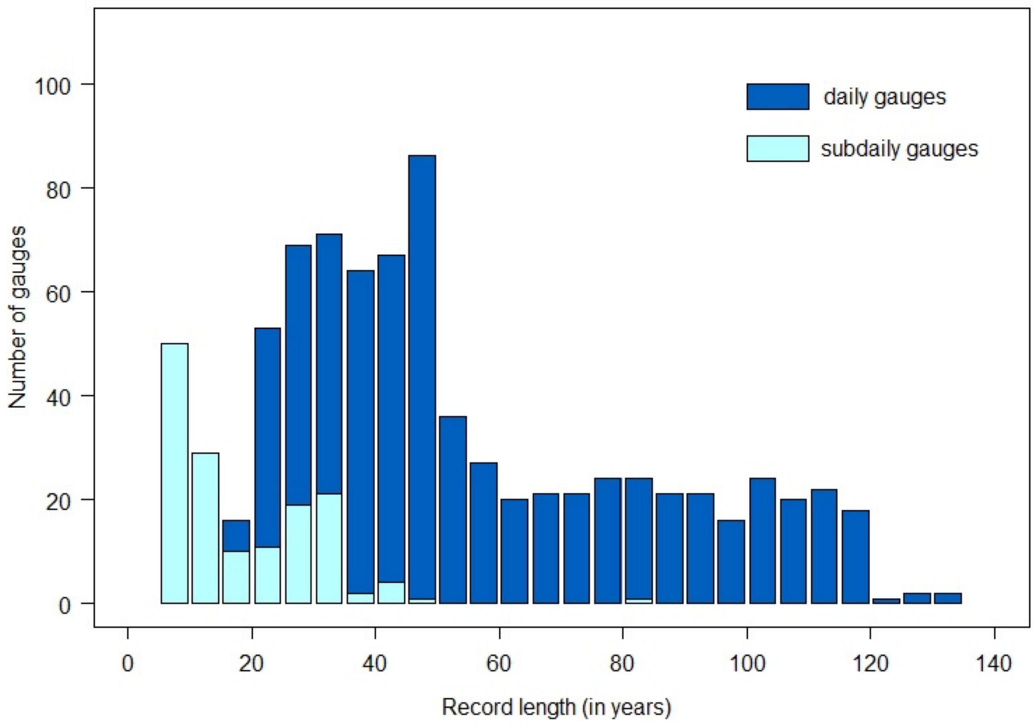


Figure 2 Updated datasets and record lengths for daily and subdaily gauges.

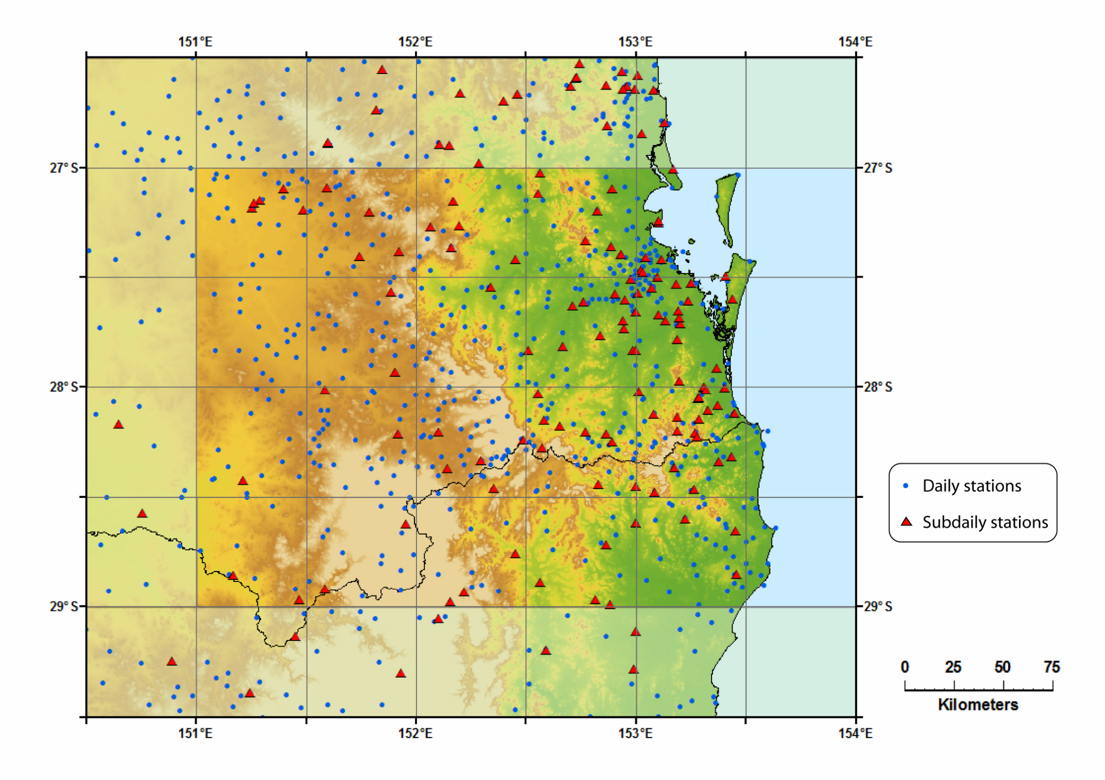


Figure 3 Location of stations used in pilot study.

1.2 Extracting Annual Maxima

Extreme value analysis was undertaken on the basis of annual maxima (AM) - the highest rainfall depth (for a given duration) was extracted for each suitable calendar year.

Seasonality of rainfall varies across Australia. For some climate zones, rainfall is distributed almost uniformly over the year whilst the tropical zones experience distinct dry and wet seasons. The Australian monsoon season lasts, depending on definition, from December to April.

As this wet season in tropical parts of Australia straddles calendar years, steps must be taken to avoid selecting events as annual maxima which occurred in the same season but in different calendar years. Alternative definitions based on 'water years' could be adopted as is used in European countries. However, considering the variability in seasonality of rainfall across Australia it is difficult to develop one universal definition of a water year for extracting annual maxima.

Alternatively (or in conjunction with annual maxima), the use of partial-duration series (PDS) (peaks over threshold) could be considered. Currently, an approximation is used to derive adjusted estimates for more frequent events (below ARI of 10 years). Other factors that make PDS appear a more suitable choice are short records and effects of non-stationarity (Khaliq et al, 2006).

Annual maxima are extracted using a dedicated set of Fortran programs, some of them developed in the Bureau of Meteorology's National Climate Centre. The programs are:

- pluvio_extract (extracts data to file for use by the following programs),
- pluvint (nominated fixed time intervals),
- pluvanal (extract monthly maxima for a station)
- ann_max_qc
- ann_max

Most of the programs are in need of being updated. The upcoming revision of design rainfall estimates makes this a high priority, creating an opportunity to tailor the programs to our needs, address known issues and include options to allow extracting partial duration series.

1.2.1 Missing Data

Since very few of the records analysed were 100% complete, a set of rules was defined to address missing data. This was done by defining a threshold of maximum percentage of missing values before a year became 'unacceptable' for extraction of an annual maximum. These thresholds were set to 25% missing data in a month (otherwise the month was rejected) and at least 10 'acceptable' months in a year in order for the year to be accepted.

1.2.2 Parameters and Thresholds used in other Countries

Different approaches are used in different countries.

The United States approach consists of a fairly complex set of criteria with details depending on duration (Bonnin et al, 2004). The three major factors are:

- number of hours missing,
- magnitude of AM,
- percentage of data missing (in particular during wet season).

The United Kingdom approach discounts years with more than 25% of data missing (Faulkner, 1999).

The New Zealand approach uses years with at least 11 months of record (which would mean more than 91% complete over the year). (Thompson, 2002)

1.2.3 'Add back in' Rule

A decision was required to judge which values represented 'valid annual maxima'. Ideally, the adopted approach would consider the number of missing values during the wet season and avoid excluding very large values as well as including very low AM (caused by large number of missing values).

When extracting annual maxima from pluviograph data, a minimum level of completeness for a given year is required. Three parameters were adopted:

- number of months for which this threshold applies (NUM_MON)
- minimum number of completeness for calendar months (PER_MON)
- number of months for the year (PER_YEAR)

For brevity, these were expressed in the form (NUM_MON **-** PER_MON **-** PER_YEAR), for example (10-75-90).

Previously used values are: NUM_MON 10 months, PER_MON 75% and PER_YEAR 90% (10-75-90). Using these settings, a significant number of years where a large event occurred which could be a valid AM (according to its magnitude) were rejected.

Additional series of AM were extracted using the following two sets of parameters:

- NUM_MON 0, PER_MON 1%, PER_YEAR 60% (0-1-60) (4 months could be missing entirely, year would still be accepted for AM extraction) and
- NUM_MON 0, PER_MON 1%, PER_YEAR 1% (0-1-1) (the year is accepted if any data is available).

Figure 4 shows the number of years judged acceptable under the different thresholds. Using the strictest set of criteria (10-75-60), 8 pluviographs were rejected because of their short

record length. There is little difference between the number of years rejected between the two less restrictive parameter sets, (0-1-60) and (0-1-1)

1.2.4 Comparing Thresholds for Rejecting Years

The annual maxima (and resulting frequency curves) were plotted for 87 pluviographs for two durations (1h and 24h). Frequency curves fitted to annual maxima selected for 0-1-1 and 0-1-60 typically show very good agreement. In most cases Average Recurrence Interval (ARI) estimates using the stricter criteria (10-75-90) are slightly higher than for the less rigid rules. This is demonstrated for a single station and duration in Figure 5. (Note that the plotting positions for the annual maxima change because the number of accepted values changes with threshold for accepting/rejecting years.)

It is suggested that the slight differences in estimates were due to the fact that some of the values considered AM following the less stringent rules have (probably rightfully) been excluded when using the more restrictive rules.

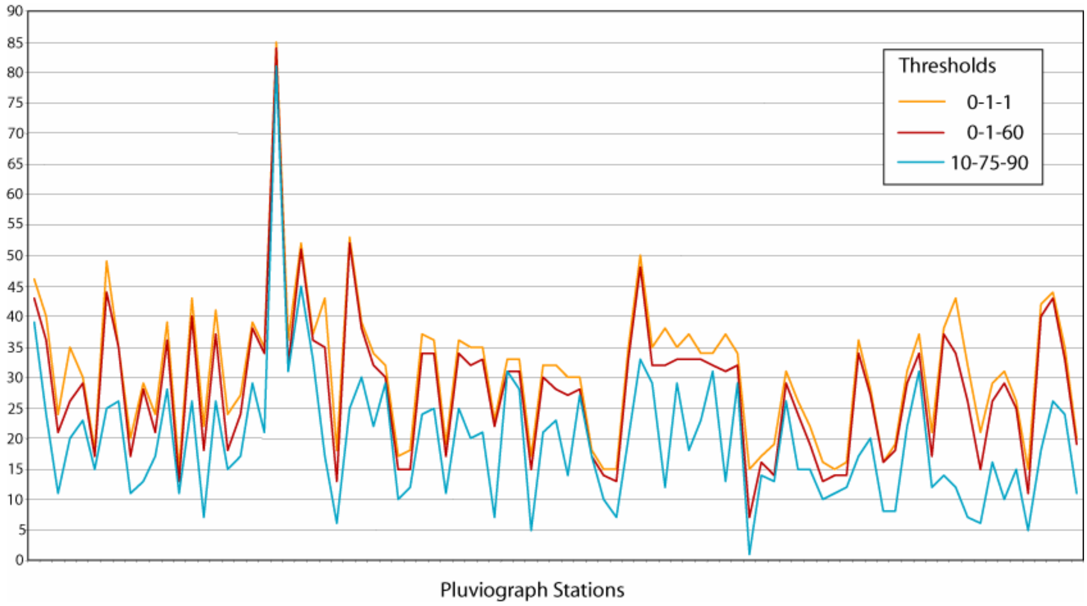


Figure 4 Number of years judged valid for extracting AM for three thresholds for missing data for all pluviographs in the pilot study area

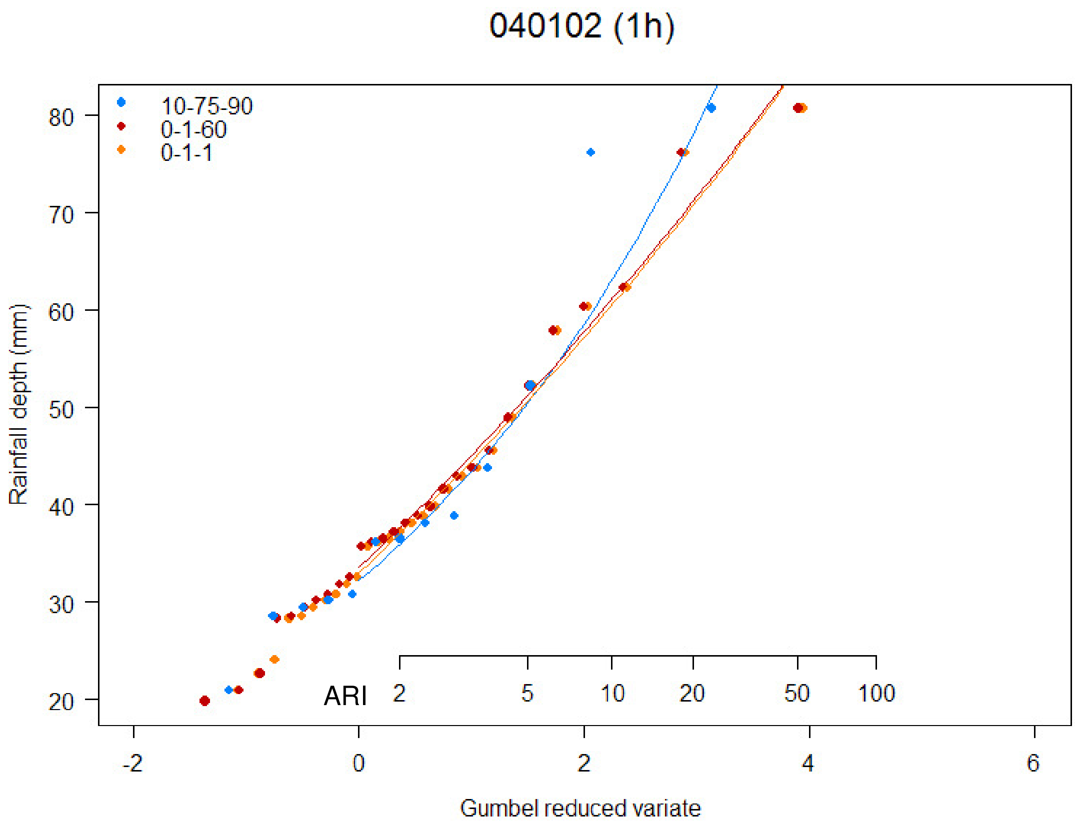


Figure 5 Annual maxima (points) and frequency curves (solid lines) derived for different thresholds for missing data for station 040102, duration 1h; different thresholds indicated by different colours.

However, in some instances the frequency curve fitted to the annual maxima extracted using the 10-75-90 combination is much lower than for the other two combinations. This appears to be caused by the fact that some of the most extreme AM were rejected. Often these would be the top 2 or 3 events, as shown in Figure 6 but these AM were always well above the median AM. Since record lengths differ considerably, a threshold in terms of percentile was chosen rather than using the top *n* events.

Excluding the 'real' AM from the series may lower the L-CV (and/or L-skew) for this station considerably which would impact on the regional estimates. Such sites might have been flagged as discordant and quality control on the AM would probably not be able to identify what is causing the problem. Figure 7 shows a case where the shape of the frequency curve is changed significantly when different thresholds for missing data are used. This appears to be the combined effect of removing some of the lowest values and some of the very high values.

Comparing the two durations that were studied (1 and 24h) it appears that this is more of an issue for the longer of the two durations although it is not quite clear what would be the cause of this difference. Using the criteria based on completeness, a year would be rejected independently of the magnitude of the maximum event for that year. This can have significant impact for some durations but might have little effect at other durations as shown in Figure 8 and Figure 9 where the most stringent criterion gives a very poor fit for the 24-hour duration but not for the 1hour duration.

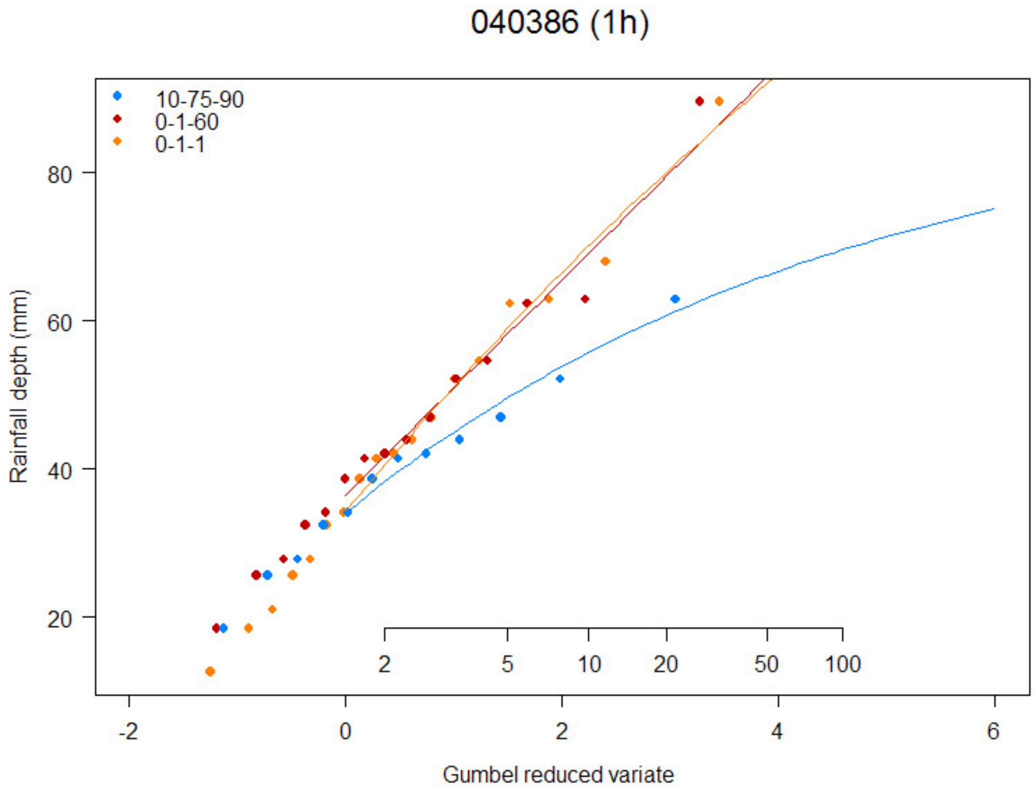


Figure 6 Annual maxima and frequency curves derived for different thresholds for missing data (station 040386, duration 1h).

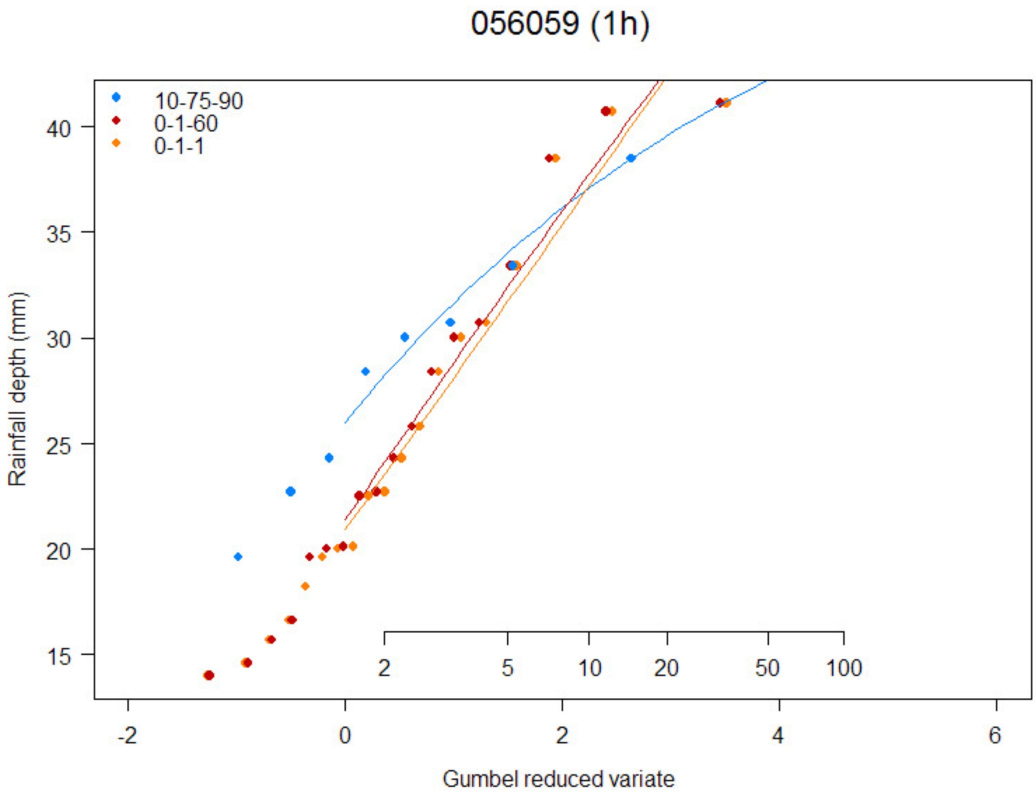


Figure 7 Annual maxima and frequency curves derived for different thresholds for missing data (station 056059, duration 1h).

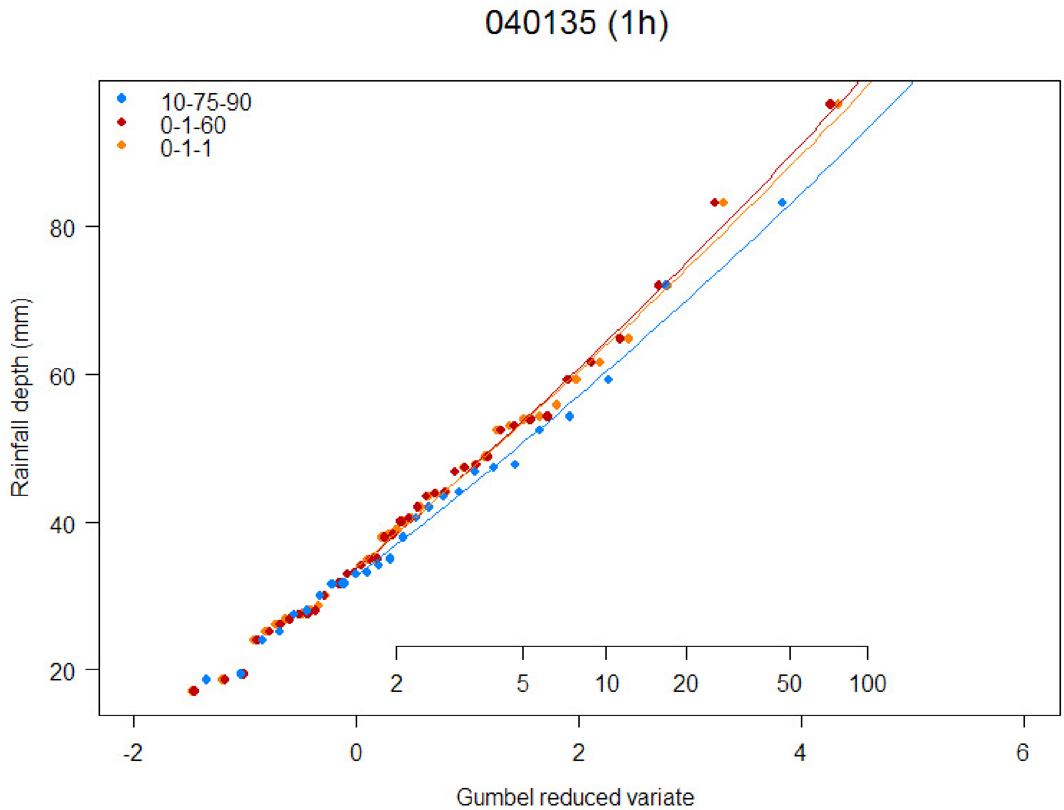


Figure 8 Annual maxima and frequency curves derived for different thresholds for missing data (station 040135, duration 1h).

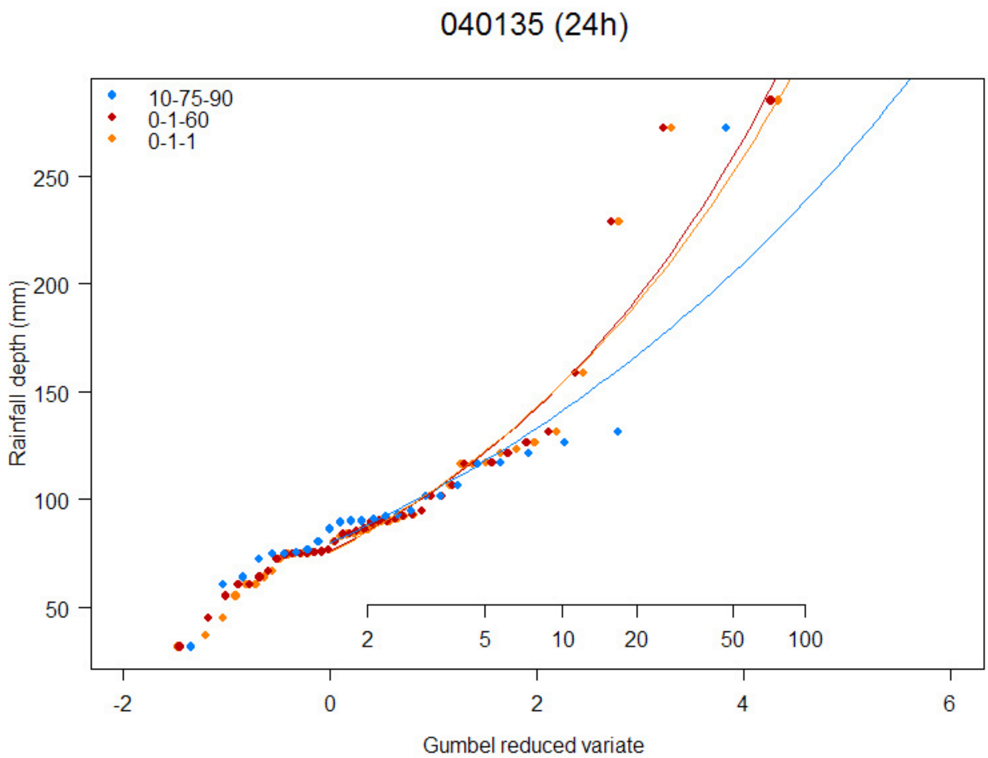


Figure 9 Annual maxima and frequency curves derived for different thresholds for missing data (station 040135, duration 24h).

Adopted rule

Using a fairly stringent approach (10-75-90), all the years where the AM was deemed 'valid' were selected. For years where the maximum value was rejected, the magnitude of this event was compared against the 90th percentile from the 0-1-1 combination. If the magnitude for the rejected value was above that threshold, it was included in the series (even though completeness for this year would not warrant its inclusion).

The number of 'additional' maxima at each station was documented and the record length for these sites was adjusted accordingly but not further investigations were undertaken.

1.3 Unflagged Accumulations

Daily rainfall is normally measured every day at 9:00 am. Occasionally, an observation is missed and the measured rainfall, when an observation is subsequently made, is then of rainfall that has occurred over a period of 2 days or more. Observers flag the period of missed observations by a -8888 and the rainfall that is then measured, which has accumulated over a period of 2 days or more, is indicated by a minus sign (-) to the left of the rainfall amount. No attempts have been undertaken in the pilot study to disaggregate these accumulations and such values were not included in the analyses.

Sometimes, accumulated daily rainfall is not flagged and rainfall on days when no observation was made, is entered as a zero (0) making such an observation indistinguishable from a day when an observation of zero rainfall was made. When an observation is subsequently made, the measured rainfall covers a period of 2 days or more. Therefore, rainfall that appears to have occurred over 1 day could actually have occurred over 2 or more days. The problem occurs most often after a weekend (on Saturdays and Sundays observations may not be made) and after public holidays, particularly long weekends. However, sometimes observations may not be made at other times of the year as well, presumably when the observer is away on holiday or is ill. (Viney and Bates, 2004)

These unflagged accumulations have the potential to affect estimates of design rainfall, especially where they constitute the highest ranking annual maximum event. Investigations were therefore undertaken to a) identify such cases, b) assess the effect unflagged accumulations have on site and regional estimates and c) experiment with methods to address the problem.

Daily rainfall data in the ADAM database is not quality controlled for unflagged accumulated daily rainfall. One way to determine if such rainfall is an accumulation is to compare the measured rainfall with that from nearby stations using a tool named 'Rainfall Near Comparison Report' (provided by NCC Data Management to quality control rainfall data in ADAM). An example of such a report is shown in Appendix A1. On Tuesday, 25 January 1927, 381 mm was recorded for station 040074. Comparison with nearby stations shows that 381 mm rainfall could not have occurred in one day. The rainfall amount of 381 mm for station 040074 was, erroneously, identified as the highest annual maximum rainfall for that station.

The annual maximum rainfall series for stations in the pilot study area was used to derive estimates of design rainfall for average recurrence intervals (ARI) of 1 year to 100 years. The inclusion of accumulated rainfalls in the annual maximum rainfall series for daily stations included in the study could potentially give results that over-estimate the design rainfall.

In the following sections, the methods adopted for adjusting unflagged daily rainfall accumulations and the effect such methods can have on design rainfall are described. It was expected that the effects of unflagged rainfall accumulations would be most marked at the 1-day duration since for longer durations, an accumulated rainfall total would approach the correct total. Hence this investigation focused on the 1-day duration. The impact of methods to adjust for flagged accumulated rainfall was not explored. In the pilot study, the effect of

unflagged accumulations has only been assessed for the region covered by the pilot study area.

However, 'the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high quality data set' was assessed by Viney and Bates (2004). The occurrence of unflagged accumulations for the 181 sites in the high-quality data set is shown in Figure 10. Each horizontal line represents one station, with the thin lines representing the years of operation for each station. Stations are grouped by state and territory.

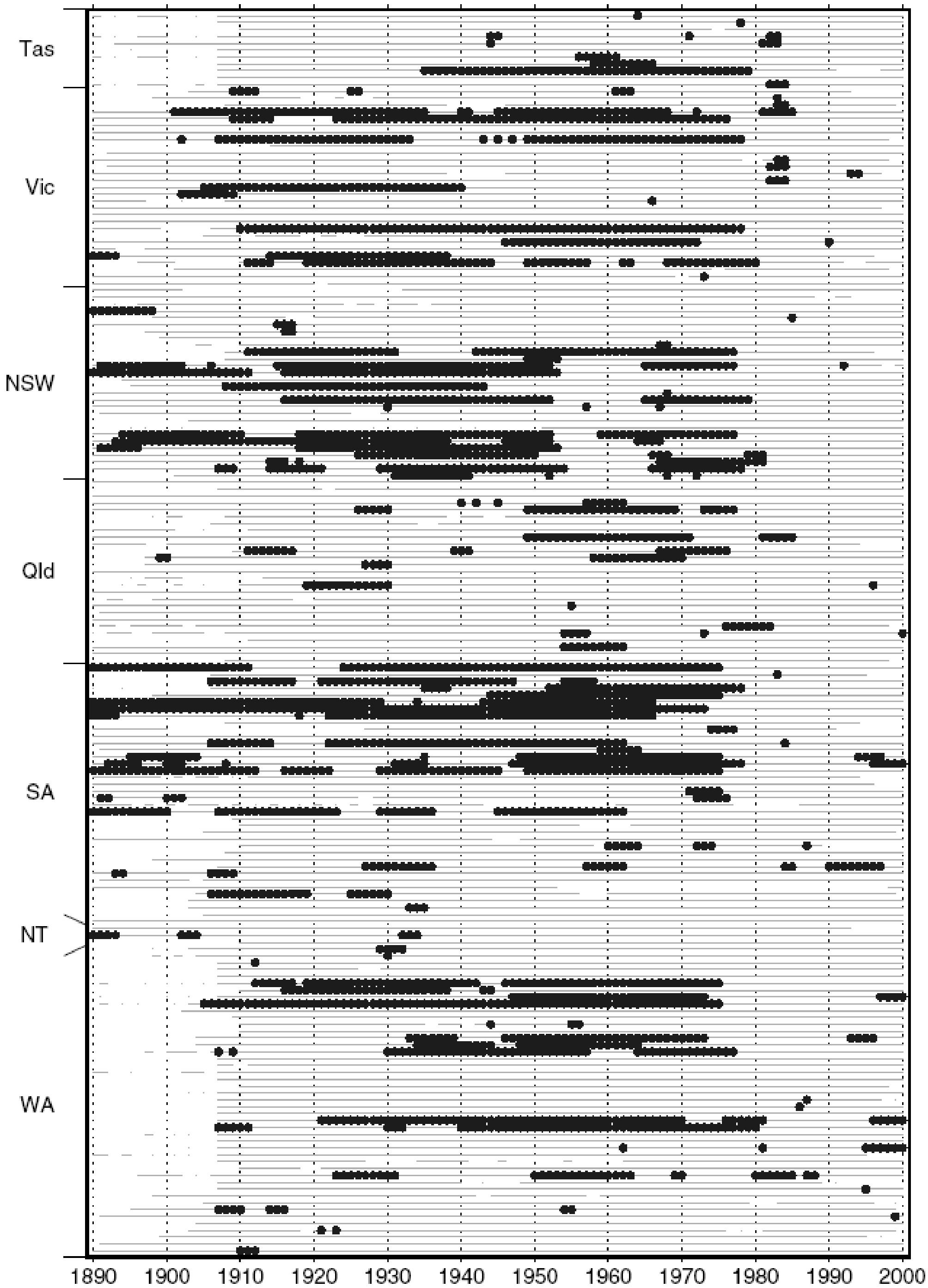


Figure 10 Occurrence of unflagged accumulations (circles) for each of the 181 gauges of the high-quality data set for the period 1890-2000. (Source: Viney and Bates, 2004)

Using the approach developed by Viney and Bates, Robert Smalley (pers. comm.) assessed the occurrence of unflagged accumulations for a more comprehensive set of sites. Figure 11 shows an example 'Abacus' plot indicating stations from the high-quality data set described in Lavery et al. (1992) with possible occurrences of unflagged accumulations on Sundays (indicated for each year with a circle) for the period 1890- 2005.

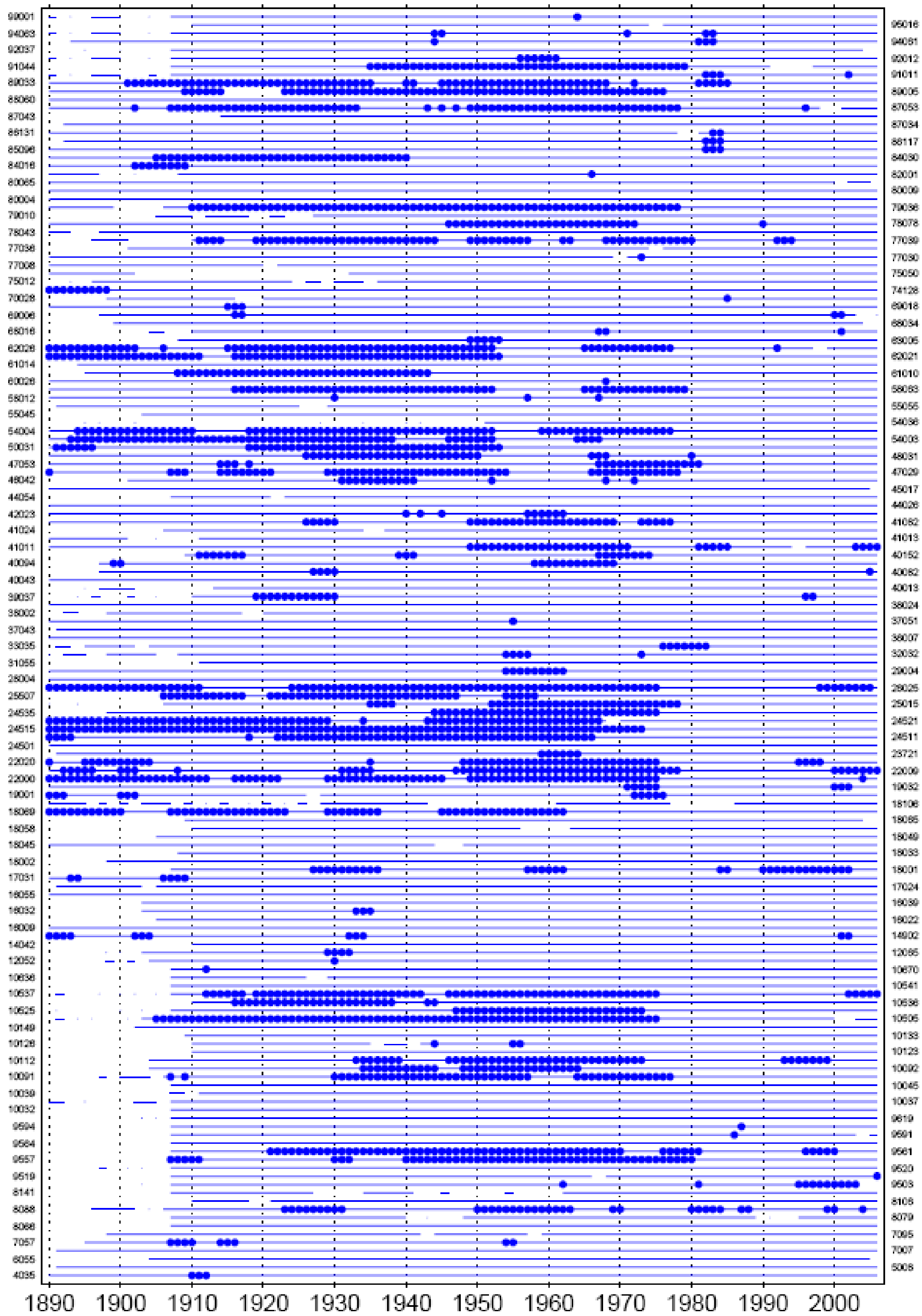


Figure 11 Occurrence of unflagged accumulations (source: Jakob et al 2008b).

1.3.1 Identifying Unflagged Accumulated Daily Rainfall

Without an existing automated system, the process of identifying unflagged accumulated daily rainfall is laborious. After extracting all the annual maxima, it would be necessary to check each annual maximum at each site and compare the rainfall with that at neighbouring sites on the date when the annual maximum rainfall was recorded as well as on the few days prior and subsequent to the event. Without an automated system it would be impractical to do this for all of the annual maxima. Typically, the highest annual maxima are the ones that have the greatest influence on the frequency curve. The presence of a very high annual maximum (which was really accumulated) in the 1-day series for a site, would push up the frequency curve particularly at the higher ARIs. However, the presence of an annual maximum rainfall which was really accumulated rainfall but of smaller magnitude would have little effect on the frequency curve because of the many other true observations of annual maxima of similar magnitude. Therefore, in this investigation, the impact of methods for adjusting for the effects of unflagged accumulated rainfall, were based on treatment of the highest of the annual maxima for each site.

To identify unflagged accumulated daily rainfall, using a station list of sites in the pilot study, the following procedure was adopted:

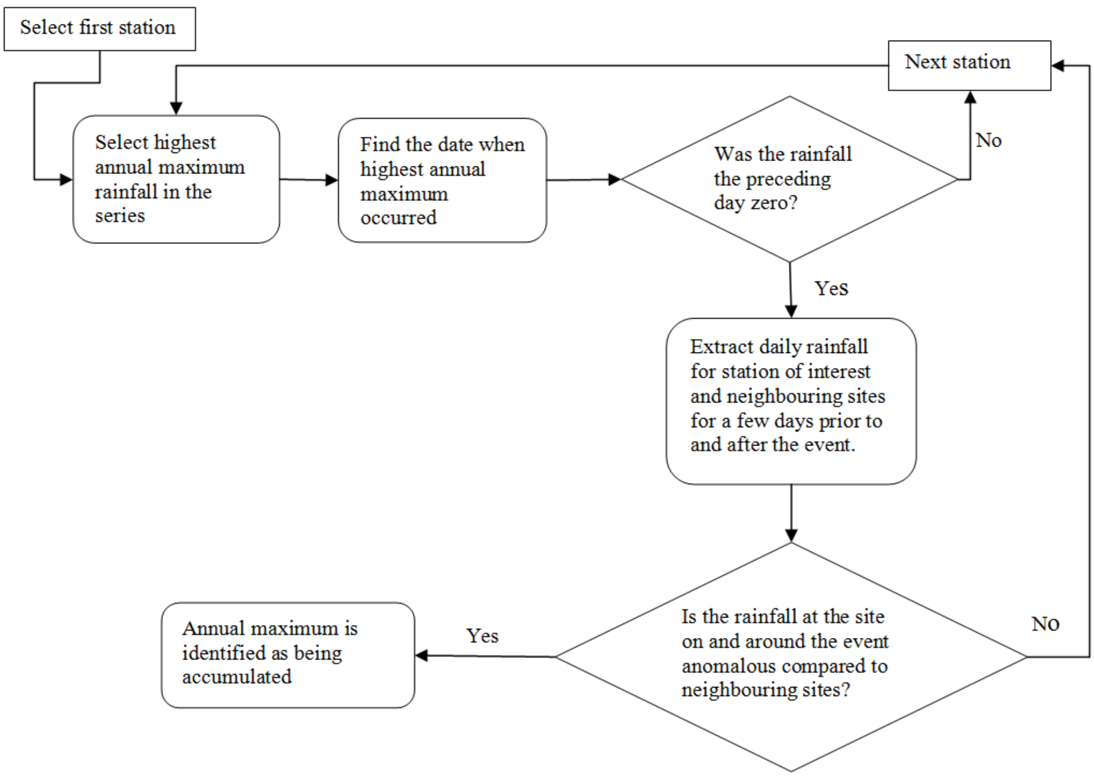


Figure 12 Identification of unflagged accumulations in the top-ranking annual maxima

Using the method shown in Figure 12, 32 stations were identified as having the highest annual maximum rainfall in the series as an accumulated total over two or more days in the 1-day dataset. The locations of the sites are shown in Figure 13. Generally, the sites identified as having the highest annual maxima as an accumulation had an even spatial distribution with many other daily sites in their vicinity.

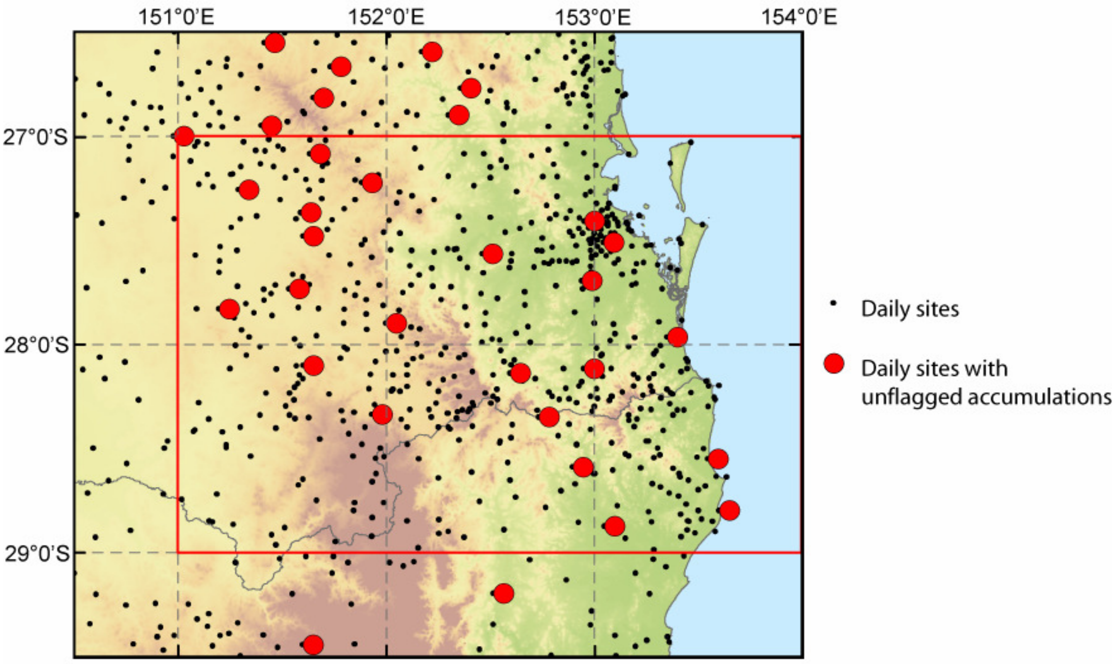


Figure 13 Locations of sites where the highest annual maximum rainfall is an unflagged accumulation

1.3.2 Methods For Adjusting For Unflagged Accumulations

Three ways for handling unflagged daily rainfall accumulations were explored:

- *ignoring* the problem, i.e. unflagged daily rainfall accumulations are included in the annual maximum daily rainfall series,
- *leaving out* any unflagged daily rainfall accumulations,
- *disaggregating* the accumulated daily rainfall by examining the rainfall at nearby stations and distributing the rainfall on successive days in percentages similar to those of nearby stations.

Ignoring the unflagged accumulations is the most straightforward approach as no adjustments are made for unflagged accumulations. This method would tend to over-estimate the design rainfall because rainfall that actually fell over two or more days would be assumed to have fallen over 1 day.

Deleting accumulations would tend to lead to under-estimation of the design rainfall because the accumulated rainfall, though not as high for 1 day, may still be high enough to be the rank one event.

Disaggregating the accumulated rainfall was explored further. It is impossible to determine the 'true' rainfall at a site where only accumulated rainfall was recorded; however, an approximate value of the rainfall can be determined by taking into account the rainfall at nearby stations. The estimate would be an approximate value, but closer to the 'true' value than values derived using either of the two previous methods (*ignore* or *delete*).

Disaggregation of accumulated rainfall involves:

1) Examination of the rainfall at *neighbouring sites* to derive an estimate of the annual maximum rainfall for that event.
2) Extraction of daily *rainfall data for the entire year* to check whether the estimate of annual maximum rainfall was exceeded on another occasion. If it was, then that higher value was adopted as the annual maximum rainfall. If it was not exceeded,

then the estimate of disaggregated rainfall was adopted as the annual maximum rainfall for the year.

3) If the new estimate of the annual maximum rainfall was lower than the *second highest annual maximum* rainfall, the latter was checked to assess whether it also was an unflagged accumulation. If not, the second highest annual maximum became the first highest annual maximum. If the second highest annual maximum rainfall was also an accumulation, it was handled in the same manner as the first highest (Step 1 above). The process was repeated until the highest ranking annual maximum rainfall was either an estimate of the annual maximum from an accumulation or it was a 'true' value, that is, it was not an accumulation.

1.3.3 Effect on Design Rainfall Estimates

The effect each method would have on regionalised and at-site design rainfall for the 24-hour duration was examined by applying the method of L-moments to derive regional and at-site parameters of the Generalised Extreme Value (GEV) distribution (which had been identified as the most appropriate for the pilot study area), in order to determine the growth curve for each region and site. The growth curves scaled by the index rainfall at each site (taken as the site median) were then used to derive regional and at-site estimates of design rainfall at standard ARIs. By comparing the estimates of design rainfall at standard ARIs using each of the three methods for adjusting for unflagged accumulations, an understanding was gained of the sensitivity of the final at-site and regional estimates to the adopted method. The comparisons were made between the estimates derived from the dataset where unflagged accumulations have been distributed and those derived using the *ignore* and *delete* approaches.

1.3.4 At-site Analysis

Figure 14 shows the differences expressed in terms of percentages as well as depths of design rainfall between at-site analysis estimates derived with unflagged accumulations included and unflagged accumulations removed compared with at-site analysis estimates derived where flagged accumulations were disaggregated. The dashed line indicates zero difference and red dots indicate the median. Percentage differences up to 60 percent (over 100 mm) at an ARI of 100 years could result from leaving unflagged accumulations unadjusted in the dataset compared with distributing the accumulated rainfall.

Table 1 and Table 2 show the number of sites affected and the magnitude of the difference in design estimates if unflagged accumulated rainfall was ignored (Table 1) and excluded (Table 2) from the dataset compared with disaggregation of the rainfall. (The Δ in the table indicates the absolute value of the percentage difference in estimates of rainfall depths.) The top-ranking annual maximum at 32 sites was identified as an unflagged accumulation. As can be seen from the tables, the differences, when compared to the disaggregated rainfall, tend to be greater in magnitude and at more sites when the unflagged accumulations are retained unadjusted in the dataset than when they were excluded.

Only 32 sites were affected by the method used to handle unflagged accumulated rainfall. However the differences at these sites could be quite large: up to 60% where unflagged accumulations were included but not adjusted and up to 32% where the top-ranking annual maxima were excluded (if due to unflagged accumulation).

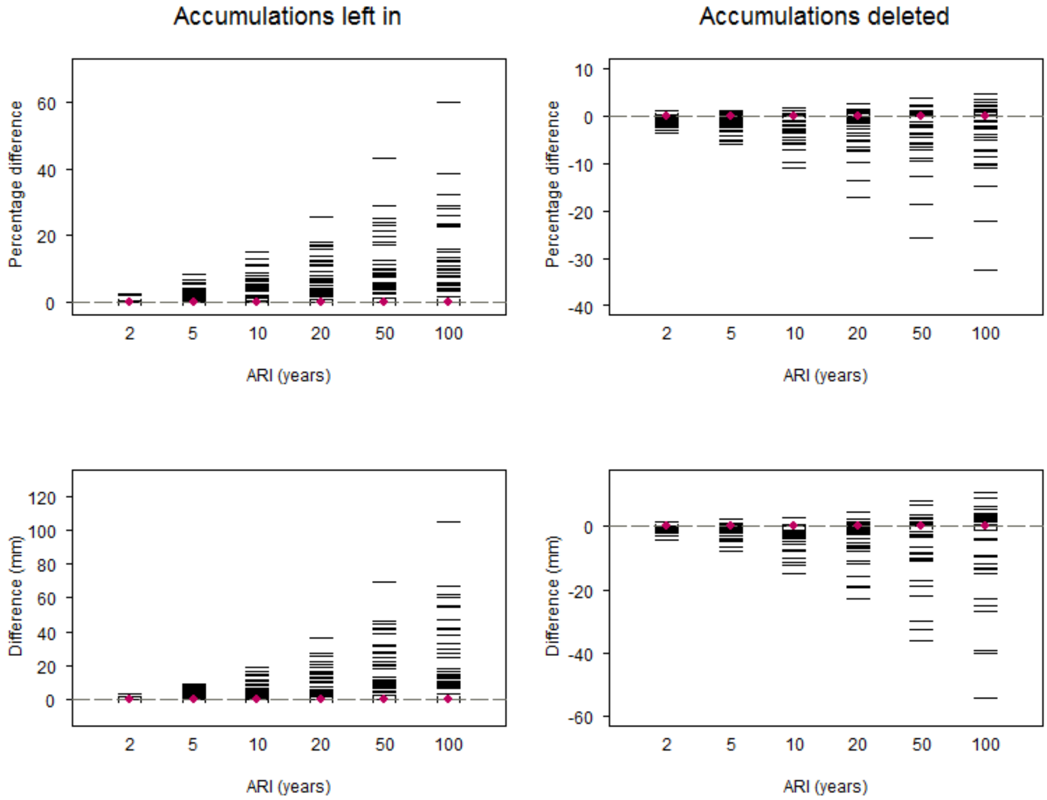


Figure 14 Effect of method of adjusting for unflagged accumulations. Percentage differences (top row) and absolute differences (bottom row) between single-site analysis estimates derived with unflagged accumulations included (left panel) and unflagged accumulations removed (right panel) compared with single-site analysis estimates derived where unflagged accumulations are disaggregated.

Table 1 Effect on at-site estimates when unflagged accumulated rainfall is included unadjusted in the top-ranking annual maxima. Δ indicates the absolute value of the percentage difference in estimates of rainfall depths.

ARI (years)	**Number of sites unaffected**	**Number of sites affected**				
		0 ≤ Δ< 3	*3 ≤ Δ< 5*	*5 ≤ Δ <10*	*10 ≤ Δ <20*	*20 ≤ Δ < 60*
2	26	6	0	0	0	0
5	0	20	7	5	0	0
10	0	12	6	9	5	0
20	0	8	5	9	9	1
50	0	3	6	10	8	5
100	0	1	4	10	8	9

Table 2 Effect on the at-site estimates when unflagged accumulated rainfall is excluded from the top-ranking annual maxima. Δ indicates the absolute value of the percentage difference in estimates of rainfall depths.

ARI (years)	Number of sites unaffected	Number of sites affected				
		0 ≤ Δ< 3	*3 ≤ Δ< 5*	*5 ≤ Δ <10*	*10 ≤ Δ < 20*	*20 ≤ Δ < 60*
2	4	26	2	0	0	0
5	0	25	5	2	0	0
10	0	23	4	4	1	0
20	0	20	4	6	2	0
50	0	18	4	7	2	1
100	0	17	4	5	4	2

1.3.5 Regional Design Rainfall Estimates

Figure 15 (top row) shows boxplots of percentage differences in regional estimates derived by *ignoring* the accumulations (method 1) and by *deleting* any accumulations (method 2) compared with *disaggregating* accumulations (method 3). Figure 15 (bottom row) shows differences expressed in terms of rainfall depth (mm).

Leaving unflagged accumulations in the dataset resulted in an over-estimation of the design rainfall (left panel) while removing them from the data resulted in an under-estimation of the design rainfall (right panel). The magnitude of the differences increased with ARI and the number of regions affected also increased with ARI.

However, when the accumulations were left in the dataset, the over-estimation of the design rainfall was never more than 3% and the worst-case underestimation when the accumulations were deleted was just over -4%. The greatest impact on the regional design rainfalls by leaving unflagged accumulations in the dataset was found to be on the region centred on site 040074. The effect for this region was that design rainfalls were 3.7 mm (2.5%) higher at an ARI of 2 years and 13 mm (2.9%) higher at an ARI of 100 years than if the unflagged accumulations were distributed. The over-estimation of the design rainfall for this region was due to an unflagged accumulated rainfall of 381 mm for site 040074 which was actually accumulated over 14 days (see Appendix A1). The actual 1-day rainfall would probably have been around 100 mm while the highest ‘true’ annual maximum rainfall for this site was 330 mm. The dashed line indicates zero difference and the red dots indicate the median.

If unflagged accumulations were removed from the dataset, the greatest difference in regional design rainfall estimates would be an under-estimation of approximately 4.1 mm (3.5%) at an ARI of 2 years and 13 mm (4.2%) at an ARI of 100 years.

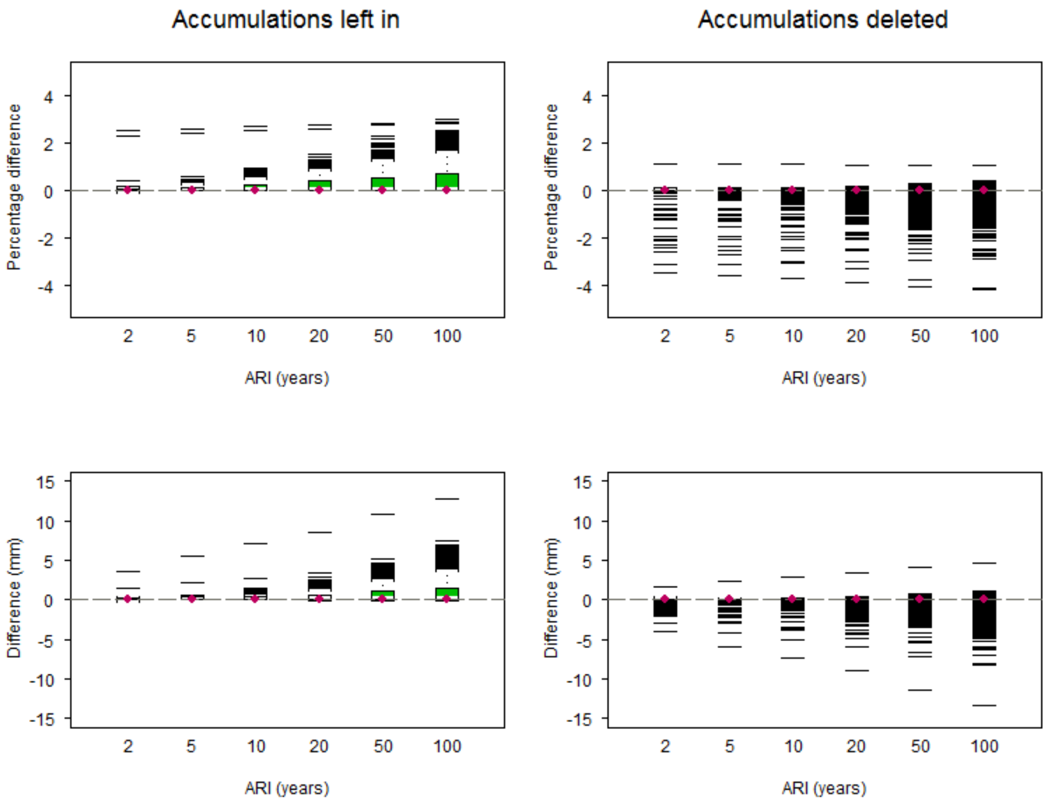


Figure 15 Effect on regional estimates of method of adjusting for unflagged data. Percentage differences (top row) and absolute differences (bottom row) between single-site analysis estimates derived with unflagged accumulations included (left panel) and unflagged accumulations removed (right panel) compared with single-site analysis estimates derived where unflagged accumulations are disaggregated.

Table 3 and Table 4 show the number of regions affected, depending on the method used to treat unflagged accumulated rainfall. Where the percentage difference is zero, regional design rainfall estimates are not affected by unflagged accumulations. The highest absolute percentage difference compared with disaggregating the rainfall was slightly higher if the accumulated annual maxima were excluded rather than included. However, the highest absolute percentage difference for the excluded dataset only exceeds 4% for two sites at an ARI of 100 years.

Approximately half of the regions considered were affected by the 32 stations with unflagged accumulations, with the number being slightly higher when the unflagged accumulations were included rather than excluded. The magnitude of the differences was also slightly higher if unflagged accumulated rainfall was included unadjusted rather than removed from the dataset. Nevertheless, most of the differences were very small, generally less than 1% for both methods. For an ARI of 100, 24 sites had a maximum difference in magnitude between 2 and 3% (if accumulations were included unadjusted). Of the total 390 sites affected for an ARI of 100 years, for 264 sites differences were less than 1%.

Table 3 Effect on regional estimates when unflagged accumulated rainfall is included unadjusted in the top-ranking annual maxima (for the pilot study area). Δ stands for the absolute value of percentage difference.

ARI (years)	Regions **centred on sites** having accumulated rainfall (total = 32)			Other regions **affected by sites** having accumulated rainfall (715)		
	Number of regions			Number of regions		
	Δ = 0	*0 < Δ < 2*	*2 ≤ Δ < 3*	*Δ = 0*	*0 < Δ < 2*	*2 ≤ Δ < 3*
2	27	3	2	671	44	0
5	0	30	2	357	358	0
10	0	30	2	357	358	0
20	0	30	2	357	358	0
50	0	30	2	357	356	2
100	0	25	7	357	341	17

Table 4 Effect on regional estimates when unflagged accumulated rainfall is excluded from the top-ranking annual maxima (for the pilot study area). Δ stands for the absolute value of percentage difference.

ARI (years)	Regions **centred on sites** having accumulated rainfall (total = 32)			Other regions **affected by sites** having accumulated rainfall (715)		
	Number of regions			Number of regions		
	Δ = 0	*0 < Δ < 2*	*2 ≤ Δ < 5*	*Δ = 0*	*0 < Δ < 2*	*2 ≤ Δ ≤ 4.5*
2	4	21	7	715	0	0
5	0	25	7	392	323	0
10	0	25	7	392	323	0
20	0	24	8	392	323	0
50	0	23	9	392	321	2
100	0	23	9	392	318	5

1.3.6 Conclusion

For the pilot study, regional design rainfall estimates could be over-estimated by up to 3% (at a duration of 24 hours) at about half of the total number of regions, if data were included unadjusted from 32 sites identified as having the highest annual maximum rainfall in the series as an accumulation over two days or more. Although only 32 sites out of a total of 747 sites were identified as having a problem with the highest ranking annual maximum rainfall, the effect of the inclusion of unadjusted data from these sites was observed at approximately half the total number of sites, due to the effects of regionalisation. However, the effect was less than 1% for an ARI of 100 years (where the greatest difference was manifest) for a total of 621 sites.

Therefore, the effect of unflagged accumulated annual maxima in the highest ranking rainfalls in the dataset for a duration of 24 hours was very small, and would be within the uncertainties in the method which arise from assumptions that may not necessarily hold over all regions, for example, the assumption of the same frequency distribution over the pilot area and the assumption of homogeneity. It would be expected that the effect of lower ranking unflagged accumulated annual maximum rainfalls would be negligible.

If the highest ranking annual maxima identified as being accumulated over a longer period were removed from the dataset, the design rainfall could be under-estimated by up to 4.5%. However, the effect was less than 1% at 687 sites when compared with estimates obtained by apportioning the accumulated annual maxima over two or more days using data at neighbouring sites.

For at-site analysis, fewer sites were affected (32 for the pilot study). However, the differences at these sites could be quite significant (up to 60% at an ARI of 100 years) if accumulated rainfall was included unadjusted. For single-site analysis, the issue of unflagged accumulated rainfall would need to be addressed.

If it were decided that the effect of unflagged accumulation on design rainfall estimates was significant enough to warrant an attempt to disaggregate at least unflagged accumulations for the highest ranking annual maxima, there would be value in investigating whether existing grids of daily rainfall might be useful in deriving a first estimate of disaggregated values, for example, SILO (Jeffrey et al, 2001) and BAWAP (Jones et al, 2007).

Alternatively, the Fortran program (daymaxplot) developed as part of the development of the CRCFORGE methods (Nadarajah Nandakumar, pers. comm.) which allows the user to identify and disaggregate accumulated values in the annual maximum series (based on data at up to 10 nearby stations) could be adopted.

1.4 Quality Controlling Data at Subdaily Durations

At subdaily durations, a variety of data sources were used, including data in HYDSYS format, data from ALERT gauges and also pluviograph information from the ADAM database. Quality controlling of the sub-daily data from each of the sources was undertaken as described in the following sections.

1.4.1 Data in HYDSYS Format

In addition to data held by the Bureau, data collected by external organisations was also sought. Data from the former NSW Department of Infrastructure, Planning and Natural Resources (DIPNR) and the former Queensland Department of Natural Resources and Mines (DNRM) were included in the pilot study.

The term ‘HYDSYS format’ refers to a specific format of electronic files used with HYDSTRA software (formerly HYDSYS), which facilitates the time series data management widely used by institutions working with hydrologic data.

Data were received in HYDSYS format (as .csv files), including the required information about the quality codes used. Figure 16 shows the quality codes used with the NSW DIPNR data. Data were loaded into HYDSYS and annual maxima were extracted for durations from 1 min to 72 hours (using HYRFREQ). Ranked annual maxima were plotted for each of the 14 durations.

Figure 17 shows an example of the plots of ranked annual maxima that were used in quality controlling the subdaily data. The highest values for the years 1994 and 1995 for the 1-minute duration were identified as suspect. Where the values were identified as suspect, the following investigations were considered:

- The synoptic charts were checked with a view to deciding whether significant rainfall might have occurred on the day in question.
- For shorter durations: the traces in HYDSYS were inspected. Figure 18 shows an example for station 54001 for 1995 where a suspect value for the 1-minute duration was identified and subsequently rejected. For longer durations: comparisons with nearby daily sites were undertaken.

Department of
Infrastructure, Planning and Natural Resources

Resource Information Unit
Armidale

Quality Codes

All Times series data produced by the department has an associated quality.
Allocation of Quality Codes to Time Series Data is based on instrument performance, reliability, calibration status and stream conditions.

The following table of Codes and characters is provided as a guide to the data you have been supplied.

If you have any questions regarding Quality Codes please contact the Resource Information Unit.

Quality	Print Character	Description
1		Highly accurate continuous records
2		Good continuous record.
3		good record processed +/- 10mm, within 30minutes
4		good record where number of gaugings >=5, 95% within 10% curve
5	D	good spot samples (daily read) +/-15minutes
6		Good gauging to AS3778
7		Good Gauging
8		Good record processed prior to 1/11/96 and previously coded1
9	B	Water level below sensor and no flow/dry stream
10	F	Fair measured data
11	F	fair measured data.
12	F	Acceptable Gauging
14	F	Fair record where 95% of flow gaugings within 20% of discharge
15	H	spot samples of fair quality, +/- 1 hour
16	J	Level estimated to +/-20mm, time unknown
17	F	Reasonable Gauging
22	F	Fair data, 1:2 exposure +/-10% but time errors caused by fog, dew etc
24	W	Wet Day in Accumulated Rainfall
25	A	Accumulated Rainfall
26	K	Fair data, 1:2 exposure +/-10%,cum.total over unread period > 24 hours
27	L	Data limited by client spec. fulfils requirement of qc10
28	F	fair record processed prior to 1/11/96, previously qc10.
30	P	Poor measured data, time +/- 2 hours.

120 Jessie St Armidale NSW 2350
Ph : 02 6772 2441 Fax : 02 6772 7862
Visit Our Website http://www.dipnr.nsw.gov.au

Figure 16 Quality codes used with HYDSYS to extract annual maxima for DIPNR data.

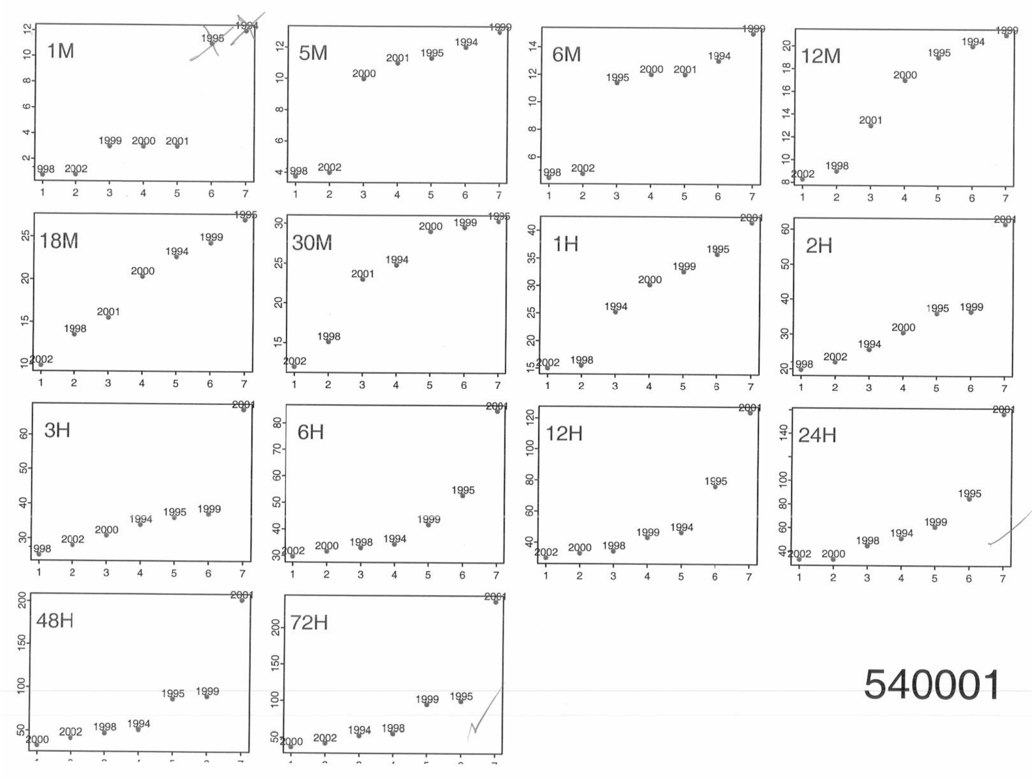


Figure 17 Example of plots of ranked annual maxima for site 54001. x-axis: running number, y-axis: rainfall depth in mm, labels above points indicate calendar year. (Highest values for the years 1994 and 1995 at the 1-min duration were identified as suspect.)

1 Min 1995 540001 1995 11mm out of the blue.

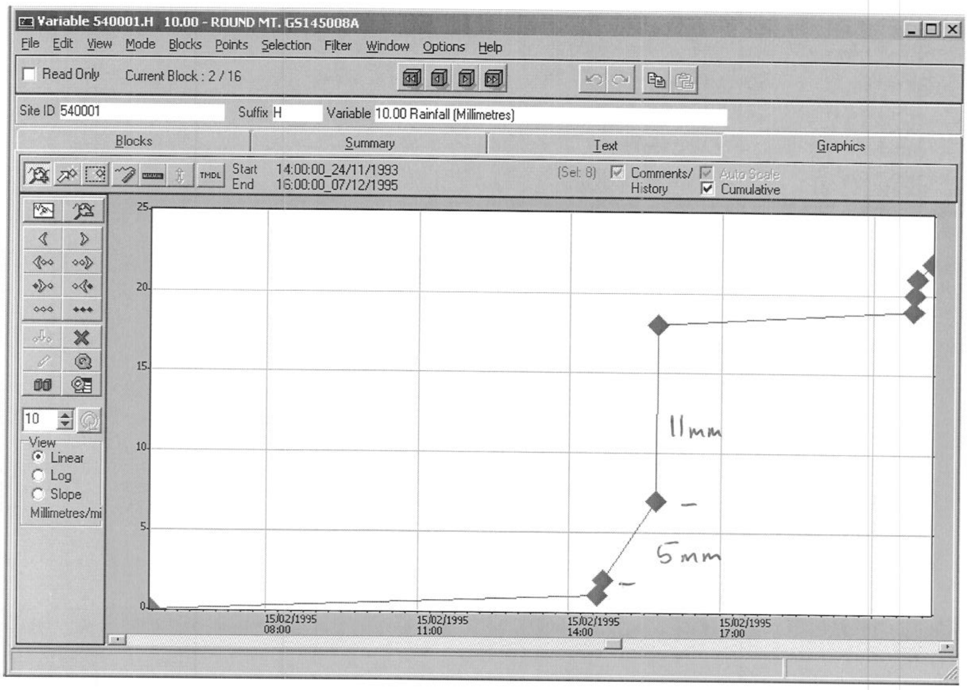


Figure 18 Example of using HYDSYS to investigate a suspect value for the 1-min duration at site 54001 (for 1995).

1.4.2 Alert Data

Similar to the data received from external organisations, ALERT data had been quality controlled during the earlier part of the pilot study. For completeness, the process is described briefly here:

It did not appear feasible (or necessary) to quality control all the available ALERT data. Instead the pilot study area was divided into four areas (for simplicity numbered 1 to 4). In each of these areas, gaps were identified which were not covered by other stations with subdaily observations. Appropriate gauges, with at least 8 years of data, were selected and an Splus function was used to plot the ALERT traces, once for all years on one page and then again one year per page. Figure 19 shows the trace for gauge 541044 for the year 2000.

For notable events, data were compared against totals from nearby daily sites. Being aware of known problems with ALERT data, helps in quality controlling data in two ways: firstly, it makes it easier to spot certain problems and secondly, it may help in deciding on possible adjustments. Some of the known problems include (Terry Malone, pers. comm.):

- Data are received as signals for each 1 mm tip of the bucket. Sometimes there is interference that causes a signal not to be received at the 1 mm interval. Therefore, the next record is the rainfall that has accumulated since the time the last signal was received. That is, usually increments are 1 mm but at times the increments are larger.

- An 8 bit signal is used and therefore a gauge will reach 2047 mm before an automatic resetting back to zero occurs. However, the gauge can be reset manually which often occurs when the equipment is being tested. It is also possible that some testing is undertaken without resetting the gauge to zero.

- Because the system is electronic, there may be times when the electronic equipment may be affected by water (i.e. a leak) which would usually occur during times of heavy rainfall. This may cause the signalling system to malfunction (i.e. large oscillations notable in the trace).

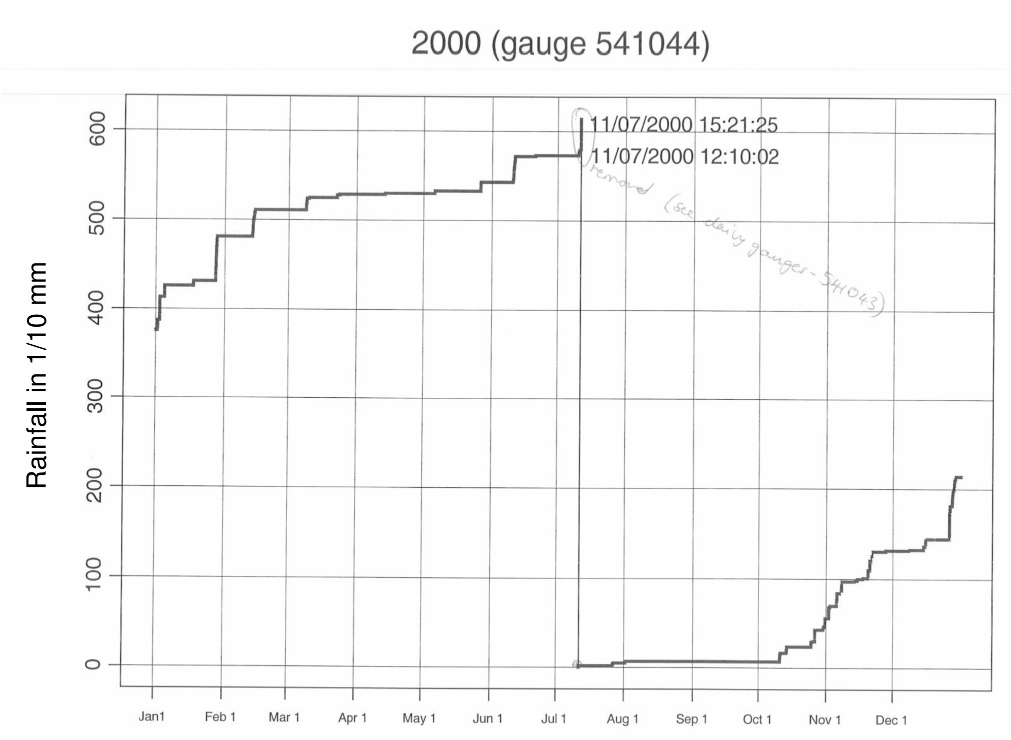


Figure 19 Example of trace plotted to quality control ALERT data. Handwritten comment reads: 'removed (see daily gauges - 541043)'.

1.4.3 Pluviograph Data

Tipping bucket raingauges (TBRGs) were installed at official Bureau stations from about 1997 onwards, replacing the long-serving Dines Pluviographs. The Dines record was pen-on-paper which required subsequent manual digitising, whereas the TBRG output is already in digital form which should facilitate the rapid loading into the database. However, it has been found that manual adjustment is required for some records. Staff shortages in the National Climate Centre (NCC) have resulted in a backlog of data that require manual adjustment and, as a consequence, the TRBG database is only complete to the end of 2006.

In order to quality control the pluviograph data, scripts had previously been written to plot the extracted annual maxima at each station for a number of durations. Visual inspection of these plots allowed checking for apparent inconsistencies across durations. Since additional annual maxima had been extracted and extra stations had been added since the first part of the pilot study, additional quality controlling was undertaken.

As part of the quality controlling for the first part of the pilot study, AM series for all twelve durations were plotted together, to allow identification of spurious values. These checks were repeated for the additional sites and suspect values were investigated. Figure 20 shows the plot for station 41056 for durations from 6 minutes to 72 hours. Gaps in the lines indicate years which had been rejected for the extraction of annual maxima.

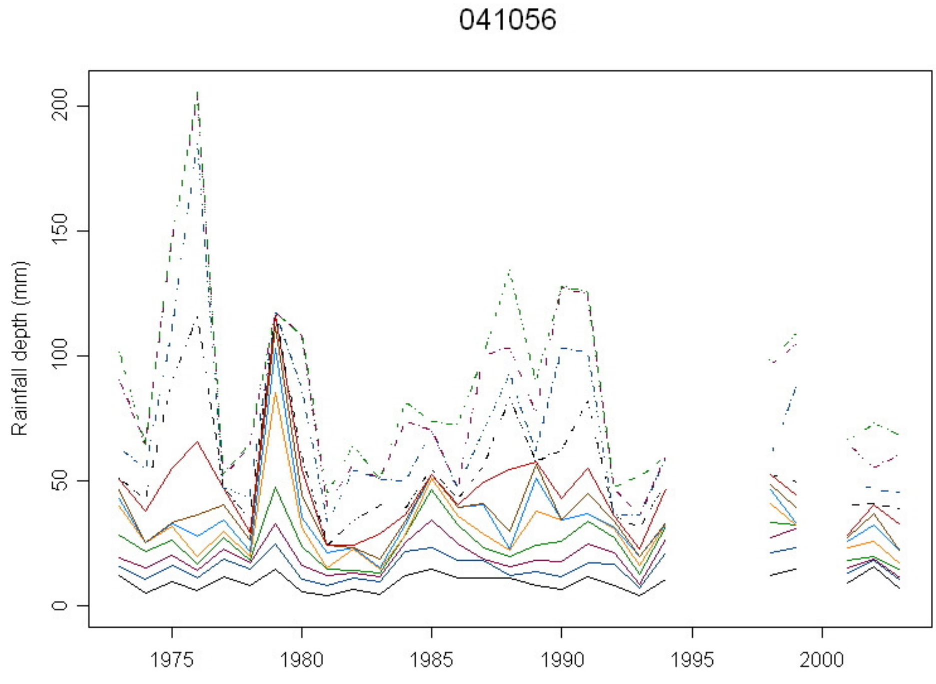


Figure 20 Example of plots used in screening annual maxima for durations from 6 min to 72 hours for station 41056). Different durations are marked by different colours/line types.

Three problems were identified during the quality controlling of pluviograph data:

- The design rainfall estimate for the 1-h duration at an ARI of 100 years for 040062 (Crohamhurst) appeared suspiciously high (227 mm) when compared to nearby values of about 100 mm. On examining output from Fortran programs used in the extraction of the annual maxima (PLUVANAL and PLUVINT), it was found that the record for 1961 contained a number of errors. There was an unflagged accumulated value of 223.8 mm compared to the next highest value of 82.6 mm and compared to the highest hourly value

observed in Australia of 230 mm at Florence (Queensland) in December 1920. In light of this, the annual maxima extracted for 1961 were removed across all durations and the number of years reduced from 23 to 22.

- For another station (040659) the annual maximum for 1979 had been excluded because the year contained insufficient data (minimum 10 months with minimum of 75% each). However, comparisons with a nearby site (40312) suggested this value should be included. Using the 'add back in' procedure this annual maximum was included in the record.

- In the course of analyses directed at deriving a method to estimate 5-min rainfall depths from 6-min rainfall depths, suspect annual maxima were also identified for one station outside the pilot study area (Inverell Research Centre, station 56018). Discrepancies were found for 5-min and 6-min annual maxima in 1949 and 1962. An unsuccessful attempt was undertaken to recover the original charts to try and reconcile the estimates and therefore the site was excluded from further analyses.

Quality controlling of pluviograph data focussed mainly on the rainfall data themselves (that is the extracted series of annual maxima) and the accompanying metadata. Investigations of grids for the median rainfall at the 1-h duration exposed how very dependent the grid is on data at one station (see section 4.2.2).

Errors introduced at the digitising stage

As part of the visual inspection of extracted annual maxima, it was considered prudent to focus on the shortest durations in order to identify data which might have been digitised incorrectly. As one of the initial checks for another study, the annual maxima for Sydney Observatory Hill (66062) were extracted and plotted together with the percentage of data missing in a given year (Figure 21). These plots revealed two very high 5-min rainfall depths. One in April 1927 (23 mm) and another in November 1984. The value in April 1927 appeared suspicious and paper charts were inspected (Figure 22). It is likely that the paper chart got soaked during the event and the needle was stuck to the paper. Paper charts are manually digitised, introducing a degree of subjectivity. The event in question was redigitised by three experienced staff in the National Climate Centre. The most likely value is considerably lower than the originally digitised value (about 8 mm in 5 minutes rather than 23 mm.)

While similar problems were not found with data used in the pilot study, mistakes at the digitising stage can potentially introduce very large errors and it might be useful to take this into account for the IFD revision.

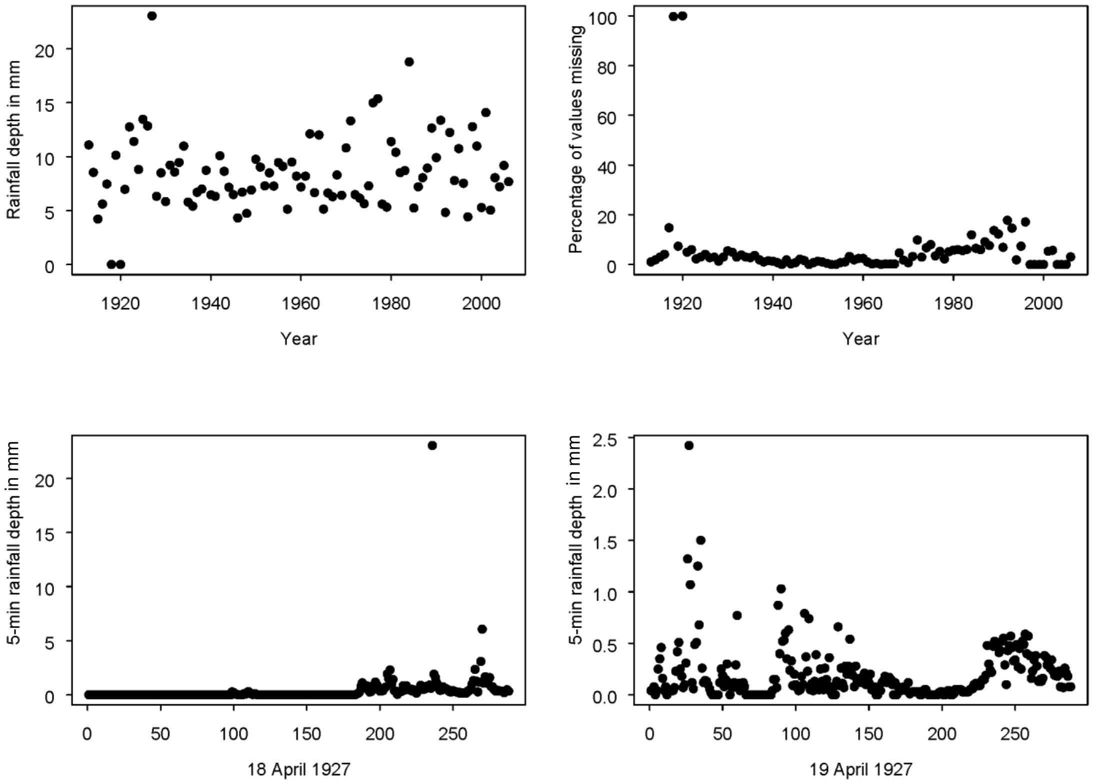


Figure 21 5-min rainfall depth for Sydney Observatory Hill (66062). Top left: annual maximum series, top right: percentage of missing values in a year. Bottom row: 5-min rainfall depths for 18 and 19 April 1927 (note different scales on y-axis for the two days).

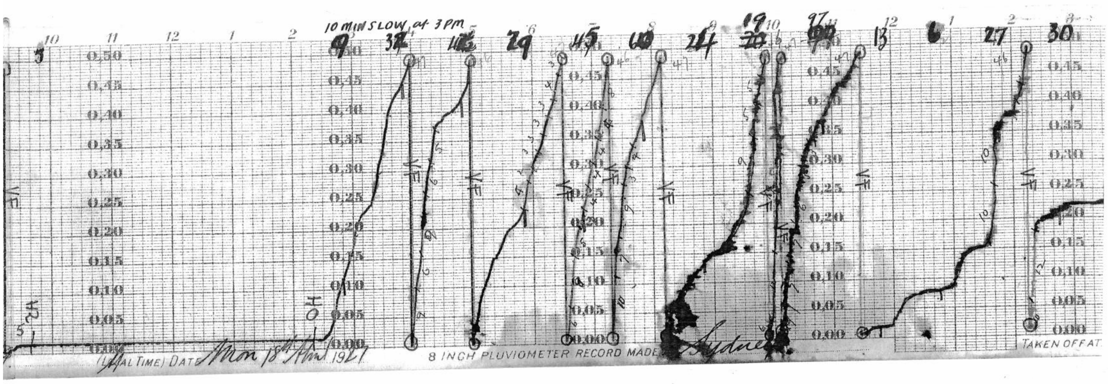


Figure 22 Paper chart for the rainfall event 18/19 April 1927 (Sydney Observatory Hill, 66062)

Additional quality checks were undertaken as part of the analyses at later stages of the pilot study, in particular discordancy tests and inspection of sites with large residuals at the mapping stage.

Discordancy tests (as part of regionalisation)

L-moments (L-CV, L-skewness and L-kurtosis) can be used to summarise the statistical properties of series of annual maxima. These L-moments are used in a number of ways - to

identify suitable regions using a ‘homogeneity measure’ and to derive regional L-moments (for regional growth curves).

Based on L-moments it is also possible to identify sites where the annual maxima exhibit different statistical properties compared to other sites (in a group or region). Such assessment can be done by plotting L-moment diagrams (for visual inspection, Figure 23). Objective tests have also been developed (Hosking and Wallis, 1997). The discordancy measure for site *i* in a group of *n* sites is compared against a critical value. Where D_i exceeds the critical value, the site is flagged as discordant. Sites flagged as discordant were inspected to decide whether a site should remain in the dataset or be excluded from further analyses.

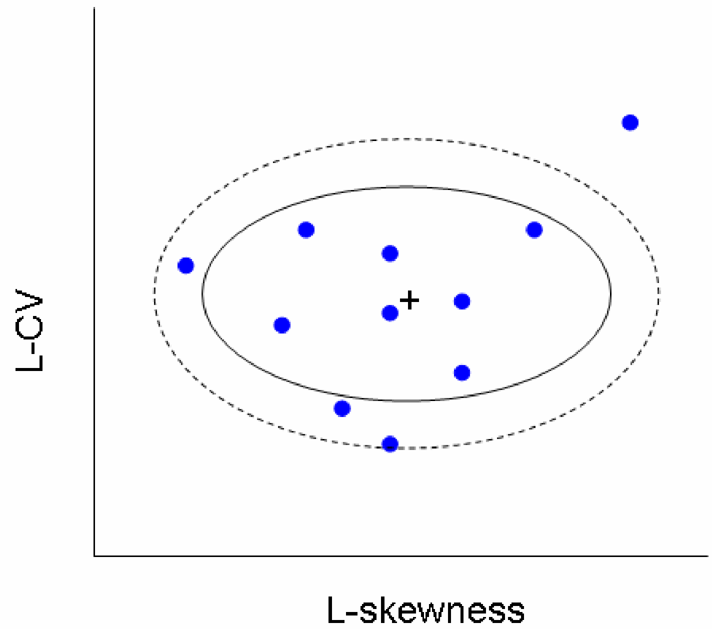


Figure 23 Definition sketch for discordancy (after Hosking and Wallis, 1997)

Largest residuals (as part of mapping)

Mapping was undertaken using the technique of thin-plate spline smoothing (ANUSPLIN, Hutchinson, 2007). As part of the mapping process lists with ranked residuals were available. Inspecting these lists led to identification of a small number of suspect sites, see HRS10 section 6.4. Typically the geographical location was found to be incorrect.

1.5 Metadata

Using station data to derive IFD estimates for any given location (at defined resolution) in Australia requires an implicit assumption: station data are representative and not unduly modified by local conditions.

When data are ingested into databases, a system of quality checks is usually developed to ensure data are internally consistent (for a location and by ‘buddy’ checking). Metadata can give useful information about the station itself (sketch, changes in instrumentation etc.). Some of this information is available from a dedicated metadata database - SitesDB; alternatively station files were used. Using these information sources, the recorded elevation and geographic location of stations were checked.

1.5.1 Elevation

Excel spreadsheets containing information such as station location (latitude, longitude and elevation) had previously been compiled. These spreadsheets were updated to take into

account increased record lengths and additional stations that were included after the updating of records in April 2007. A number of cases were identified where information from different sources showed inconsistencies. For example, the elevation from SitesDB/ADAM for some stations showed differences of more than 10% when compared with elevations derived from a digital elevation model (DEM) using available latitude and longitude data. These cases were followed up and a decision was made about the validity of the station elevation. Table 5 shows details). The last of six stations with suspect elevation data listed in Table 5, station 41361, was not included in the final data set as the record length was shorter than the minimum of 8 years.

Table 5 Checking suspect elevations for additional pluviograph sites

Station (name/number)	Lat.	Lon.	Elev.	DEM	Diff. in %	Comment
COOLANGATTA BOWLS COMP (40052)	-28.183	153.533	6	83	93%	location doubtful, likely location -28.181 and 153.534, elevation at likely location 53 m, use DEM?
ESK DALE WEST (40076)	-27.153	152.168	501	376	-33%	location could be correct, elevation from Google Earth 382 m, stick with DEM
IPSWICH (40101)	-27.612	152.761	39	4	-875%	location might be correct, near road and river, Height from Google Earth is 25 m, much lower on other side of the river (8 m), stick with ELEV (39 m)
MT STANLEY FORESTRY (40148)	-26.600	152.167	350	310	-13%	Google Earth has 349 m for this location - use ELEV (350 m)
THORNTON BVRT (40202)	-27.821	152.381	200	159	-26%	Google Earth has 189 m, stick with ELEV (200 m)
PITTSWORTH DPI (41361)	-27.717	151.633	608	511	-19%	Google Earth has 521 m for this location - use DEM (511 m)

1.5.2 Geographic Location of Stations

Station locations can be derived from ADAM, however, for some stations, the accuracy of the latitude and longitude was very low with co-ordinates being quoted to only one decimal place.

In addition, the latitude/longitude co-ordinates in the ADAM database for some of the daily stations used in the pilot study were different to the co-ordinates in ADAM when the pilot study was commenced. For a number of stations slight differences between the old and new co-ordinates resulted from the use of three decimal places in the old latitude/longitudes compared with the four decimal places adopted for the new latitudes/longitudes. Such differences are due to rounding and are trivial. Other differences not due to rounding were

evident for 59 stations out of 747. Many of these differences were, again, quite small, but for 17 stations there was a difference of .01 latitude or longitude degrees (about 1.7 km) and for 4 stations there was a difference of .05 latitude or longitude degrees (about 6 km). These shifts could be significant as they could affect the stations which were/were not included in each circle for regionalisation.

The station 58087 was examined because it showed the greatest shift in its latitude/longitude co-ordinates. The latitude/longitude changes resulted in a station shift of 10 km to the southwest from its previously given position. The effect this had on the group of stations included for regionalisation is shown in Figure 24. This shows quite a different grouping of stations based on station 58087's new location compared with the grouping used with its original given location. Five stations included previously for regionalised estimates for 58087 were not included and seven new stations were included.

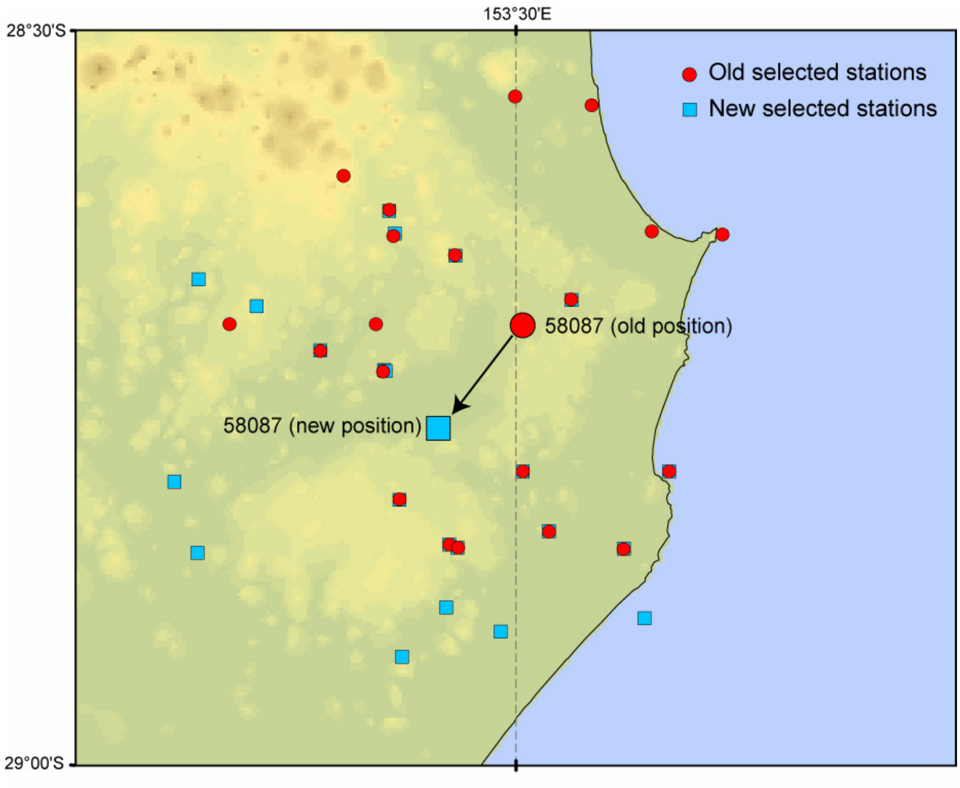


Figure 24 Old and new location for station 58087 and set of sites in region.

The process of regionalisation was carried out with the new latitude/longitude co-ordinates to assess the effects of these latitude/longitude shifts on the design rainfall estimates. The procedure used was the same as before with the same 747 daily sites being used but with their updated latitude/longitude co-ordinates.

The regionalised estimates using these updated latitude/longitude co-ordinates differed from the previous estimates in 111 out of 747 regions. Figure 25 shows the effect of the change in station location on the regional estimates. The left panel shows the difference in rainfall depth while the right panel presents the percentage difference. For 90% of sites, differences are less than 7 mm, however, for 6 sites the change in the site locations affected the ARI 100 years, 24-hour regionalised estimates by more than 10 mm with maximum differences of up to 20 mm (see Figure 25, left panel). Percentage differences of up to 5% were seen but were typically much lower with the differences for 90% of sites not exceeding 2.3%. The regional estimates for the region centred on the site with the largest change in location (58087) showed large differences in both the absolute difference (19 mm) and the percentage difference (5%).

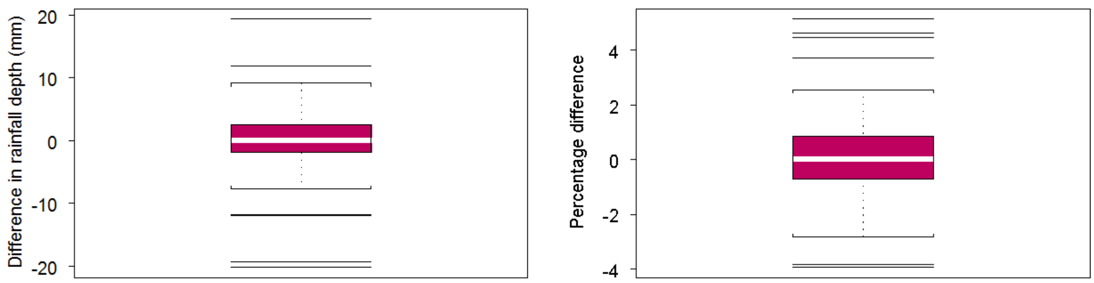


Figure 25 Effect of changes in station location on regional estimates of 24-h ARI 100 year rainfall depths.

A number of sites with large differences in estimates (based on the new and old locations) lie in the same geographic region to the southeast of the pilot study area and as a result the gridded estimates would also be affected. It was decided to re-run the pilot study area analysis based on the new latitude/longitude co-ordinates for all durations. The pluviograph sites were also checked but no significant aberrations were found between the old and the new locations.

There is a possibility that other stations' latitude/longitude co-ordinates might be changed at some future point in time. It is noted that there are some other stations that have suspect co-ordinates with zeros in the second, third and fourth decimal places in both latitude/longitude pairs. However, it is not known whether any of these sites will have their location updated. The effect of changes in site locations due to more accurate latitude/longitude positioning will need to be taken into account in any current and future hydrometeorological studies.

2 REGIONALISATION FOR SUBDAILY DURATIONS

HRS10 describes the fundamental regionalisation approach used in the pilot study. Based on the annual maximum series at a site, L-moments and medians (of the annual maximum series) were calculated. Regional estimates of L-moments were derived as weighted averages of at-site L-moments. Design rainfall estimates at a location were derived by multiplying the regional growth curve (based on regional L-moments) by the at-site median.

Two different approaches to defining regions were explored: a cluster approach and a circles approach (HRS10, section 9 ‘Regionalisation’). An alternative approach based on ellipses was tested to address the fact that the relief has a strong impact on rainfall extremes (although it might have been computationally simpler to use rectangles instead).

Regionalisation at subdaily durations was based on the methods developed for the daily durations; however the procedure is more complex. At subdaily durations, fewer sites and typically shorter records than for daily durations were available (see Figure 2 for a comparison) and therefore it was necessary to make use of the information from longer durations. Methods were developed to derive ‘predictions’ of index rainfall, L-CV and L-skewness - these are described in sections 2.1, 2.2 and 2.3.

In deriving regional estimates of L-CV and L-skewness, both ‘predictions’ (based on the 24-hour duration) and ‘direct estimates’ (derived directly from annual maximum series at subdaily durations) were used. A suitable weighting scheme was devised to combine these two sets of observations as discussed in section 2.4.

2.1 Derivation of Predictions

Three different methods for inferring statistics for sub-daily durations from information at longer durations were tested, including a Principal Component Analysis (PCA) similar to the one used for the development of ARR87; a simple scaling approach (Menabde et al, 1999); and an approach based on *Partial Least Squares Regression* (PLSR). 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 (Geladi & Kowalski, 1986). The technique is related to both canonical correlation and factor analysis. A set of orthogonal factors (referred to as ‘latent variables’) is extracted from the predictors. One of the particular strengths of this approach is that it can deal with a large number of independent/dependent variables and data sets where correlation might otherwise be an issue. *Predictive Error Sum of Squares* (PRESS) values give a good indication of the performance of a model. While the maximum number of factors is prescribed by the number of independent variables (predictors), the optimum number of factors was identified using the PRESS values.

The prediction of different statistics might be considered including: average recurrence intervals (as for ARR87), distribution parameters (location, scale and shape for three-parameter distributions) or L-moments and index rainfall (Alila, 1999; Wallis et al, 2007). For this study it was decided to use a Partial Least Squares Regression approach to predict the index rainfall (median of annual maxima) and the first two L-moment ratios (in the following referred to simply as ‘L-moments’): L-CV (the coefficient of L-variation) and L-skewness.

2.2 Prediction of Index Rainfall at Sub-Daily Durations

Two decisions have to be made in developing a PLSR model to predict index rainfalls for sub-daily durations: the identification of an optimum set of predictors and a suitable set of stations. R^2 was used in the choice of independent variables; candidates included latitude, longitude, elevation, slope, aspect and distance from coast as well as index rainfall at durations of 24, 48 and 72 hours. For the final set latitude, longitude and distance from coast together with the index rainfall at 24 hours duration were selected.

Two different sets of stations were trialed for deriving the model: all pluviographs with at least 30 years of data, regardless of their geographical location within Australia ('entire set') and a smaller set of pluviographs with the same minimum record length, located in and around the pilot study area ('reduced set', see Figure 26). While the second of these data sets is smaller than the first, a model developed using these data might be expected to be more suitable for predicting estimates in the pilot area.

In the following, 'predictions' refers to estimates of index rainfall derived using the PLSR approach, whereas 'direct estimates' refers to estimates derived from annual maximum series directly (using pluviograph data). The suitability of these two sets of stations was assessed by comparing predictions against what might be deemed 'best estimates': direct estimates from stations in the pilot study area and estimates from ARR87 (at a resolution of 0.025 degrees of latitude and longitude). Comparisons were made for three durations (1, 3 and 12 hours) and at locations where all four different estimates were available (ARR87, direct estimates, predictions from *entire set* and *reduced set*).

Note that *index rainfall* for the pilot study was derived as the median of the annual maximum series at a site. Therefore, index rainfall corresponds to a design rainfall estimate for an ARI of 2 years. For comparison with the ARR87 values, estimates were adjusted from an annual maximum scale to a partial duration series scale using the adjustment factors described in Jakob et al (2005a).

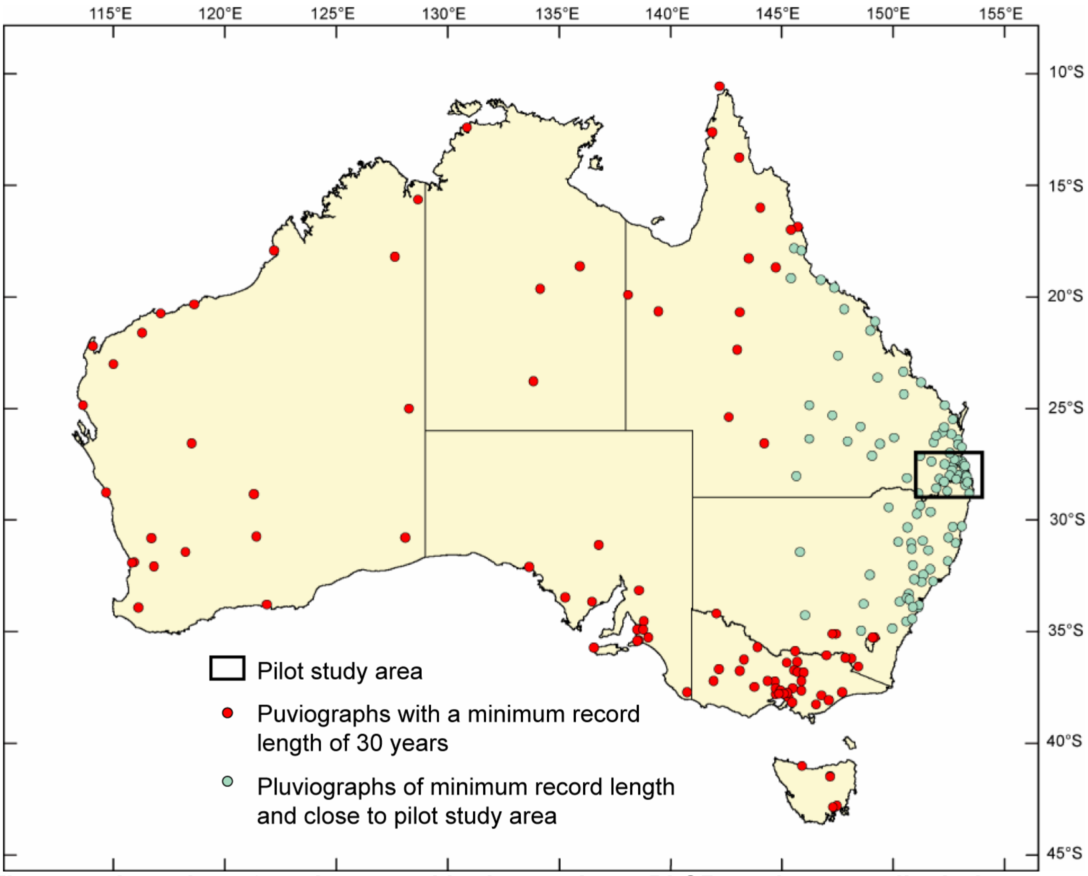


Figure 26 Location of stations used in developing a PLSR model to predict index rainfall at sub-daily durations.

Figure 27 shows boxplots of these four different estimates for three durations (1, 3 and 12 hours). The box-whisker plots indicate the range of index rainfall estimates for a given set. The horizontal line dividing each of the boxes is the median. The box itself indicates the first and third quartile. The distance between these quartiles is also referred to as the 'interquartile range' (IQR). The horizontal short lines at the ends of the vertical dashed lines are plotted at 1.5 IQR. Values outside this range are considered outliers and are plotted as open circles.

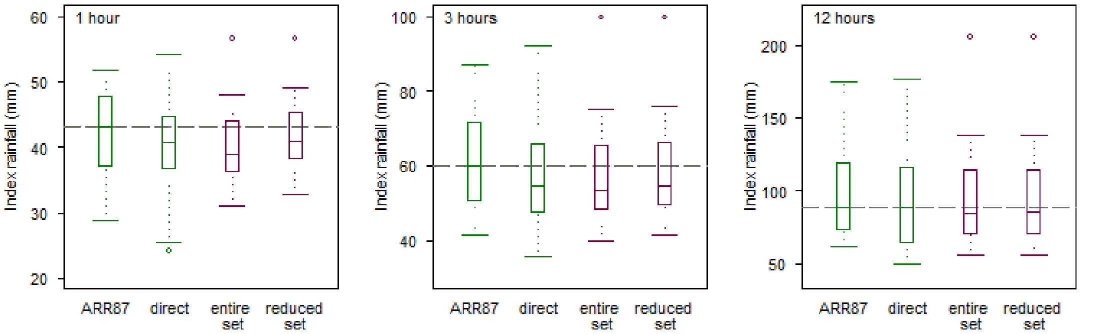


Figure 27 Boxplots of index rainfall for locations in pilot study area where four different estimates were available

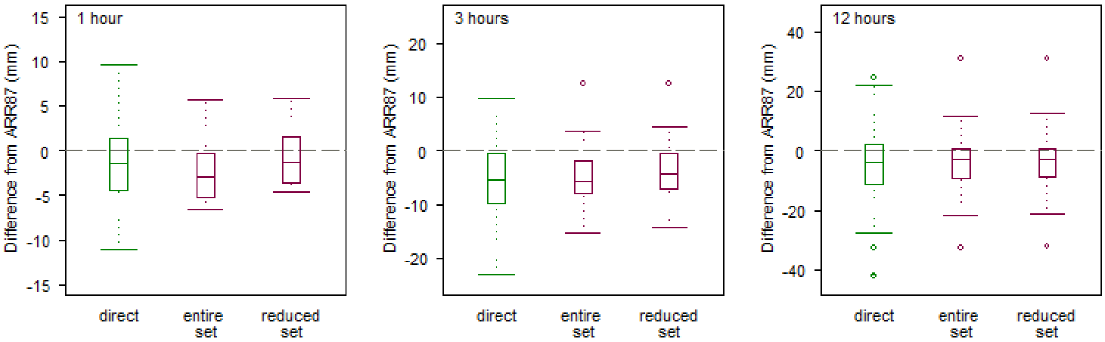


Figure 28 Differences between new estimates and ARR87.

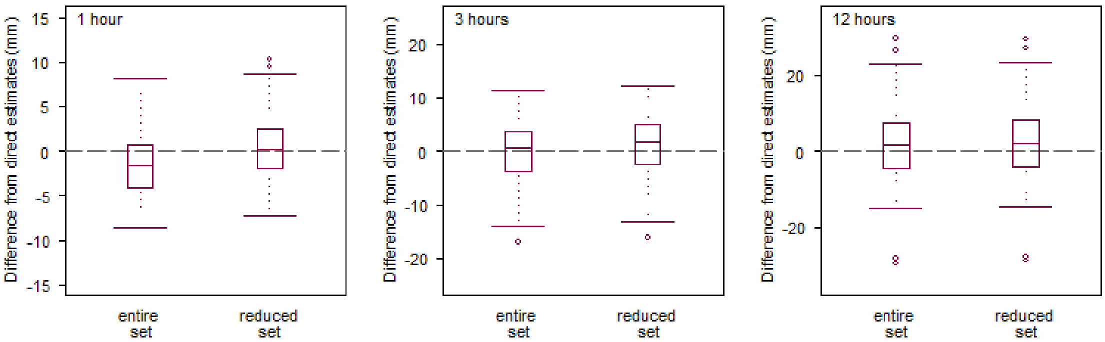


Figure 29 Difference between direct estimates and the two sets of predictions.

Typically the range of values is narrower for the ARR87 estimates than for direct estimates (although this could be partially due to smoothing during mapping procedures.) Apart from one outlier, the range for both sets of predictions is narrower than the range for both the ARR87 and the direct estimates. Both the direct and predicted estimates for the locations considered have a tendency to be lower than the ARR87 estimates, as can be seen from Figure 28 where negative values indicate that the new estimates are lower than the ARR87 estimates. Differences between the ARR87 and the direct estimates do not appear to be entirely random. There is a region where the ARR87 estimates tend to be lower than the direct estimates as shown in Figure 30.

Differences between the ARR87 and direct estimates (Figure 28) are of magnitudes similar to differences between direct estimates and the predictions, as shown in Figure 28 are of magnitudes similar to the differences between the direct estimates and the predictions seen in Figure 29. These magnitudes might give some indication of the uncertainties involved in estimating index rainfall. Differences between direct estimates and predictions for the entire set and the reduced set respectively are largest for 1 hour and negligible for the 12-hour duration. Judging from Figure 29, the predictions from the reduced set appear preferable at 1 hour although at 3 hours the predictions from the entire dataset might be marginally more suitable.

Based on these considerations, it was decided that the reduced set would be used for deriving predictions of index rainfall in the pilot study area. However, it was noted that choosing the entire set would yield very similar estimates and that the R^2 values shown in Table 6 are slightly higher for the larger dataset.

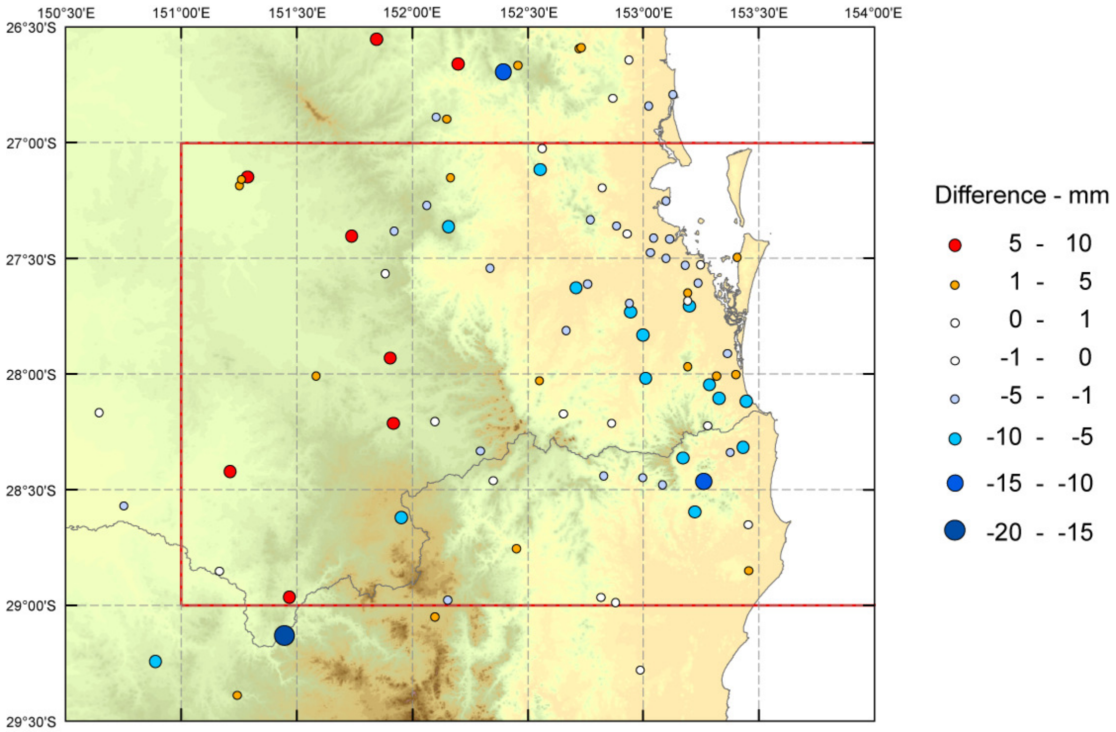


Figure 30 Differences between ARR87 estimates and direct estimates for 1-hour duration and an ARI of 2 years. Red shades indicate where ARR87 estimates are lower than direct estimates.

2.3 Prediction of L-CV and L-skewness at Sub-Daily Durations

The L-CV and L-skewness for sites in the pilot study area are not independent of each other. An improvement in model performance for predicting sub-daily values was achieved by predicting these variables jointly to exploit the additional information. The L-CV and L-skewness at 24 hours, latitude and distance from the coast were chosen as predictors in the PLSR model.

Previous investigations (HRS10, section 8.1.1 'estimating L-skewness') showed that the availability of long records is vital for deriving reliable estimates of L-skewness in particular. Therefore, the set of stations used in developing the PLSR model differed from that used for the index rainfall with only sites with at least 45 years of data being adopted. Previously this set consisted of only 30 sites but this number had increased to 40, mainly due to concerted efforts to digitise the backlog of Dines pluviograph paper charts. The R^2 values presented in Table 6 show that the model to predict L-CV and L-skewness did not perform nearly as well as that used to predict index rainfall. (R^2 is the fraction of the variation in the values of the predictand that is explained by the regression on the predictors.)

Table 6 R^2 values for Partial Least Squares Regression (PLSR) models to predict index rainfall, L-CV and L-skewness for durations from 1 to 12 hours.

Duration	1 hour	2 hours	3 hours	6 hours	12 hours
Index rainfall (entire set)	0.92	0.93	0.94	0.96	0.98
Index rainfall (reduced set)	0.87	0.86	0.89	0.94	0.97
L-CV	0.72	0.70	0.75	0.84	0.92
L-skewness	0.38	0.32	0.35	0.48	0.67

The option of assuming L-skewness as constant across durations was explored but rejected. Similarly, when inspecting boxplots showing variation of L-skewness with duration as shown in Figure 31, it could be concluded that variability with duration could be ignored. However, preliminary maps showed some spatial coherence in estimates of L-CV and L-skewness respectively. Figure 32 and Figure 33 show grids of predicted L-skewness and L-CV. L-CV values typically increase towards the coast. These maps also suggested that variation of L-skewness with duration is not entirely random although the pattern is perhaps less obvious for L-skewness (Figure 32) than it is for L-CV (Figure 33).

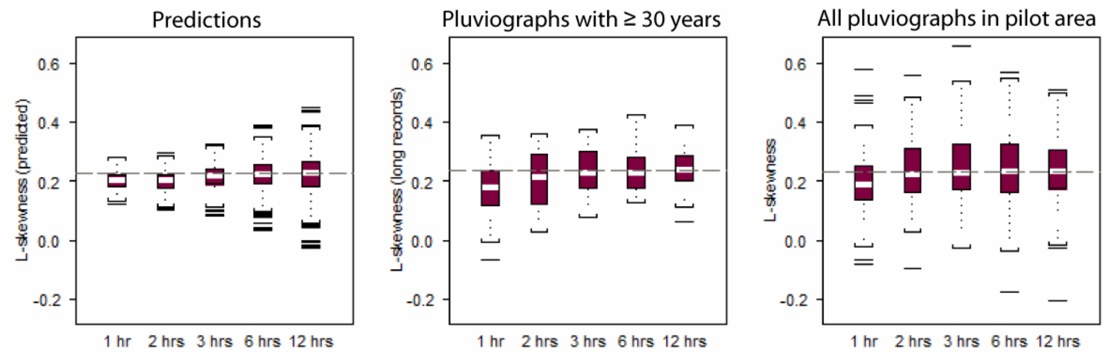


Figure 31 Boxplots of L-skewness estimates for pilot area. Dashed line indicates respective median for 12-hour duration.

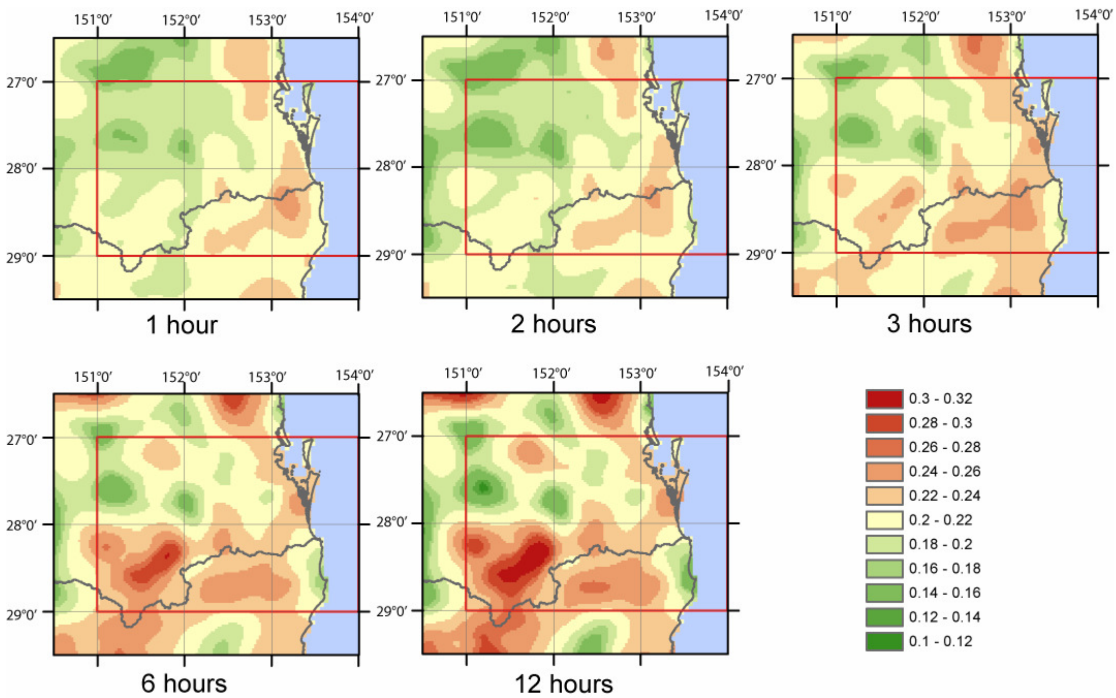


Figure 32 Predicted L-skewness for durations of 1, 2, 3, 6 and 12 hours.

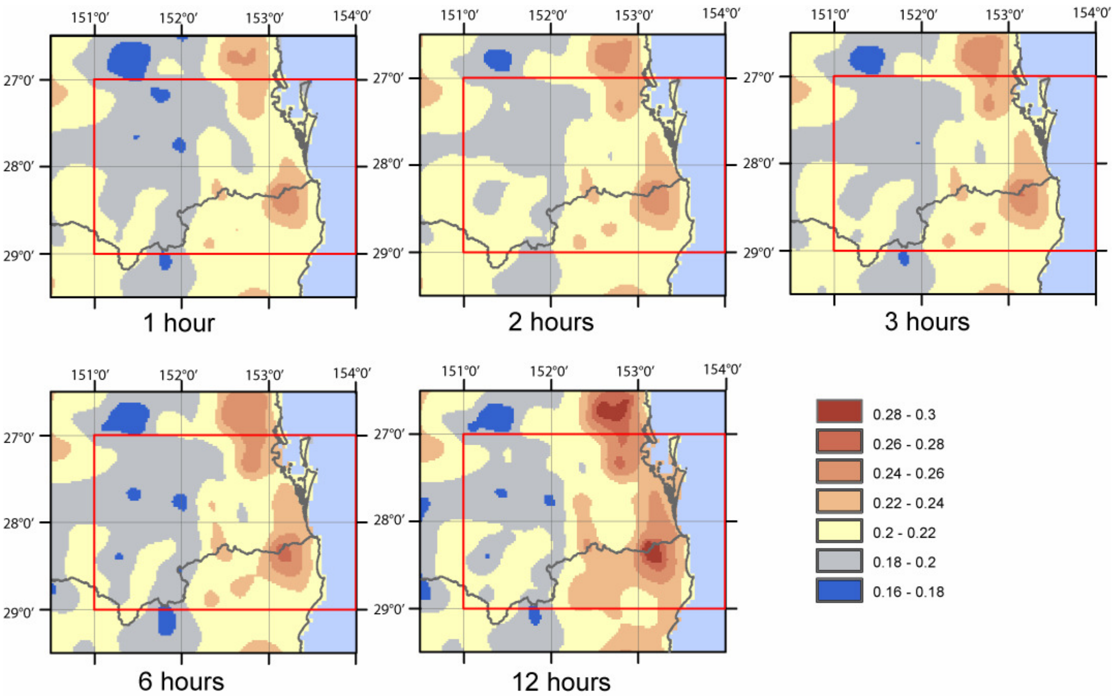


Figure 33 Predicted L-CV for durations of 1, 2, 3, 6 and 12 hours.

It should be noted that the estimates discussed here are at-site estimates rather than regional estimates. It is not suggested that at-site estimates should be considered the yardstick against which to measure the regional estimates, however, some comparisons might be useful in validating the adopted methods.

Typical growth factors for ARIs of 10 and 20 years are 1.6 and 1.8 respectively. Differences between at-site and regional growth factor estimates for these ARIs at 50 locations in the pilot study area are presented in Figure 34. The average differences between the regional and the at-site growth factor estimates for both ARIs are close to zero. Although there is a tendency for differences to increase with ARI there is little change from shorter to longer durations.

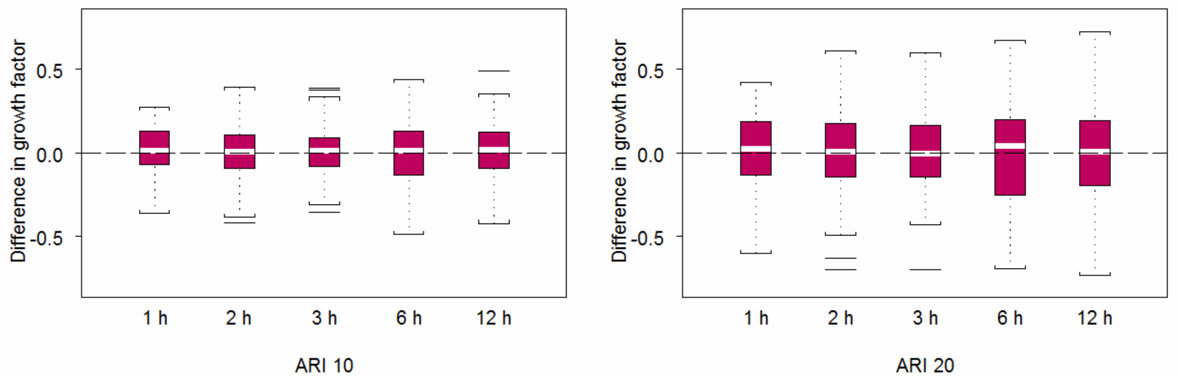


Figure 34 Differences between regional and at-site growth factor estimates for durations of 1, 2, 3, 6 and 12 hours at 50 locations in the pilot study area. Left for ARI 10 years, right for ARI 20 years.

2.4 Weighting Scheme

Regional L-moment estimates were calculated as the weighted average of at-site L-moments for sites within a region. At daily durations, the weights were based on the record length of the sites. None of the more complex weighting schemes explored, which included the distance from the focal site, showed any marked improvement and therefore were not adopted.

At subdaily durations, two sets of estimates with different degrees of uncertainty and different abundance had to be combined: a) estimates derived directly from pluviographs and b) predictions derived using the PLSR model. Direct estimates were only available at the location of pluviographs which typically have record lengths that are much shorter than for daily sites. In comparison, predicted estimates had been derived for all daily gauges. Weights based simply on record lengths (as for daily durations) would have placed too much emphasis on the predicted estimates resulting in the regional estimates being swamped by predictions. In addition, direct estimates were considered more reliable than predictions. The R^2 values from the PLSR model give an indication of the skill in predicting a variable and were therefore considered a good starting point to define weights according to the uncertainty attached to predictions. The final approach addressed both issues (abundance of data and confidence in estimates) as discussed in the following section.

2.4.1 Procedure for Deriving Regional Estimates of Regional L-CV and L-skewness at Subdaily Durations

1. The available pluviographs within a maximum radius (starting at 10 km and increasing by 5 km increments to 40 km) were searched until either the maximum circle size was reached or 200 station years (the sum of the record lengths of the pluviographs within circle) was found.

2. If the total number of available years fell short of 200 station years, predictions were used to make up for that shortfall. Station-years from daily sites were converted to 'pluviograph equivalent'; to take into account that there is less confidence in these values they were multiplied by a weighting factor, as discussed below. Starting from the focal site and a 10 km radius, the circles were extended until the sum of the number of station years from the pluviographs and the equivalent number of station years from the predictions reached 200 station years. A maximum circle size of 40 km was again used but circle sizes where predictions were included were typically much smaller than for direct estimates.

3. Regional estimates of L-CV and L-skewness were calculated based on: a) pluviographs and b) predictions. The two estimates were combined, using the number of station years (for pluviographs) and the equivalent number of station years (predictions) as weights and the following relationship.

$$L-CV_{regional} = \frac{\text{station years}_{pluvio} * L-CV_{pluvio} + \text{weighting factor} * \text{station years}_{prediction} * L\text{–}CV_{prediction}}{\text{station years}_{pluvio} + \text{weighting factor} * \text{station years}_{prediction}}$$

where

station years$_{pluvio}$; *station years*$_{prediction}$
are the station years for pluviographs and predictions respectively

L-CV$_{pluvio}$; *L-CV*$_{prediction}$
are the weighted average L-CV from pluviographs and predictions respectively (weighted by station years of individual sites)

Weighting of Predictions

The R^2 values as produced by the PLSR, and shown in Table 7, were considered as natural weighting factors. Their values are shown in Table 7. As expected the R^2 values increase with duration and are greater for L-CV than for L-skewness. It was considered that using different values for different durations might add to the known problem of inconsistencies between durations and therefore it was decided that the use of different weights for L-CV and L-skewness was unnecessarily complex and a single value should be used instead. The sensitivity of growth factor estimates to choice of weight was tested using values from 0.2 to 0.6 (Figure 35). Differences in growth factor estimates were typically less than 0.5%. On the

basis of these tests and taking the R^2 values into account, a weighting factor of 0.4 was chosen.

Table 7 R^2 values from PLSR model for L-CV and L-skewness and durations from 1 to 12 hours

	1h	2h	3h	6h	12h
L-CV	0.72	0.70	0.75	0.84	0.92
L-skewness	0.38	0.32	0.35	0.48	0.67

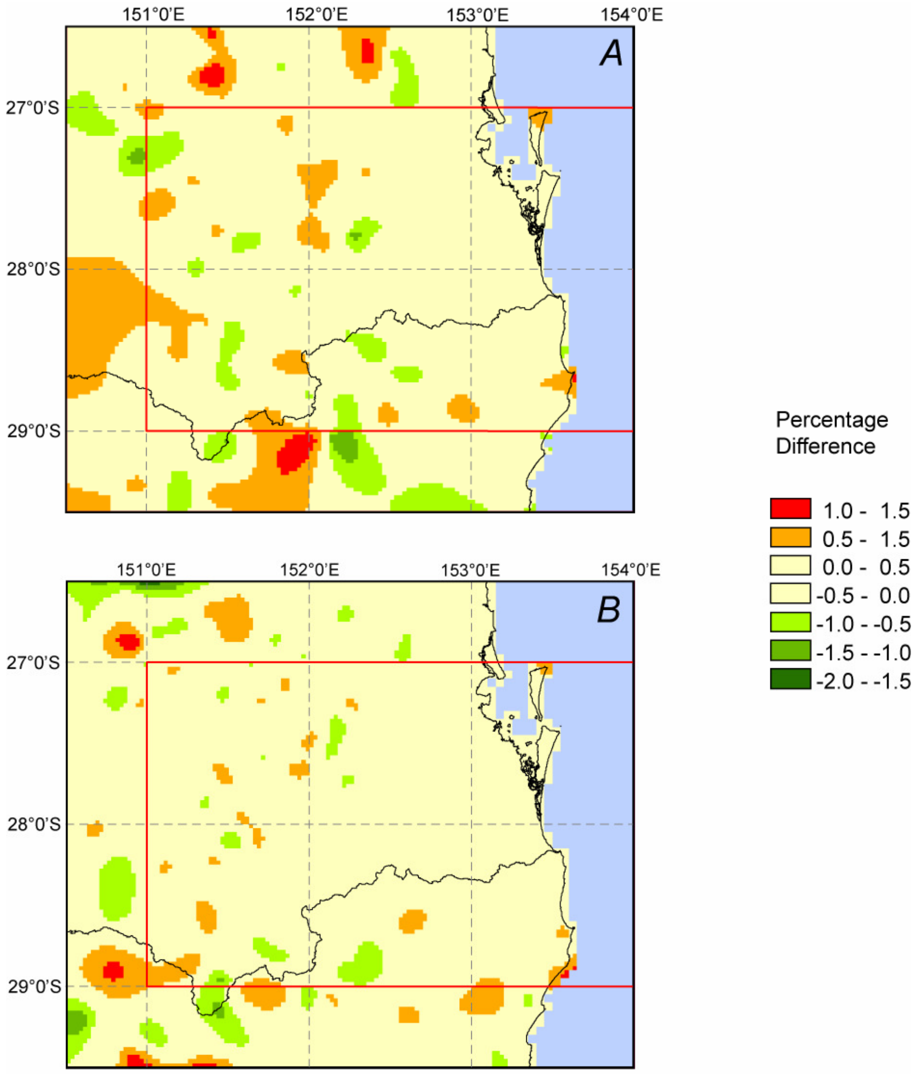


Figure 35 Percentage difference in growth factors depending on weight for the 3-h duration and an ARI of 5 years. A: comparing weighting factor of 0.2 and 0.4, B: comparing weights of 0.4 and 0.6.

3 VERY SHORT DURATIONS

3.1 Prediction of Sub-Hourly Design Rainfall

3.1.1 Introduction

Design rainfalls, depending on application, are required for a large range of durations. In Australia, density of stations recording daily rainfall as well as the lengths of these records is typically better than for sub-daily durations.

Previously, a statistical approach had been developed to use information about daily rainfall to ‘predict’ sub-daily design rainfall. This model performs well for the longer durations (12 and 6 hours) but only satisfactorily for the 1-hour duration. Both from a statistical point of view and considering meteorological processes, it appears sensible to develop a separate approach to estimate design rainfalls for sub-hourly durations.

For some countries (e.g. New Zealand), there are sufficient data to derive sub-hourly design rainfall estimates directly (without requiring the use of information from hourly design rainfalls). For other countries (e.g. UK), this issue is not specifically addressed.

ARR87 allows the derivation of estimates of design rainfalls for durations as short as 5 minutes, which are required, in particular, for urban applications.

3.1.2 Ideas for Deriving Estimates of 6-Minute Design Rainfalls

Methods for deriving design rainfall estimates for durations of less than 1 hour have not been finalised. This first section will introduce the problem, discuss a possible approach and highlight some of the issues. In subsequent sections, approaches adopted in Australia and overseas are summarised.

Objective:
Predict subhourly (6, 12, 18, 30 min) design rainfalls for ARIs of 2, 5, 10, 20, 50 and 100 years for the pilot study area based on 1-h estimates for ARIs of 2, 5, 10, 20, 50 and 100 years.

Approach:

- Preliminary findings show that considering all Australian data might be preferable to focussing on the pilot study area as there are very few stations upon which to build a model.
- Work has commenced on the derivation of a model based purely on at-site estimates (6 min and 1 h, up to ARI 100 years).
- To date, two datasets have been considered, one with 329 stations and another with 203 sites (with minimum record length of 30 years). These typically give very similar results.

Reasoning:

- In deriving 1-h estimates regional information has been taken into account.
- Use has been made of information at longer durations (24 h, PLSR).
- Currently, suitable regional 1-h estimates are not available outside the pilot study area.
- Using estimates from ARR87 as a replacement may not be appropriate as they will probably be quite different to estimates that might be derived as 1-h grids.
- It is considered that there is little value in attempting to predict L-moments and index rainfall for subhourly durations because of uncertainty in 1-h estimates but rather the 1-h maps should be ‘scaled’ in an appropriate way. This is essentially what was done

for ARR87 and by the Hydrometeorological Design Studies Centre (HDSC, NOAA) and is basically a scaling approach.

Issues:
- Site estimates for an ARI of 100 years from short records are doubtful.
- Maps showing ratios of 6-min to 1-h depths at ARI of 100 years are 'erratic' - single events may drive these estimates.
- The GEV may not be appropriate for the 6-min duration – a poor fit would result in problems when estimating at an ARI of 100 years.

Main issue:
Based on maps of 1-h design rainfalls, maps of 6-min design rainfalls for the pilot study area are required to be produced. However, for building the predictive relationship it is most likely that it will be necessary to move outside the pilot study area in order to increase the sample size and suitable gridded or even regional estimates will not be available. The only feasible way might be to develop a model based on at-site estimates (6-min and 1-h). For the predictions, regional or gridded 1-h estimates could be used as predictors. The inclusion of direct, at-site 6-min estimates at the mapping stage should be considered.

3.1.3 Data

The analyses were based on at-site estimates. Therefore, only stations with at least 30 years of data were considered. Rainfall at sub-daily durations is recorded by Dines pluviographs and Tipping Bucket Raingauges (TBRG). The pluviograph records rainfall on a paper chart which has to be digitised later. For high rainfall intensities, it is difficult to digitise rainfall depths accurately. It is suggested that the lowest sensible resolution in time is about 6 minutes. The Tipping Bucket Raingauge converts each fill (0.2 mm of water) into an electric pulse which is measured by a counter and sent to an Automatic Weather Station (AWS). Rainfall intensities can be derived for 1-min intervals. For quality control purposes, daily accumulations from pluviograph and/or TBRG are routinely checked against rainfall measured by a daily raingauge.

Considering the paucity of data for durations below 6 minutes, a dedicated approach will be required to derive estimates at these durations.

3.1.4 Approaches

Australia

In ARR87, for durations of less than one hour, the logarithm of 6-min intensity was related to the logarithm of the 1-hour intensity by a scaling factor and a constant ('geographical coefficient'). Intensities were derived for the two key ARIs of 2 and 50 years. The 6-min intensities for the key ARIs were calculated from the 1-h intensities of the same ARI. Values for durations of 10, 20 and 30 min were derived using interpolation between two of the ARI standard durations of 6 minutes and 1 hour. 5-min values were derived by extrapolation based on the 6-min and 1-h intensities.

Canada

Nguyen & Nguyen (2008) suggest using a simple scaling approach for Quebec (Canada) based on their analyses of 15 stations. Different scaling properties were found for durations between 1 hour and 24 hours and between 5 minutes and 1 hour as shown in Figure 36.

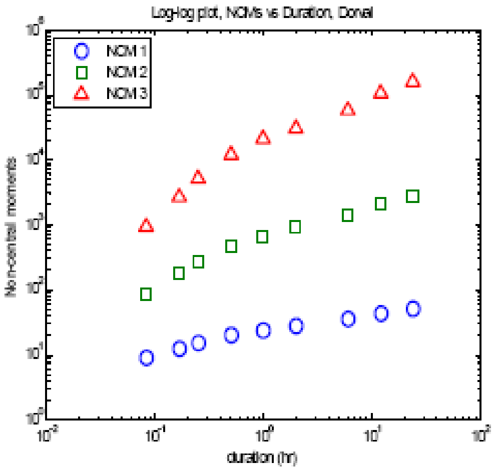


Figure 4 – The log-log plot of maximum rainfall non-central moments versus rainfall duration for Dorval station.

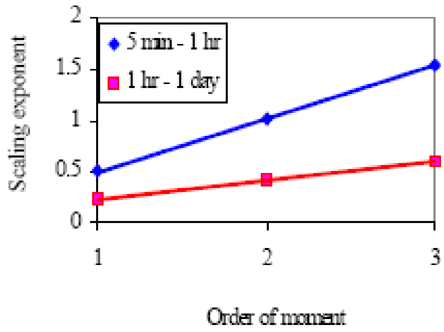


Figure 5 – The plot of the scaling exponent $\beta(k)$ versus the order k of maximum rainfall non-central moments for Dorval station.

Figure 36 Simple scaling for deriving sub-hourly estimates (source: Nguyen and Nguyen, 2008).

3.1.5 Results from NOAA

One of the simplest approaches is based on calculating ratios of 1-hour design rainfall depth (in mm) to 6-min design rainfall. For the US, these ratios were calculated for a number of average recurrence intervals, (Bonnin et al 2004a). Ratios decrease with decreasing duration and with increase in ARI. For example, for the 'northern region', the ratios range from 0.815 (for 30 min, ARI 2 years) to 0.261 (for 5 min and ARI 1000 years). The 'northern region' comprising of locations in Illinois, Indiana, Ohio, West Virginia and additional stations in neighbouring states (see Figure 4.1.5 I NOAA Atlas 14, Volume 2, Version3).

According to NOAA Atlas 14 (Volumes 1, 2 and 3) ratios of design rainfall estimates were derived for a range of ARI (1 to 1000 years). The ratios are calculated as:
$depth_{n\ minutes}/depth_{60\ minutes}$ (for given ARI).

If rainfall fell uniformly over time, one would expect to find for instance ratios of 0.5 for the 30-minute duration and 0.1 for the 6-minute duration. According to the numbers presented in NOAA Atlas 14, Volume 2 (Bonnin et al 2004b) one might expect that

a) these ratios depend on geographic location (northern and southern region) and
b) that ratios decrease for higher ARI.

Table 8 presents the ratios of rainfall depths for the northern region of NOAA Atlas 14, Volume 2, Version 3: for durations of 5, 10, 15 and 30 minutes to the 60-minute duration. (Bonnin et al 2004b, Table 4.1.3). The 1.58-year value was computed to equate the 1-year average recurrence interval (ARI) for partial duration series results.

Table 8 N-minute ratios for the northern region of NOAA Atlas 14, Volume 2, Version 3.

AEP	*5-min*	*10-min*	*15-min*	*30-min*
1 in 1.58	0.325	0.505	0.619	0.819
1 in 2	0.319	0.498	0.609	0.815
1 in 5	0.305	0.474	0.582	0.797
1 in 10	0.298	0.46	0.566	0.786
1 in 25	0.289	0.442	0.546	0.771
1 in 50	0.283	0.429	0.531	0.759
1 in 100	0.277	0.417	0.518	0.748
1 in 200	0.272	0.406	0.505	0.737
1 in 500	0.266	0.391	0.488	0.723
1 in 1000	0.261	0.38	0.475	0.712

3.1.6 Ratios for Australian Sites

Using the approach adopted by NOAA, ratios were estimated for Australian sites in the following manner:

Annual maxima (for durations down to 6 minutes) with record length of at least 30 years were available at 203 sites in Australia. For each duration (6, 12, 18, 30 and 60 minutes) at-site L-moments were calculated and GEV were fitted to estimate ARIs of 2, 5, 10, 20, 50 and 100 years. For ARIs of less than 10 years, the estimates were adjusted using the AM/PDS adjustment factors.

Figure 37 shows the average ratios for durations of 6, 12, 18 and 30 minutes. For durations longer than 6 minutes, ratios show a decline with increasing ARI. For the 6-minute duration however ratios appear to (on average) increase with ARI. For comparison the ratios derived by the Hydrometeorological Design Studies Center (HDSC, NOAA) for the northern region are shown (note the different durations). There is some general agreement for the 30-minute duration although the decrease for higher ARI appears stronger for NOAA.

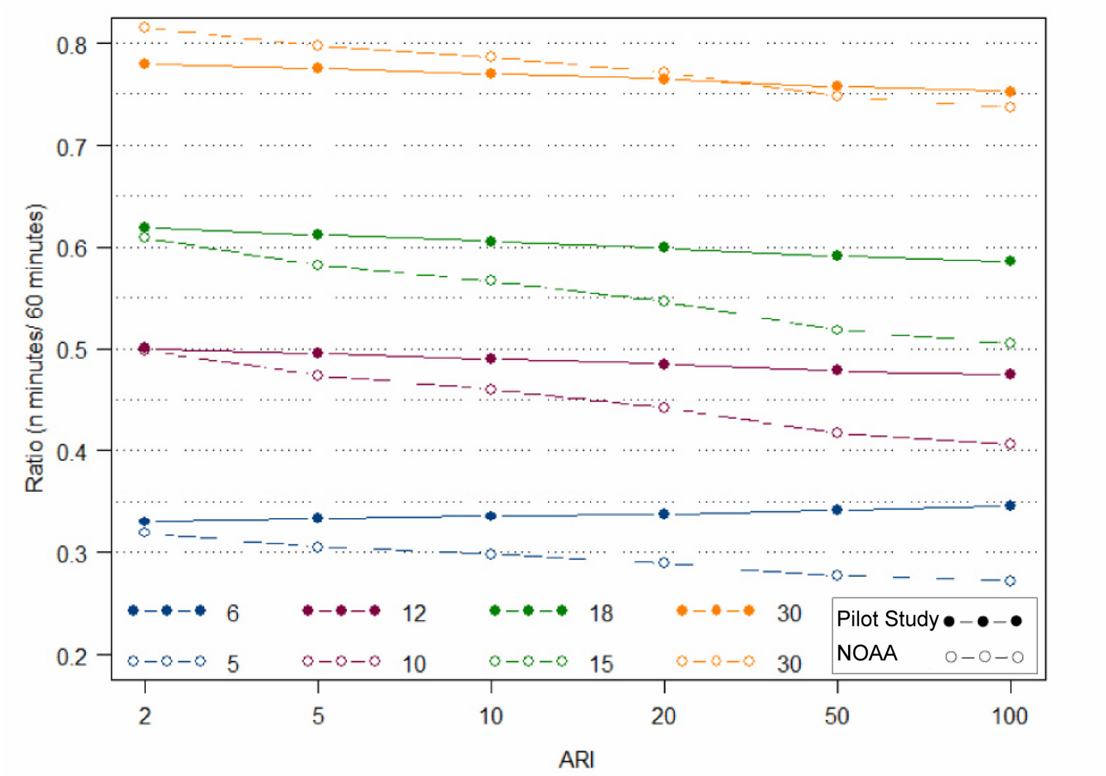


Figure 37 Ratios of design rainfall depths of duration *n* to design rainfall depth at 60 minutes

Boxplots were prepared to give an idea of the spread in ratios for given durations as shown in Figure 38. The dashed lines indicate the relevant ARI 2 year estimate. From these plots it becomes obvious that at any given duration there are always a number of stations where the ratio at ARI of 100 years is higher than the ratio for ARI of 2 years. It should also be noted that for 30 minutes there are some ratios at ARI of 100 years that exceed 1. It is considered that these inconsistencies require correcting.

Following the argument made in NOAA, it would be conceivable that certain regions of Australia show a different tendency of ratios with ARI compared to other regions. The location of stations was therefore mapped, using colour to indicate whether the ratio at ARI of 2 years is above or below the ratio at an ARI of 100 years as shown in Figure 39. The pattern is remarkably similar from one duration to the next which could be partially due to the fact that the events might not be independent of each other as the durations increase.

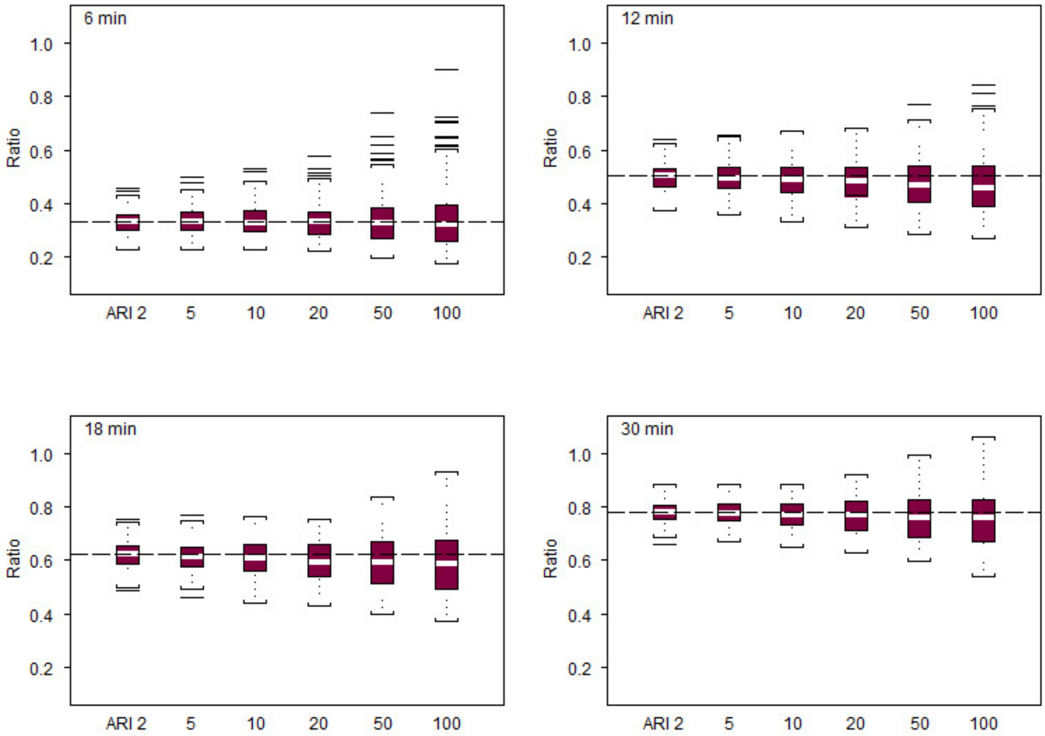


Figure 38 Boxplots of ratios of design rainfall estimates of duration n to design rainfall estimate at 60 minutes, by duration and ARI.

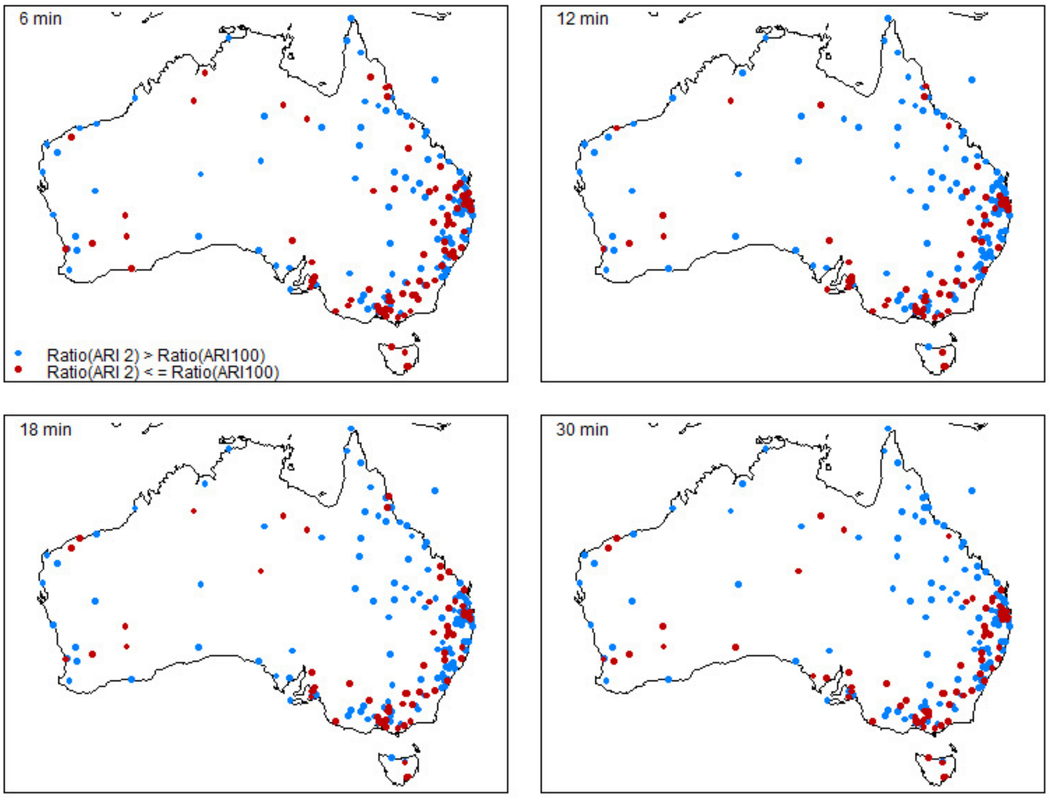


Figure 39 Location of stations where ratio increases from ARI of 2 years to ARI of 100 years.

Techniques

Assuming subhourly rainfall estimates can be derived from hourly rainfall depths by applying a factor that varies only with duration and ARI (but not with geographical location), one can set out to identify the most appropriate techniques for estimating these factors.

The simplest method is to calculate ratios of say 6-min to 60-min rainfall depths at a sufficiently large number of stations (with records of sufficient length to derive estimates of up to ARI of 100 years). This analysis was undertaken based on all-Australian pluviographs with at least 30 years of data (192 sites). During the course of investigation, at three sites the highest ranking annual maximum were identified as suspect and these three sites were removed from further analyses. The excluded sites were gauge numbers 55194, 56018 and 82042.

Rather than calculating the average of ratios at all stations (say ARI 2, 6 min to ARI 2, 60 min), the ratio of averages can be calculated (average ARI 2, 6 min across all stations to average ARI 2, 60 min across all stations). The second approach leads to a more robust estimate and appears superior based on the percentage error shown in Figure 40. The 'percentage error' here refers to the RMSE divided by the average value for a given ARI and duration. Model performance was worst for the highest ARI (ARI of 100 years) and the shortest duration (6 minutes), reaching about 30%.

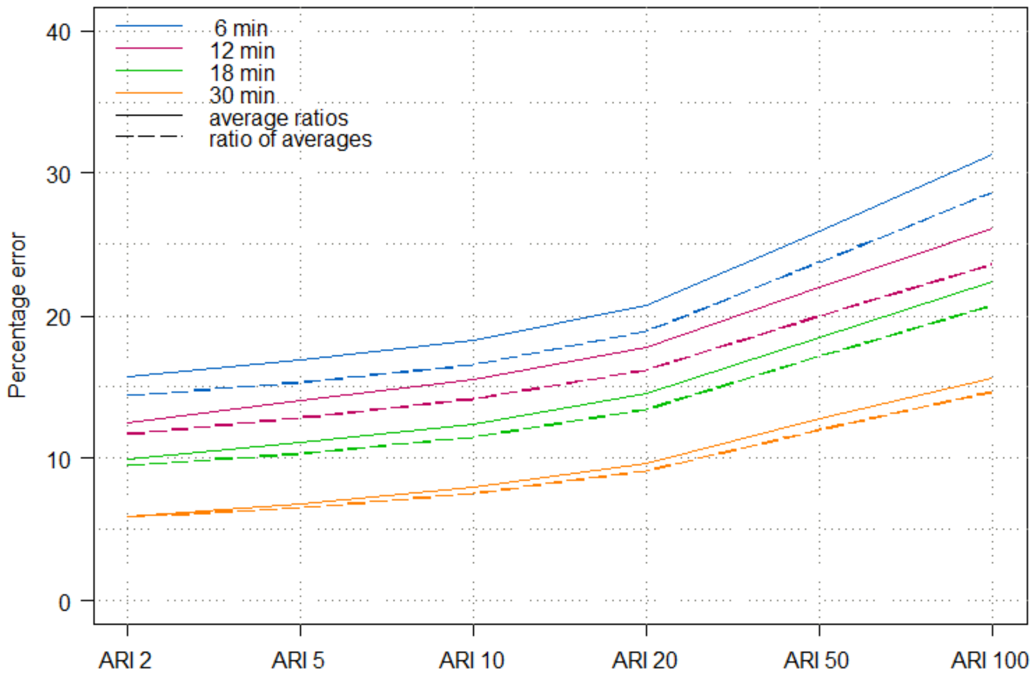


Figure 40 Percentage difference for subhourly rainfall depths predicted using average ratios and ratio of averages.

Initial assessments using scatterplots of rainfall depths showed that the relationships were not linear but neither did they follow a log-log relationship very closely. An approach based on a power law assumption was unsuccessfully trialled. While typically performing equally well or better than *average ratio* and *ratio of averages*, percentage error at ARI of 100 years for the 6-min duration exceeded 50%.

Alternative approaches based on two regression techniques were tested to try and improve the predictive skill for subhourly durations: non-linear least squares (NLS) regression and generalised least squares (GLS) regression. Using a GLS approach, errors are 'allowed' to be correlated and/or have unequal variance, while for NLS errors are assumed Gaussian and independent. The two techniques perform similarly well, with GLS showing slightly superior performance at lower ARIs.

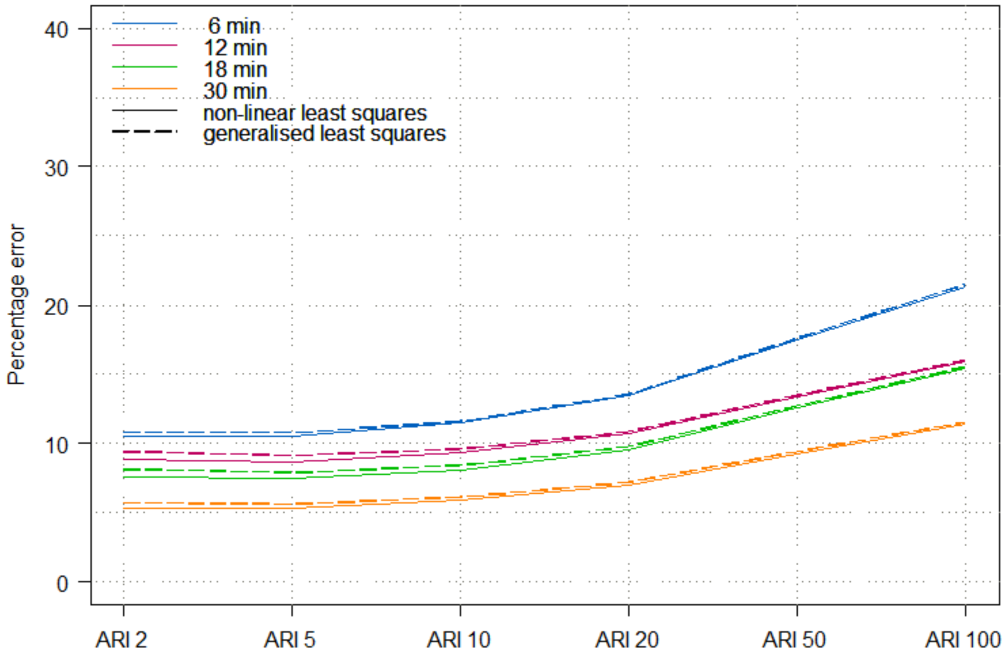


Figure 41 Percentage error in subhourly rainfall depths predicted using non-linear least squares and generalised least squares regression.

Differences in model performance between the average ratio and the ratio of averages on the one hand and non-linear least squares and generalised least squares regression on the other hand are particularly large at an ARI of 100 years for the 6-min duration. For this combination of ARI and duration, the absolute residuals (6-min rainfall depths derived based on annual maximum series and predicted from the 1-h rainfall depth) were plotted for each of the 4 approaches as shown in Figure 42.

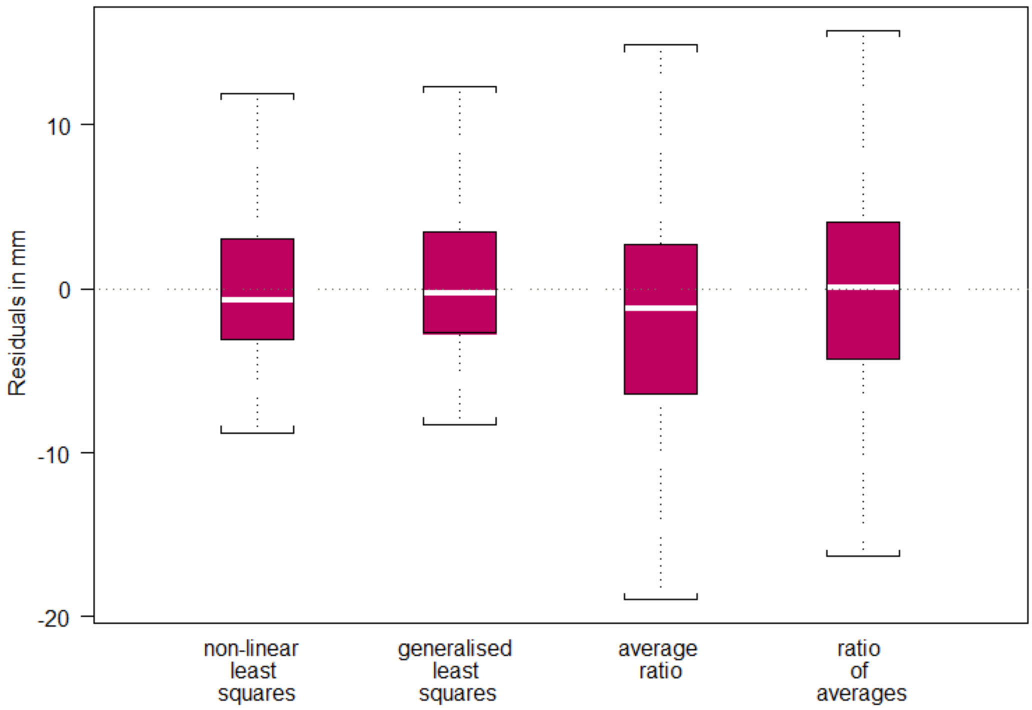


Figure 42 Absolute error in predicted 6-min rainfall depths at ARI of 100 years for all Australian data (minimum record length 30 years).

On the basis of these analyses, the GLS approach was selected. To validate the model, predictions were made for 104 sites in the pilot study area with a minimum record length of 8 years. Absolute differences in direct and predicted estimates exceeded 20 mm but this was typically only a problem at ARI of 100 years as shown in Figure 43. Given the available record lengths, direct estimates for ARI of 100 years can only be considered a rough approximation. The largest percentage error of 28% was associated with predictions for 6 minute, ARI100 estimates as shown in Figure 44.

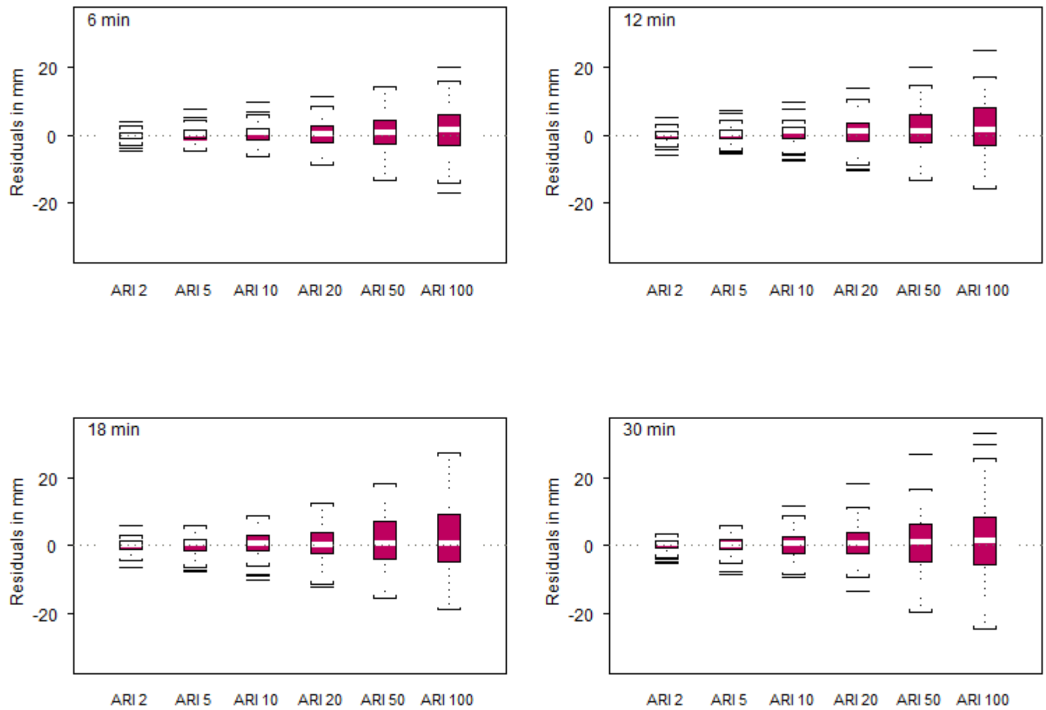


Figure 43 Boxplots of absolute error for rainfall depths predicted using a GLS approach. For stations in the pilot study area with a minimum record length of 8 years; for 6, 12, 18 and 30 minutes and for ARI of 2, 5, 10, 20, 50 and 100 years.

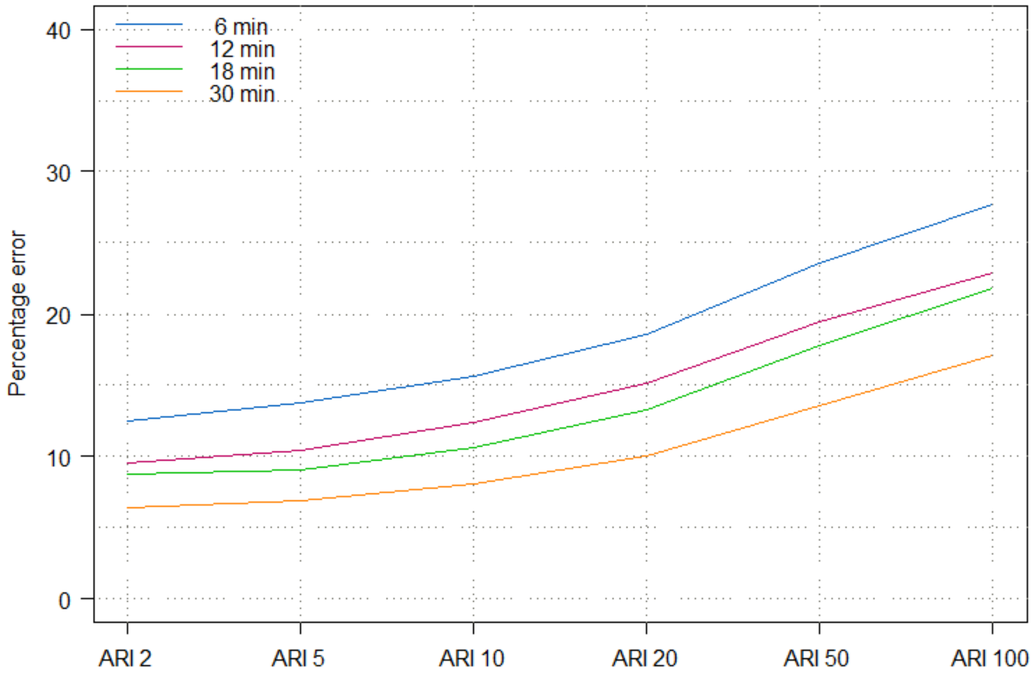


Figure 44 Percentage error for subhourly rainfall predicted using a GLS approach (for stations in the pilot study area with a minimum record length 8 years).

Recommendation

Judging by percentage and absolute errors, approaches based on *generalised least squares* regression and *non-linear least squares* regression result in better estimates than *average ratios* and *ratios of averages*.

There are two concerns: one is to avoid introducing inconsistencies when converting from 1 h to 6, 12, 18 and 30 min. The other is that while for most stations, the ratio of 1h depths to shorter duration depth decreases with higher ARI, for other stations the opposite is true. This may not be a statistical artefact from the fitting of distributions or sampling issues but could be due to meteorological factors. Data are insufficient to develop sensible maps that would assist in exploring this further. However, there would be value in investigating whether additional information, such as catchment boundaries to describe barrier effects, should be made use of.

3.2 Inferring 5-minute Depths from 6-minute depths

It is considered that it is not possible to obtain genuine 5-minute data from most digitised Dines pluviometers. As a consequence, it is likely that the maximum 5-minute depth will often be an interpolated version of the maximum 6-minute depth, having the same intensity and therefore be 5/6 (0.83) of the 6-minute depth (for the same event). To estimate the 5-minute depths, therefore, it has been necessary to confine the data to TBRGs which are easily capable of 5-minute resolution – and less.

Annual maxima were derived for 236 TBRG (islands were excluded from the analysis). Annual maxima were extracted for years where at least 10 months had at least 75% of data available. For the following analyses only stations with at least 9 annual maxima were considered – a total of 116 stations.

It was found that the 5-min and 6-min depths were likely to be from the same storm and the 5-min burst was likely to fall within the 6-min burst. In 90% of cases (1744 events out of 1926 events) the end time of the two bursts differed by less than 3 minutes.

The ratio of 5-min to 6-min rainfall depths was calculated for all pairs of annual maxima, although these were not always from the same event. The results are shown in Figure 45. Ratios range from 0.780 to 0.996, with a median of 0.884. (Results are virtually identical if the ratios were calculated only for rainfall depths from the same event.)

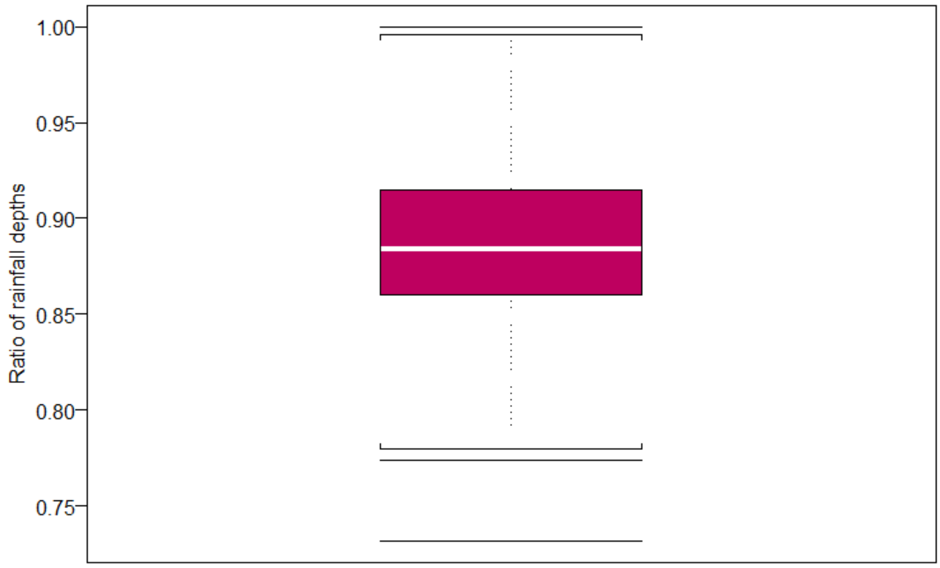


Figure 45 Ratio of rainfall depths for 5-min and 6-min durations for pairs of annual maxima.

Based on this analysis, it appears justifiable to use a constant factor to derive estimates of 5-min rainfall depths from 6-min rainfall depths regardless of ARI. To test this hypothesis 2, 5 and 10-year ARIs were estimated by fitting GEV distributions. The results of the analysis are presented in Figure 46.

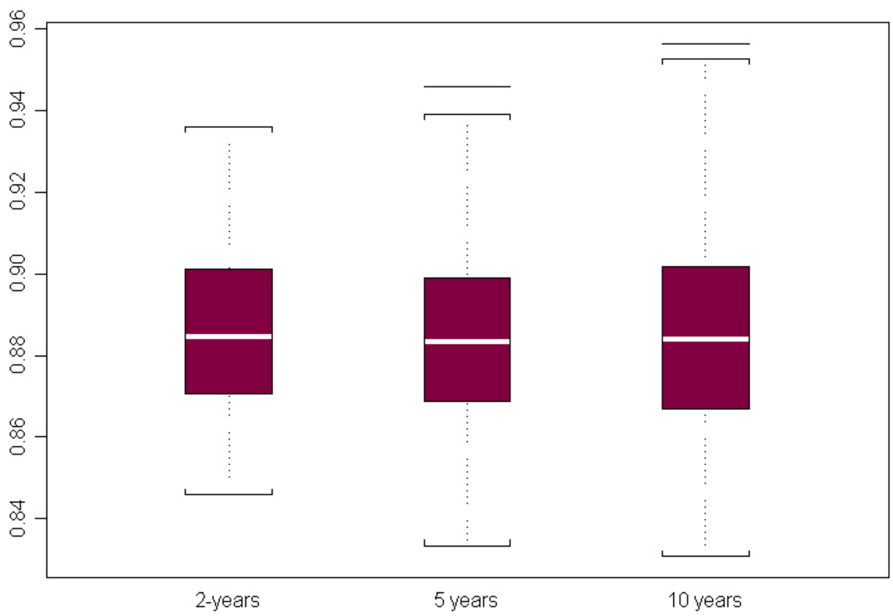


Figure 46 Boxplots of ratios for 5-min to 6-min rainfall depths

The median values are: 0.885, 0.883 and 0.884 for ARIs of 2, 5 and 10 years respectively. The variation with ARI is not monotonic (based on 3 values). Hence, on the basis of the evidence available, it is reasonable to use a factor of 0.88 for the ratio of 5-minute to 6-minute depths for all ARIs. This means that the 5-minute and 6-minute isohyets will have exactly the same shape. However it is felt that, given the closeness of these durations this is not unreasonable.

It is recommended that the 5-minute depths be calculated as 0.88 of the 6-minute depths for all ARI.

4 MAPPING

A variety of mapping approaches were used in developing currently valid procedures for design rainfall estimation. Examples include an approach based on geo-regression (Faulkner, 1999), thin-plate spline smoothing (Thompson, 2002) and a combination of objective mapping techniques and expert knowledge (Daly et al, 1997). For ARR87, a complex procedure was adopted that involved the derivation of final estimates based on combining master charts, an approach that involved a considerable degree of subjectivity. Despite certain short-comings, there may be reason to prefer objective methods over subjective approaches.

Design rainfall estimates from ARR87 are supplied at a resolution of 0.025°. Design rainfall estimates developed in the pilot study were derived at the same resolution. For the pilot study, a thin-plate spline smoothing approach as implemented in the software package ANUSPLIN (Hutchinson, 2007 and Tait et al, 2006) was used. While users have to select appropriate predictors and required transformations, the degree of smoothness for the surface can be selected automatically using Generalised Cross Validation (GCV). In this approach, each observation is left out in turn, the surface is fitted without this observation and the residual at this location is calculated. The GCV is then calculated as the (weighted) sum of squares of these residuals and the optimum smoothing parameter is found by minimising the GCV. Hutchinson (1998) discusses the choice of order of spline, independent spline variables versus covariates, appropriate resolution of digital elevation models (DEM) as well as choices for transformation and scaling for mapping daily rainfall.

For the subdaily durations, a number of different maps were prepared: grids of median (index rainfall), grids of k and β (parameters of the GEV distribution) and grids of growth factors. Grids of preliminary design rainfall estimates were derived using ArcGIS. All the grids were derived at a resolution of 0.025°, which matches that of the Computerised Design IFD Rainfall System (CDIRS), developed by the Bureau of Meteorology to allow automatic determination of a full set of IFD curves, and at the same extent as the pilot study area (including the buffer zone): 150.5°E to 154.0°E and 26.5°S to 29.5°S.

A set of trials was undertaken to select the independent variables/independent covariates, order of spline and smoothing directive. These trials used the 12-hour medians (index rainfall) initially, as this was considered the most reliable duration; being closer to the data rich 24-hour duration. Candidates for the independent variables included elevation and mean annual rainfall (MAR), both of which were available at a resolution of 0.025°. A coarser DEM grid with a resolution of 0.5° was used.

In addition to the independent variables (or covariates) adopted, there may be value in considering additional variables such as 'orographic effectiveness of terrain' and 'coastal proximity' as done by Daly et al (2002).

4.1 Choice of Variables to be Mapped

Apart from the mapping technique, approaches vary in the choice of variables to be mapped. Thompson (2002) selected the mapping of distribution parameters, while Wallis et al (2007) mapped index rainfall, regional L-CV and L-skewness.

Three sets of characteristics that could be mapped are:

a) regional growth factors (for ARI of 2, 5, 10, 20, 50 and 100 years)
b) regional GEV parameters (β and k)
c) regional L-moments (L-CV and L-skewness)

The considerations in choosing a set of characteristics to be mapped include:

1) the number of maps/grids required,
2) the spatial consistency of parameters (pertaining to the sensitivity of results to choice of mapping procedure) and
3) the requirement to ensure consistency across durations.

If the growth factor at a location is higher for a shorter duration than at a longer duration (for the same ARI), it does not necessarily imply an inconsistency in design rainfall estimates. It is therefore not sensible to attempt adjusting maps of growth factors for inconsistencies and choosing to map growth factors on these grounds is not a convincing argument. The number of grids/maps required at each duration (in addition to index rainfall) is two for the GEV parameters (or L-moments) but 5 for the growth factors.

Sensitivity tests were undertaken to test how sensitive the resulting grids of growth factor estimates were to the choice of variable. Based on comparisons of estimates for the 6-h duration at an ARI of 20 years, differences in gridded growth factors were minimal. It was therefore decided to map regional k and β (rather than mapping growth factors directly) because fewer grids are required.

4.2 Mapping Index Rainfall at Subdaily Durations

A set of trials was undertaken to select the independent variables/independent covariates, order of spline and the smoothing directive which would produce the best compromise between reproducing the values at the location of sites and producing a smooth map without lots of 'bullseyes', that is, individual sites at the centre of a target of concentric isolines. Candidates for the independent variables included elevation and mean annual rainfall. These trials used the 12-hour medians (index rainfall) initially; this being considered the most reliable duration as it is closer to the data rich 24-hour duration. The trials were then run again for the 1-hour medians to test if the results were reasonable. After about 15 trials, a set of inputs for ANUSPLIN was settled upon. Details of these trials may be found in Appendix 0. Medians for both durations (1 and 12 hours) were derived from the annual maximum series of rainfalls.

There were 88 direct estimates at sub-daily rainfall recording sites as shown in Figure 47. The figure demonstrates the paucity of direct estimate sites especially in the west. This prompted the incorporation of elevation and mean annual rainfall into the mapping process. Figure 48 shows the distribution of the 646 daily rainfall observing sites on which the subset of predictions is centred (see Figure 26 for details). The sites are fairly well distributed and sufficiently dense to map without additional inputs.

Values of the grids produced for the various trials were extracted at the input sites and compared with the original values in a spreadsheet; the aim being to approach a version with both the smallest differences in values and the least number of large differences. At the same time, a smooth gradient of values across the mapped region needed to be maintained. As mentioned previously, trying to achieve an exact match between grid and original values would result in an unrealistic series of bullseyes.

4.2.1 ANUSPLIN Options

The chosen set-up for ANUSPLIN to map the direct estimates and produce grids of index rainfall used the following programs and inputs:

- Programs used were **SPLINA**, the simpler of the splining options as this is suitable for data sets with up to 2000 points, and **LAPGRD** which is the program used for generating grids. The output grids were in *.asc* format suitable for reading directly into ArcGIS and at a grid spacing of 0.025°.

- ***Longitude and latitude*** were used as independent variables within the limits of 150.5°E to 154.0°E and 26.5°S to 29.5°S.

- ***Elevation*** was used as an independent variable. A DEM in metres and at a fine grid spacing (0.0025°) was manipulated in ArcGIS to create a coarser grid. This was achieved by aggregating the elevations over a larger area in order to dampen out very local features; the averaging was done over a spacing of 0.5° but the grid has an actual spacing of 0.025° to match all the other grids. Figure 47 shows this grid. The elevations were divided by 1000 in the SPLINA transformation parameters section as recommended in the Fitting Climate Surfaces section of the 'ANUSPLIN version 4.37 User Guide' (Hutchinson, 2007). The base DEM was originally downloaded as 'GEODATA 9 Second DEM' from the Geoscience Australia website http://www.ga.gov.au/.

- ***Mean annual rainfall*** (MAR) was used as an independent variable. The grid was based on one obtained from the Bureau's National Climate Centre and has a grid spacing of 0.025°, see Figure 48. The SPLINA transformation option to use the logarithm of the MAR was chosen.

- A ***square root transformation*** was applied to the dependent variable, that is, the direct rainfall estimate. This is recommended in the *Dependent Variable Transformations* section of the 'ANUSPLIN version 4.37 User Guide' (Hutchinson, 2007). The dependent variable had the transformation applied to it in SPLINA and was back-transformed in LAPGRD when the grid of values is generated.

- ***A Spline order of 3*** was selected (orders 2 to 4 were tested).

- As smoothing directive, the ***fixed signal to noise ratio*** option was chosen. This option was pursued after a recommendation from Craig Thompson (pers. comm.) of the National Institute of Water and Atmospheric Research, New Zealand. Trials suggested that a ratio of 4:1 worked best.

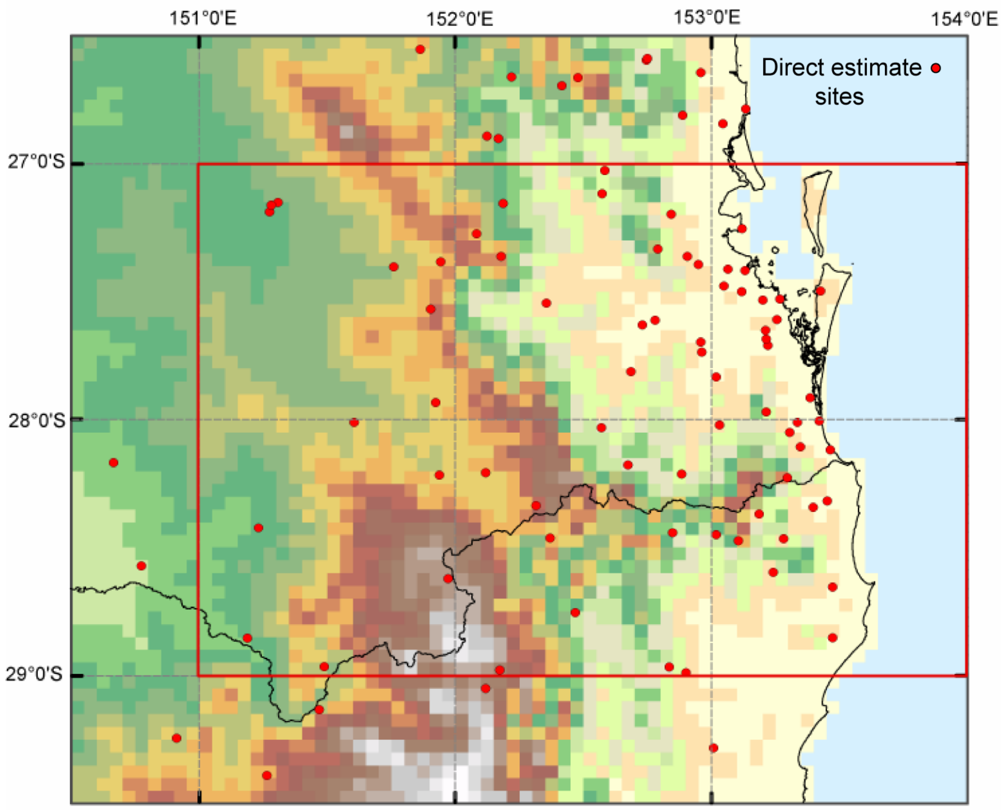


Figure 47 Direct estimate sites and elevation grid shown on the DEM coarsened to 0.5°.

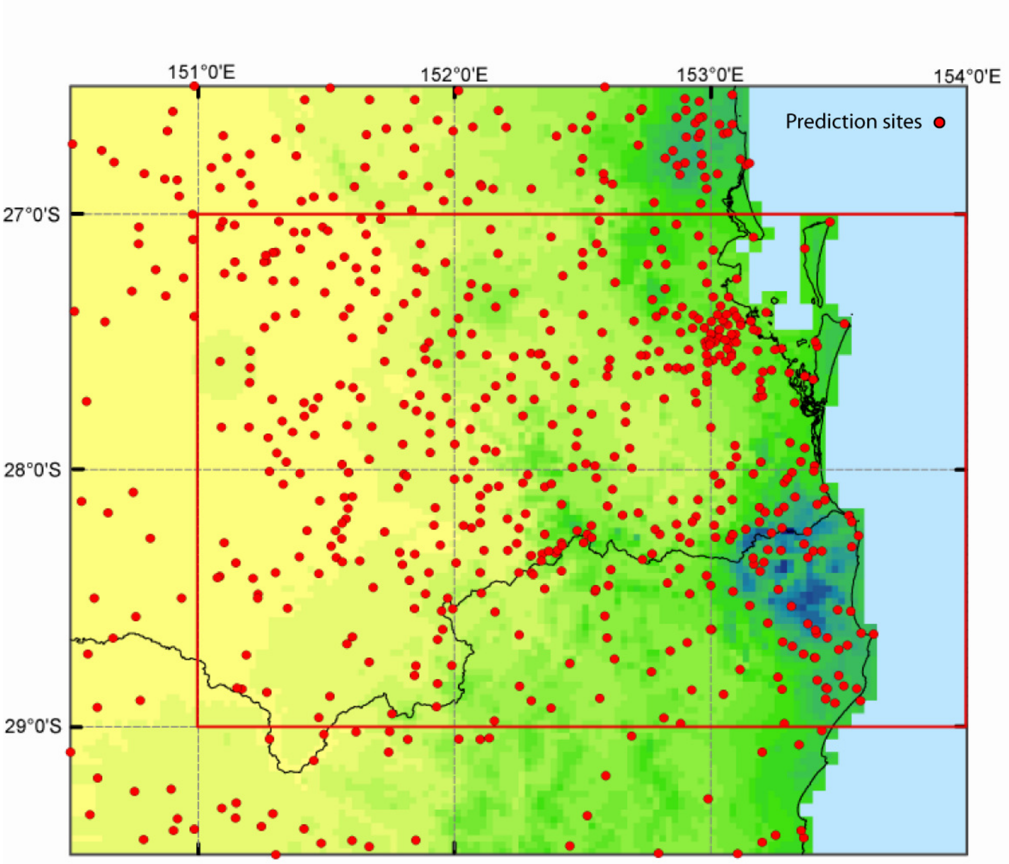


Figure 48 Prediction sites and mean annual rainfall grid.

- Predictions from a PLSR model were used as a dependent ***covariate.***

- A ***grid of predicted values*** was generated as the initial step of the mapping exercise. Only the prediction values were input into SPLINA as the 646 sites shown in Figure 48 gave sufficient coverage to obtain a smooth and accurate grid. The splining parameters were:
  - Spline order 3,
  - Square root as dependent variable transformation and
  - Generalised Cross Validation (the ANUSPLIN default) as smoothing directive.

- It was discovered that ANUSPLIN ignored the direct estimate values (dependent variable) in all cases where there was no prediction value (covariate) for the same site provided. In other words, ANUSPLIN handled a ***missing covariate value*** of -99 by not using any of the data for that site. This situation arose where subdaily observations were not collocated with a daily rainfall observation site. As a quick solution to this problem, values of prediction were extracted at the relevant sites from the grid of prediction values previously created as input to LAPGRD and used to replace the missing data in the SPLINA input data. Prediction values based on the subdaily data available at these sites could have been created but considering time constraints this solution was considered satisfactory.

- A ***masking grid*** was used in LAPGRD to confine output grid values to only those over the land. Any grid which had values where required and nodata values (eg -999.0) over the sea could be used.

Appendix 0 contains a set of the ANUSPLIN logs for SPLINA and LAPGRD used for the 12-hour duration index rainfall mapping.

Using these inputs with ANUSPLIN gave reasonable agreement between the direct estimates used as input data the values extracted from the output grids at those sites and gradients that appeared plausible across most of the pilot study area. However, in one area the mapping was judged unrealistic and prompted further investigation.

4.2.2 Sensitivity of Grids to Spurious Data

The grid values of index rainfall in the south-west corner of the pilot study region rose to a peak in an area where there were no direct estimates as input and the maps of predictions, mean annual rainfall and elevation all had a flat gradient.

Suspicion fell on the station 56041 (Bonshaw (Monkstadt)) which had a direct estimate of 10 mm at the 1-hour duration and a nearby station 41430 which had a value of 35 mm. The difference between the two stations caused ANUSPLIN to fit a steep gradient between them, which then ‘overshot’ into the area north of the two sites where there was no modifying real direct estimate data, thus producing a large and spurious region of high values.

Mapping the direct estimates without data from 56041 gave a reasonable result for all durations and it was decided to adopt this approach for the pilot study and to flag 56041 as a problematic data source. Figure 49 shows the two mapping versions with the left panel showing the suspect grid when station 56041 is included and the right panel showing the revised grid when the station is ignored.

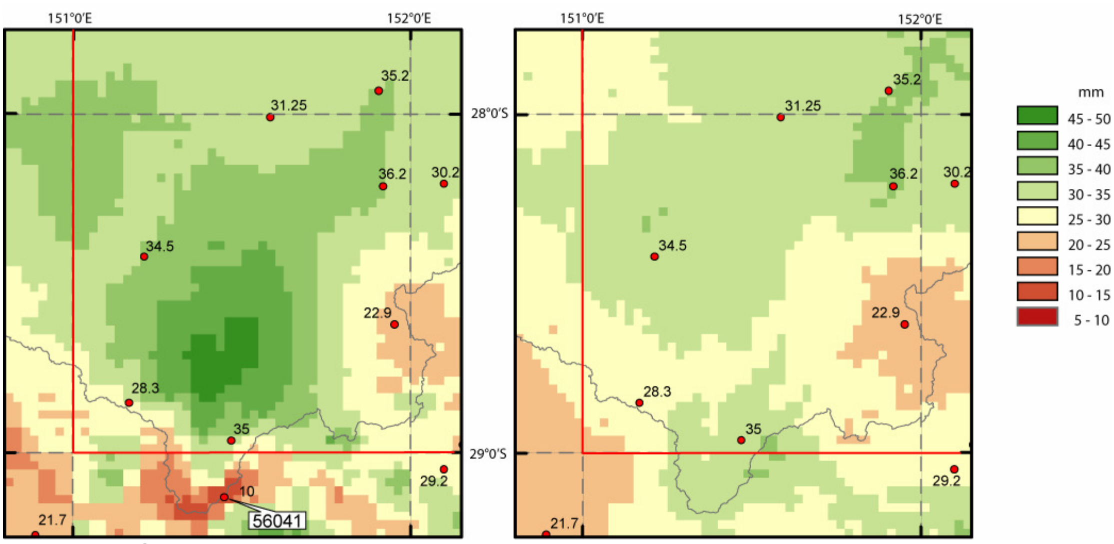


Figure 49 Sensitivity of grid to spurious data for index rainfall for 1 hour duration.

This problem demonstrated two things: firstly, that there is a lack of data in the west of the study area which enabled the spline results to dip or peak unrealistically even though grids of prediction values, mean annual rainfall and elevation data were included in the program for guidance and, secondly, that mapping at subdaily durations is very sensitive to the input data. This reinforced the importance of quality controlling data as discussed in sections 1.4 and 1.5.

4.3 Mapping Growth Factors at Subdaily Durations

Regional GEV parameters were calculated on the basis of regional L-moments. Once the regional GEV parameters had been derived, growth factor estimates, $x(F)$, were calculated according to equation (1) below. (The ArcGIS command version of this equation is shown in Appendix 0.)

In the following $x(F)$ is the growth curve, where F is the non-exceedance probability, 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. Table 9 lists the required values for ARIs from 5 to 100 years.

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

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

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

Table 9 Input to GEV equation for the various ARI values.

ARI	5 years	10 years	20 years	50 years	100 years
T	5	10	20	50	100
F	0.8	0.90	0.95	0.98	0.99

Once satisfactory grids of k and β had been obtained, they were imported into ArcGIS and manipulated within it to generate grids of growth factors and, subsequently, rainfall quantiles. For ARIs less than 10 years, grids were adjusted to a PDS scale using the values given in Table 10 (see HRS10 section 4.2.3).

Table 10 Conversion factors - Annual maximum to partial duration series scale

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

4.3.1 Direct Mapping of Growth Factors at Subdaily Durations

There are two possible approaches to deriving maps of growth factors:

1. Calculation from the estimates of regional GEV parameters (k and β) and subsequent mapping of results to create a grid of growth factors or
2. Mapping of the k and β values and then calculation of growth factors using the resultant grids.

Initially the first approach was applied and a series of twelve trials was run to evaluate the best way to apply ANUSPLIN to map the growth factors.

Regional growth factors were calculated, centred at sites inside the pilot study area itself that met a set of criteria including circle size, number of station-years etc.. This did not provide a coverage that was adequate to generate a spline-fitted surface to points in the pilot study area as there were edge effects. These effects were usually overcome by using additional points in the buffer zone which were calculated using new, less rigorous criteria e.g. fewer station-years of data being required for inclusion as these additional sites in the buffer zone only provide guidance at the edge.

An area in the central west of the pilot study was identified that had no data, so additional regions were included after application of another set of 'weakened' criteria. These additional values filled the gap in the coverage.

One major finding of the trialling was the necessity to include a sufficient number of regions to provide adequate coverage of the area for ANUSPLIN to fit a reasonable surface over it. To address this, two sets of additional regions (circles) with less stringent criteria applied to them were combined with the original set of values in a spreadsheet and, after eliminating duplicate regions by selecting the one with the best quality data, a total of 693 regions were used in the mapping.

One of the options trialled in an attempt to improve the results was the use of the combined data set as input to a grid used as the dependent covariate when mapping just the better quality pilot study area sites. However, it was found that this approach did not improve the results.

4.3.2 Mapping k and β at subdaily durations

It was decided to map the regional GEV parameters k and β rather than mapping the growth factors directly. The gridded values were used in ArcGIS to calculate growth factors, and then rainfall quantiles, to allow the consistency across durations to be assessed.

The concept of using values from locations which did not meet the most stringent criteria but which were located in 'gaps' in the coverage, as discussed in section 4.3.1, was continued and three types of input were used as discussed below and shown in Figure 50.

The criteria which determined the quality of the input data to the k and β calculations for a circle centred on a site were,

A. at least 200 station-years of actual sub-daily data were required (Figure 50, red dots),
B. at least 200 station-years of data required, with prediction data allowed to be included along with actual sub-daily data (Figure 50, orange dots) and
C. for the buffer zone only, no minimum station-years of data were required, all prediction data and actual sub-daily data (if any) within a 40 km radius were used (Figure 50, blue dots).

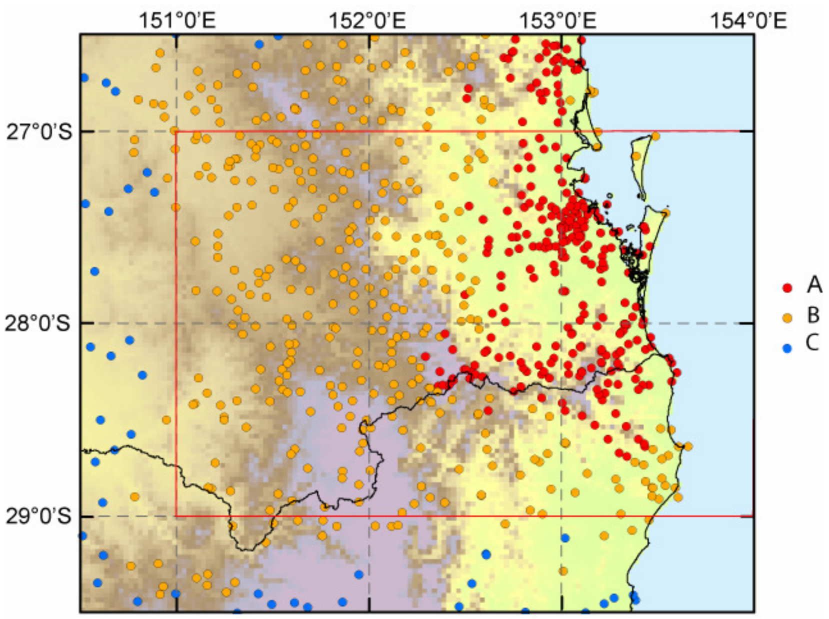


Figure 50 Location of sites for which criterion A, B or C was used in calculating k and β.

4.3.3 ANUSPLIN Options

The chosen set-up for ANUSPLIN to map the k and β parameters used the following programs and inputs:

- Programs used were **SPLINA**, the simpler of the splining options as this is suitable for data sets with up to 2000 points, and **LAPGRD** which is the program used for generating grids. The output grids were in .asc format suitable for reading directly into ArcGIS and at a grid spacing of 0.025°.

- **Longitude and latitude** were used as independent variables within the limits 150.5°E to 154.0°E and 26.5°S to 29.5°S.

- No transformation was applied for the dependent variable.

- A **Spline order** was chosen as 2.

- The smoothing directive was set to obtain a **fixed signal to noise ratio** with a ratio of 3:1.

- A masking grid was used with LAPGRD to confine output grid values to only those over land.

Appendix 0 contains a set of the ANUSPLIN command files for SPLINA and LAPGRD used for the mapping.

4.3.4 Choice of Weights for Combining Predicted and Direct Estimates of Regional GEV Parameters

All the work on growth factors thus far had assumed a weighting of 0.6 for the predictions in relation to actual data, refer to the discussion on Weighting Scheme in section 2.4.

The k and β factors were recalculated for 9 representative durations and ARIs, using the weights of 0.2, 0.4 and 0.6 for the prediction values. As k and β are difficult to interpret, they were used to generate, firstly, the growth factors and, secondly, the quantiles. Grids of the quantile differences were also derived.

As shown in Figure 50, the inputs to evaluating the k and β at a site had to meet one of 3 sets of criteria. Those sites meeting set A did not require any prediction data as input and, therefore, the sensitivity evaluation was concentrated on the non-costal regions that used the B and C criteria. A prediction weighting of 0.4 was adopted after visual inspection of the resultant quantile and quantile difference maps.

4.4 Mapping Index Rainfall at Daily Durations

Based on mapping of the subdaily index rainfall, a small series of trials was undertaken to choose the ANUSPLIN inputs which produced the best compromise between reproducing the values at the input data sites and creating a smooth map.

Daily index rainfalls (medians) were derived from the annual maximum series. Estimates of index rainfall were available for 24-, 48- and 72-hour durations at 747 daily rainfall sites in the study area (see Figure 3). These data were considered sufficiently well distributed and dense enough to map without the need for the additional inputs used to support the 88 subdaily direct estimate sites.

Values of the grids produced for the trials were extracted at the circle centres and compared against the original values in a spreadsheet. At the same time, the maps were checked to assess whether a smooth gradient of values across the study region was being maintained.

4.4.1 ANUSPLIN Options

The ANUSPLIN setup chosen to map the daily annual maximum series data at an ARI of 2 years and produce grids of index rainfall used the following programs and inputs:

- The programs used were **SPLINA**, the simpler of the splining options as this is suitable for data sets of up to 2000 points, and **LAPGRD** which is the program used for generating grids. The output grids were in .asc format suitable for reading directly into ArcGIS and at a grid spacing of 0.025°.

- **Longitude and latitude** were used as independent variables within the limits 150.5°E to 154.0°E and 26.5°S to 29.5°S.

- A square root transformation was applied to the dependent variable, that is, the direct rainfall estimate. This is recommended in the 'Dependent Variable Transformations' section of the 'ANUSPLIN version 4.37 User Guide' (M.F. Hutchinson, 2007). The dependent variable has the transformation applied to it in SPLINA and it is back-transformed in LAPGRD when the grid of values is generated.

- The **order of spline** was set to 3.

- The smoothing directive was set to obtain a **fixed signal to noise ratio**. This option was pursued after a recommendation from Craig Thompson of the National Institute of Water and Atmospheric Research, New Zealand. Trials suggested that a ratio of 4:1 worked best.

- A masking grid was used with LAPGRD to confine output grid values to only those over the land. Any grid which had values where required and 'nodata' values (e.g. -999.0) over the sea could be used.

Appendix 0 contains sample input data and ANUSPLIN command files for SPLINA and LAPGRD as used for the 24-, 48- and 72-hour duration daily index rainfall mapping.

4.5 Mapping k and β at Daily Durations

A similar approach to mapping the subdaily k and β values was used to map the daily values. Regional estimates were available at 747 locations.

4.5.1 ANUSPLIN Options

The chosen set-up for ANUSPLIN to map the k and β parameters used the following programs and inputs:

- The programs used were **SPLINA**, the simpler of the splining options as this is suitable for data sets of up to 2000 points, and **LAPGRD** which is the program used for generating grids. The output grids were in .asc format suitable for reading directly into ArcGIS and at a grid spacing of 0.025°.

- **Longitude and latitude** were used as independent variables within the limits of 150.5°E to 154.0°E and 26.5°S to 29.5°S.

- No transformation of the dependent variable was applied.

- The **Spline order** was set to 2.

- The smoothing directive was set to obtain a fixed signal to noise ratio with a ratio of 3:1.

- A masking grid was used in LAPGRD to confine output grid values to only those over the land.

Appendix 0 contains a set of the ANUSPLIN command files for SPLINA and LAPGRD as used for the mapping.

4.6 Deriving Estimates for ARI of 1 Year

HRS10 lists recommended adjustment factors to convert from an annual maximum scale to a partial duration series scale for ARIs of 2, 5 and 10 years (Table 10); no adjustment is required for higher ARI. These conversion factors were applied once quantile estimates had been derived. ARR87 used a similar approach but different conversion factors (based on Miller 1973) which is based on an approximation and cannot be used to derive an adjustment factor to convert an ARI of 1 year from an annual maximum scale to a partial duration scale.

To address this limitation, a special step was included in the ARR87 procedures (step 6, page 25) to derive estimates for an ARI of 1 year. These estimates were derived based on ARIs of 2 and 50 years for the duration being considered using an equation with 3 different constants. Although the precise derivation of this equation is not clear, it may have been based on a regression approach.

The UK ‘Flood Estimation Handbook’ (Faulkner, 1999) does not explicitly cover an ARI of 1 year in the text book except to explain that it can not be derived from annual maxima. However, it is possible to derive estimates for an ARI of 1 year using the accompanying software but only because a regression model had been built.

NOAA (Hydrometeorological Design Studies Center, US) supply estimates based on annual maxima and starting at an ARI of 2 years through their webpages as shown in Figure 51. In the accompanying documentation, there is a section on deriving 1-year estimates (see Figure 52, available from this webpage: http://hdsc.nws.noaa.gov/hdsc/pfds/docs/NA14Vol1_4analysis.pdf). The recommendation is to use the 1.58 year event on an AM scale to represent the 1-year event on a PDS scale..

POINT PRECIPITATION FREQUENCY ESTIMATES FROM NOAA ATLAS 14

Ohio 40.083 N 81.607 W 797 feet

from "Precipitation-Frequency Atlas of the United States" NOAA Atlas 14, Volume 2, Version 3
G.M. Bonnin, D. Martin, B. Lin, T. Parzybok, M.Yekta, and D. Riley
NOAA, National Weather Service, Silver Spring, Maryland, 2004
Extracted: Wed Jul 23 2008

Confidence Limits | Seasonality | Location Maps | Other Info. | GIS data | Maps | Help | Docs | U.S. Map

Precipitation Frequency Estimates (inches)

AEP* (1-in-Y)	5 min	10 min	15 min	30 min	60 min	120 min	3 hr	6 hr	12 hr	24 hr	48 hr	4 day	7 day	10 day	20 day	30 day	45 day	60 day
2	0.36	0.57	0.69	0.92	1.13	1.33	1.41	1.67	1.96	2.29	2.66	3.02	3.65	4.19	5.82	7.28	9.29	11.16
5	0.47	0.72	0.89	1.22	1.53	1.81	1.92	2.24	2.61	3.03	3.47	3.91	4.69	5.32	7.28	9.01	11.33	13.50
10	0.53	0.82	1.01	1.41	1.79	2.15	2.28	2.67	3.10	3.58	4.08	4.54	5.43	6.10	8.24	10.12	12.60	14.92
25	0.62	0.95	1.17	1.65	2.14	2.64	2.81	3.29	3.81	4.38	4.94	5.42	6.42	7.14	9.45	11.48	14.12	16.57
50	0.68	1.03	1.27	1.82	2.40	3.03	3.23	3.79	4.40	5.04	5.64	6.11	7.22	7.93	10.35	12.46	15.18	17.70
100	0.74	1.11	1.38	2.00	2.67	3.45	3.69	4.35	5.05	5.77	6.40	6.86	8.05	8.76	11.23	13.42	16.20	18.75
200	0.80	1.19	1.49	2.17	2.94	3.90	4.19	4.95	5.76	6.57	7.23	7.65	8.94	9.61	12.11	14.35	17.16	19.73
500	0.88	1.30	1.62	2.40	3.31	4.55	4.91	5.84	6.83	7.77	8.46	8.78	10.18	10.77	13.26	15.55	18.36	20.91
1000	0.94	1.37	1.71	2.57	3.61	5.09	5.52	6.60	7.74	8.79	9.48	9.72	11.20	11.70	14.12	16.43	19.21	21.74

* These precipitation frequency estimates are based on an annual maxima series. AEP is the Annual Exceedance Probability.
Please refer to the documentation for more information. NOTE: Formatting forces estimates near zero to appear as zero.

Figure 51 Example for range of ARI covered in HDSC.

4.6.2. 1-year computation

The 1-year average recurrence interval (ARI) precipitation frequency estimates were computed for this project. ARI is the average period between exceedances (at a particular location and duration) and is associated with the partial duration series (PDS). Annual exceedance probability (AEP) is the probability that a particular level of rainfall will be exceeded in any particular year (at a particular location and duration) and is derived using the annual maximum series (AMS). An AEP depth or intensity may be exceeded once or more than once in a year. (Section 3.2 provides additional discussion on this topic.)

A 1-year AEP estimate, associated with AMS, has little meaning statistically or physically. However, the 1-year ARI, associated with PDS does have meaning and is used in several practical applications. The equation $T_{PDS} = [\ln(\frac{T_{AMS}}{T_{AMS}-1})]^{-1}$ (Chow et al., 1988), which is distribution free, provided a mathematical base for converting between frequencies for the AMS data and the PDS data. Here, T_{AMS} and T_{PDS} stand for the frequency associated with the AMS data and the frequency associated with the PDS data, respectively. The equation can be transformed into the following:

$$T_{AMS} = \frac{1}{1 - e^{-\frac{1}{T_{PDS}}}}$$

Therefore, T_{AMS} = 1.58-year when T_{PDS} = 1-year from the equation. This means that a PDS 1-year event is equivalent to an AMS 1.58-year event. This relationship was used to calculate the 1-year ARI from AMS data for this project. Appendix A.9 provides the regional growth factors computed for the 1.58-year AMS results. However, for all ARIs other than 1-year, the results were obtained by analyzing both AMS and PDS data separately, averaging ratios of PDS to AMS quantiles and then applying the average ratio to the AMS results (Section 4.6.4).

Figure 52 Description of computing ARI of 1 year (source: http://hdsc.nws.noaa.gov/hdsc/pfds/docs/NA14Vol1_4analysis.pdf)

5 SMOOTHING ESTIMATES

For ARR87, IFD curves were smoothed across durations using a 6th order polynomial. In addition, a special scale was used to plot IFD curves. This was most likely done to make them as close to straight lines as possible to facilitate the application of the graphical approach. It is assumed that this is no longer a relevant constraint.

Smoothing might be expected to assist in ensuring consistency across durations, reducing the need to correct such inconsistencies later on. While smoothing would reduce some of the inherent 'noise' in estimates, which is a desirable result, it might also reduce some discontinuities (unevenly spaced differences in design rainfall estimates at neighbouring durations) which are due to the methodology - in particular the use of prescribed durations rather than natural event length.

In the following sections, the terms 'order' and 'degree' of polynomial are used synonymously to mean the highest order power of a polynomial.

5.1 Deriving Smoothed IFD Curves

Figure 53 shows plots of log-transformed rainfall depths (in mm) against log-transformed durations (ranging from 6 minutes to 72 hours) for an ARI of 100 years for the example station of 2012. Each of the nine plots shows fitted polynomials for order ranging from 2 to 10. Similar investigations could be undertaken for rainfall intensity.

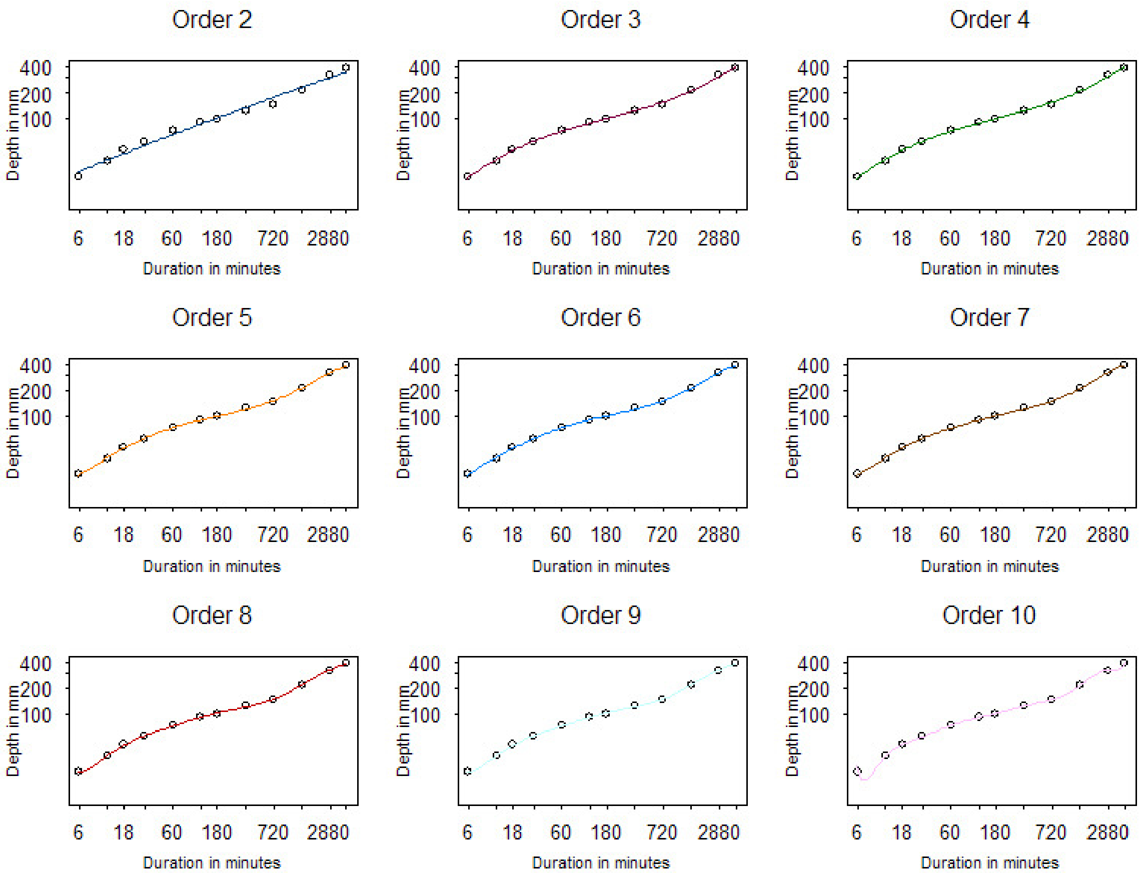


Figure 53 Example of fitting polynomials to estimates of rainfall depths with an ARI of 100 years at station number 2012.

Investigations were undertaken to answer the following questions:

- Should the IFD curves be smoothed?
- If so, which approach should be used?
- Considering that 6th order polynomials were used in the development of ARR87 estimates, is there an optimum order of polynomial?

From Figure 53 some of the criteria for selecting the order of polynomial to fit become apparent:

1) The fitted curve should show little deviation from the straight line connecting neighbouring points. For higher order polynomials this criterion was not always satisfied (for example fit for order 10 between 6 and 12 minutes in Figure 53).

2) The fitted curve should match the data points. For lower orders, this criterion was not satisfied. For testing, thresholds were defined for the maximum and average residuals (point versus fitted curve).

Polynomials up to order 10 were investigated. Splus functions *poly* and *poly.transform* were used to fit the polynomials. (Polynomials were originally derived to be orthogonal and needed to be transformed to the conventional form.)

$y = \sum_{k=0}^{n} \beta_k x^k + \varepsilon$ where n is the degree (or order) of the polynomial

Splus allows the inspection of the statistical significance of the polynomial of a given order. It was found that p-values did not decrease monotonically. Judging from this plot, there was little advantage in using polynomials of order higher than 7.

Using criterion 1 for each station, each ARI and orders from 1 to 10; the maximum difference between the fitted curve and straight lines connecting points was calculated. These values were then averaged over ARI to assess the typical difference for a given order polynomial at a site. Histograms were prepared showing the frequency of the order of polynomial associated with the minimum difference at a station.

In less than 5 cases did fitting the lower orders (1 to 4) result in the minimum differences (measured against straight lines): likewise, order 10 appeared unsuitable. Polynomials of order 5 to 9 did merit consideration, however, order 6 and 7 were judged most suitable and finally order 6 was selected because it is of the lower order.

Figure 54 and Figure 55 suggested that degree 6 would be a plausible choice, pending an assessment of maximum and typical differences when using the smoothed curves to derive estimates (criterion 2).

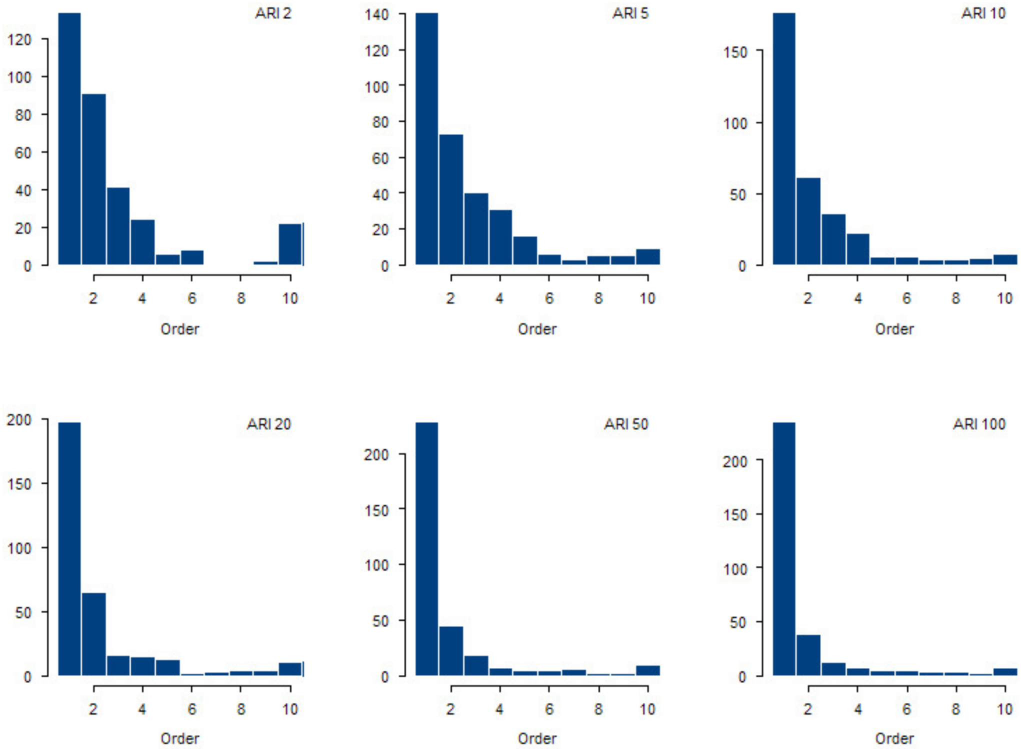


Figure 54 Histograms of counts of fitted polynomials with p-value < 0.02 (statistically significant) for 329 stations and for ARI from 2 to 100 years.

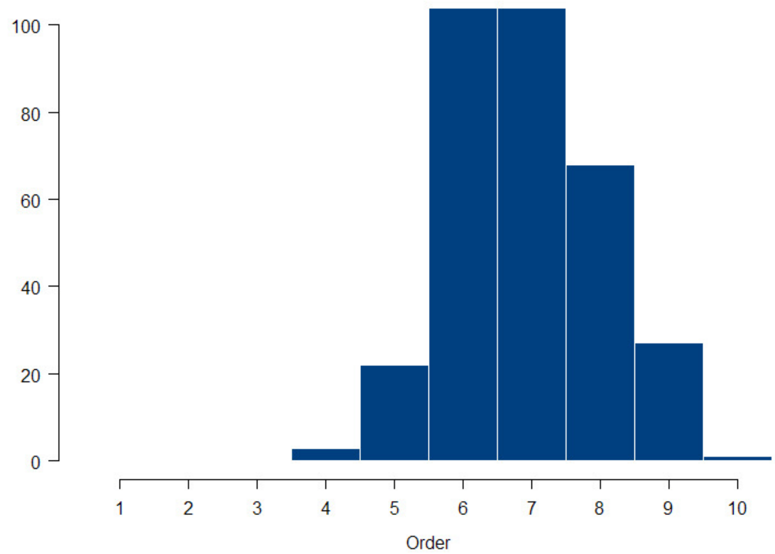
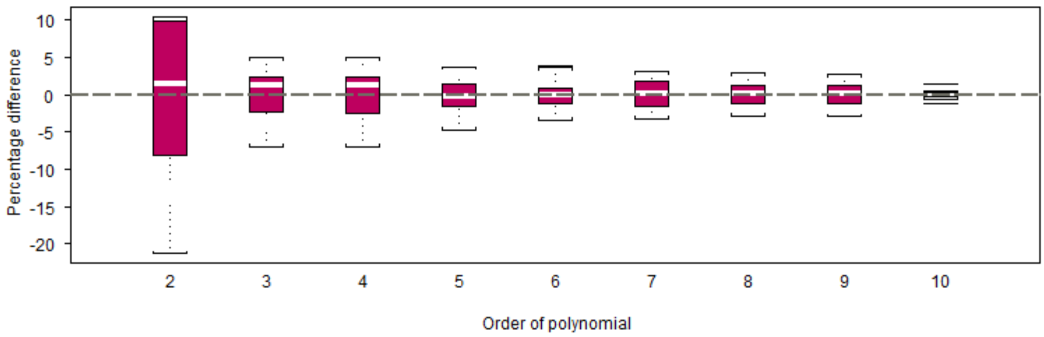


Figure 55 Histogram of counts of degree of polynomial showing frequency of minimum deviation from straight lines (based on all stations and ARIs).

Differences between the original and the smoothed values for station 2012 for an ARI of 100 years are shown in Figure 56. The top row shows boxplots of residuals for degrees of fitted polynomials of order 2 to 10 averaged over all durations. The bottom row shows the average percentage differences. The average error dropped to below 2% for polynomials of degree 5 and above (Figure 56, bottom row); the maximum absolute error was less than 5% for degrees of 5 and above (Figure 56, top row). Considering all 329 stations and all durations (6 minutes to 72 hours), typical differences were less than 2%. However, differences were larger for an ARI of 100 years (Figure 57). Horizontal lines indicate 1, 2 and 3% error. Red dots stand for 6th degree polynomial.



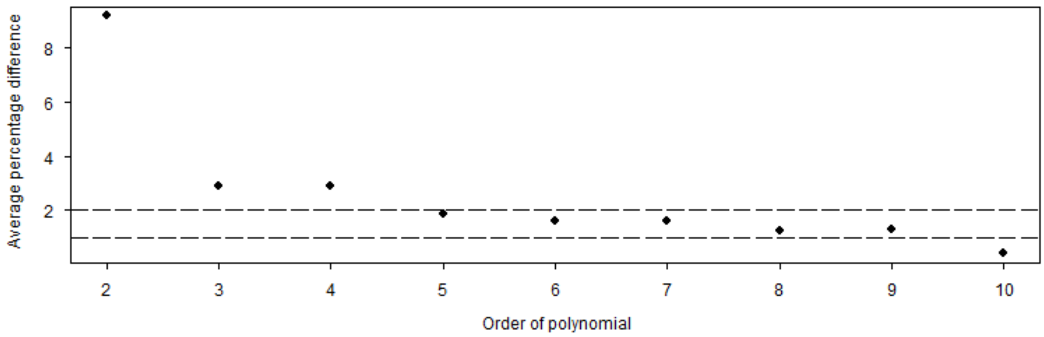


Figure 56 Boxplots of percentage difference between rainfall depth estimates and fitted curve (top) and average percentage difference (bottom) for station 2012 and an ARI of 100 years for durations from 6 minutes to 72 hours.

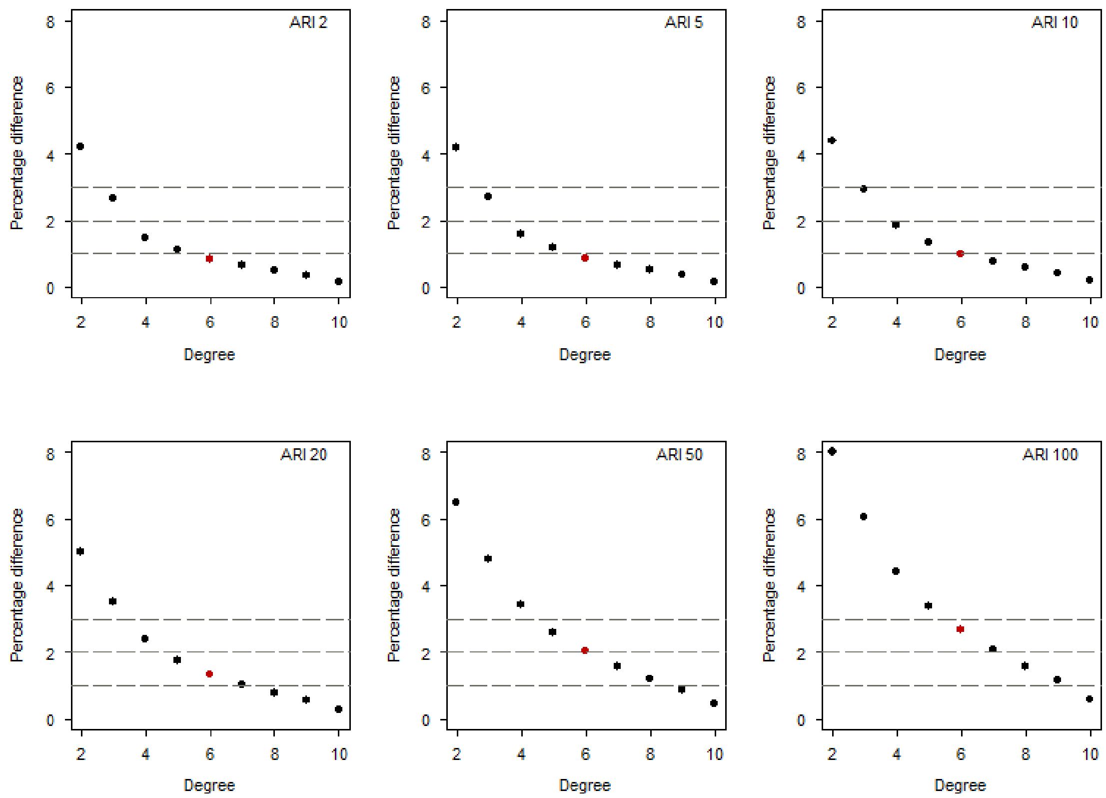


Figure 57 Percentage difference between original values and fitted curve averaged over all stations and durations.

Recommendation

It appears reasonable that users should be able to derive estimates for durations that are not standard durations. If on this basis smoothed IFD curves are required, 6^{th} degree polynomials could be used (as in ARR87). The difference introduced in the smoothing process would typically be less than 2%, which is consistent with the value quoted in ARR87 table 1.1.

On the basis of this analysis, a decision was made to smooth the preliminary quantile estimates for a given ARI across durations. Initial investigations were undertaken in Splus, however, for computational speed, the approach was implemented in Fortran. Tests undertaken showed good agreement between curves fitted using an Splus code and Fortran code respectively.

5.2 Fitting Polynomials - Excluding Very Short Durations

The set of durations over which the polynomial is fitted is likely to affect the fit of the polynomial. Maps of quantile estimates for durations shorter than one hour (to include durations of 6, 12, 18 and 30 minutes) were not developed. It was therefore not possible to include these durations in the smoothing procedure. Analyses were undertaken to test what effect the choice of starting duration had on the fitted polynomial.

For 329 pluviographs, ARI estimates for the various durations (from 6 minutes to 72 hours) were derived based on at-site data only. Previous analyses showed that polynomials of order 6 were appropriate for deriving smoothed IFD when considering all eleven durations, however, it was considered likely that for fewer durations polynomials of lower order might be appropriate. This question was not assessed. Instead, it was decided to explore differences between the two fitted polynomials (across stations, durations and ARI). Figure 58 shows boxplots (one per ARI, starting at an ARI of 2 years) for these differences. For an ARI of 100 years, differences reached up to 60 mm.

Figure 59 shows results for two stations: station 2012, which had been analysed previously, and station 32064 which was selected because it showed particularly large differences between the two fits. These large differences were at least partially due to the fact that the quantile estimates themselves were high (nearly 1000 mm for the 24-h duration at an ARI of 100 years). The difference between the two smoothed estimates for this station exceeded 5% (at 24 h and an ARI of 100 years). It is assumed that regional estimates are less sensitive to the effects studied here. However, it is expected that analyses will have to be repeated once the full set of durations is available. That is after deriving initial estimates, smoothed IFD estimates would have to be derived which in turn would have to be adjusted to address any remaining inconsistencies.

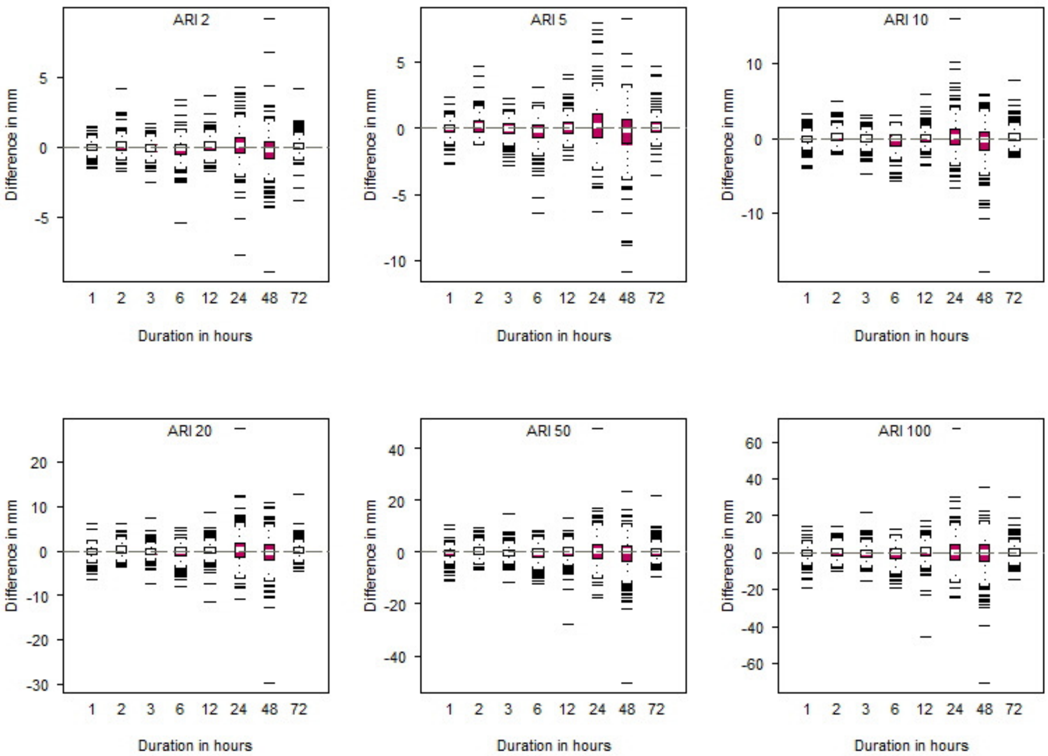


Figure 58 Differences between polynomials fitted to quantile estimates (site data) for the full set of 11 durations (starting at 6 minutes) and for 8 durations (starting at 1 hour) respectively.

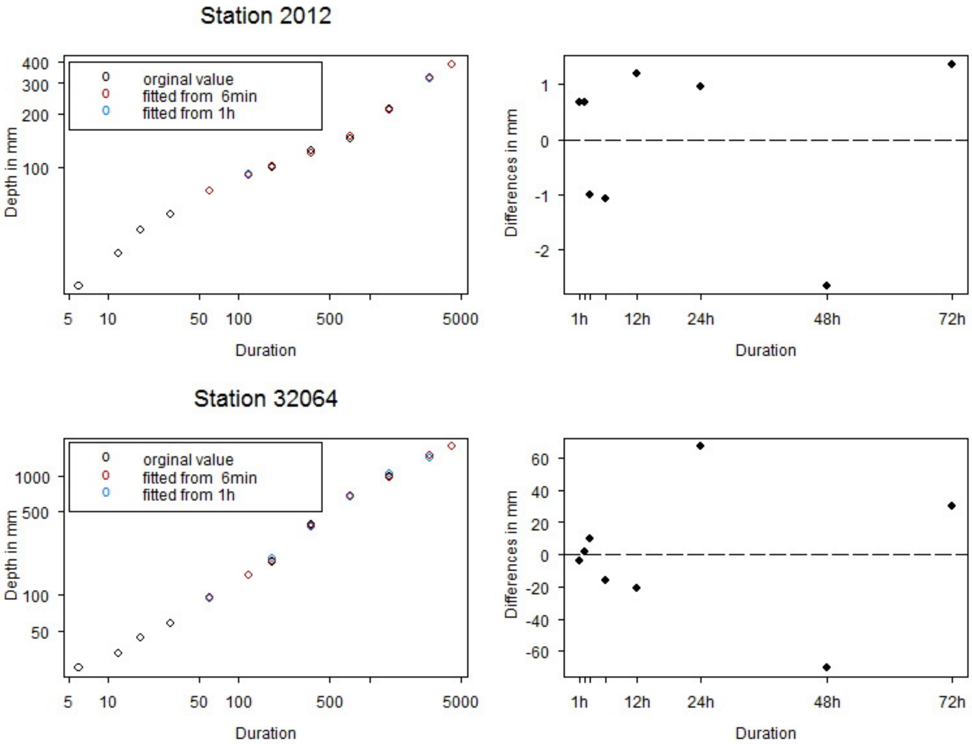


Figure 59 Examples for differences in fit depending on starting duration

Differences in original and smoothed estimates

Differences between the original and the smoothed estimates were calculated to assess how smoothing affected the estimates. These comparisons were undertaken once for the absolute differences (in mm) and then for percentage differences.

It was found that the percentage differences varied strongly from one duration to the next but were fairly constant across ARIs. Percentage differences were smallest for the two end points (1 hour and 72 hours). Largest differences were found for the 6 and 12 hour durations (up to 10%, as seen in Figure 60). In addition to percentage changes, absolute differences were assessed, although these are not shown. The locations where absolute differences were largest tended to be in areas where quantile estimates were particularly high.

There was a tendency for the average percentage differences between original and smoothed estimates to oscillate from one duration to the next with negative values at 1h, positive values at 2h etc. (Figure 60). Investigations revealed that these oscillations were most likely caused by the original data rather than the smoothing procedure.

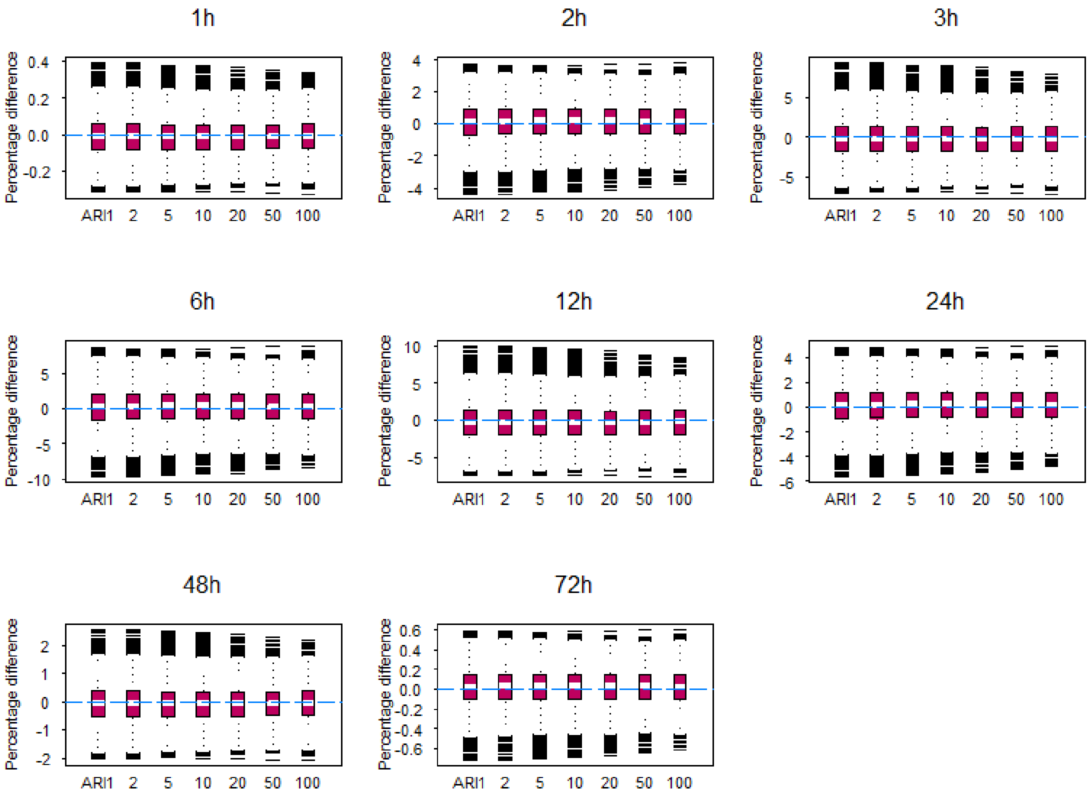


Figure 60 Boxplots showing the percentage difference between original and smoothed estimates. Each panel shows differences for one duration across the range of ARIs.

6 INCONSISTENCIES

Once a full set of grids of daily and subdaily index rainfalls and growth factors had been created, estimates of rainfall depths were derived using ArcGIS by multiplication of the grids (see section 4 'Mapping').

There is a requirement to guarantee the internal consistency of design rainfall estimates across durations and ARI with regards to both rainfall depth and intensity. Inconsistencies could be due to:

- meteorological factors (e.g. the annual maximum series for 3 hours in a region were derived from events that were basically contained within 2 hours),
- the change in method between daily and subdaily analysis or
- the use of adjustment factors (e.g. fixed to sliding durations).

Inconsistencies across ARI could also have been introduced through spatial smoothing in the gridding phase or through smoothing across durations. Although this was deemed unlikely, estimates of rainfall depths were assessed for consistency across ARI (section 6.2). Another question that needed to be addressed was whether the smoothing was likely to help correct inconsistencies or whether smoothing would lead to additional inconsistencies. A ratio of 1.0 or above was considered an inconsistency.

The consistency across durations was examined by dividing the grids of rainfall depths at adjacent durations for a subset of ARIs (5, 20 and 100 years). Figure 61 shows the pattern of inconsistencies, (which were similar for the three ARIs tested) for ARI of 5 years *prior to smoothing* across durations. In Figure 61, the magnitude of inconsistency is indicated by the range of colours. The pale yellow indicated areas where the depth for the shorter duration is lower than at a longer duration and therefore the estimates are considered to be consistent. The region of interest focussed on the pilot study zone (excluding the buffer zone indicated by the thin red lines in Figure 61). Whilst a small number of grid points within it exhibited substantial inconsistencies (estimates at shorter duration exceeded those at the longer duration by up to 50%), the average inconsistency was much lower at about 5%.

6.1 Assessment of Inconsistencies after Smoothing across Durations

Sixth order polynomials were used to smooth quantile estimates across durations. This was achieved using a Fortran program which read grids of original quantile estimates prepared in ArcGIS. The 'original' quantile estimates were compared with the smoothed estimates, excluding those estimates in the buffer zone. Again, ratios of quantile estimates for neighbouring durations were calculated and the results for ARI of 5 years are shown in Figure 62. Differences in spatial extent and magnitude of inconsistencies prior to and after smoothing were particularly apparent between 2 hours and 3 hours, and between 12 hours and 24 hours.

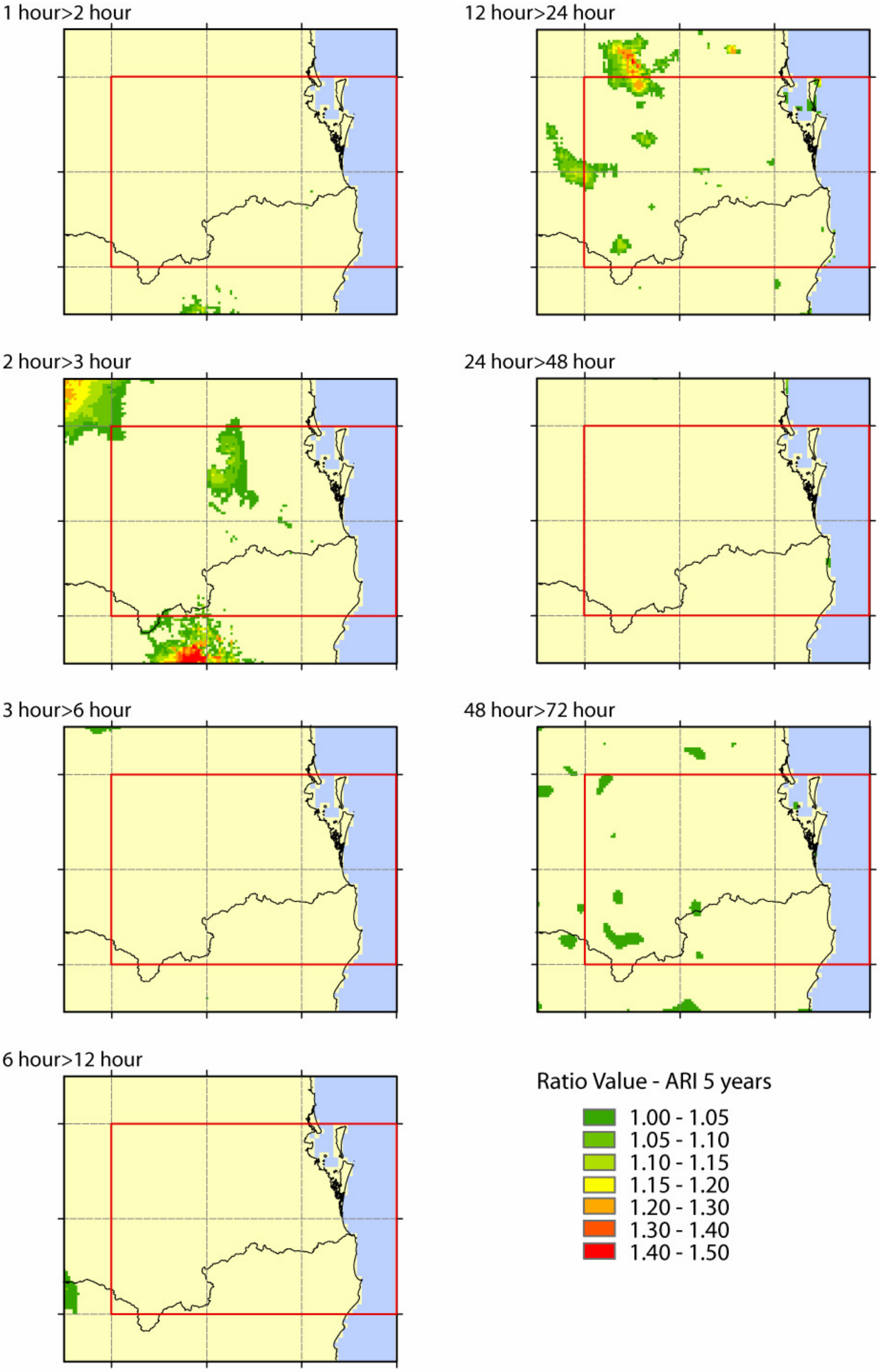


Figure 61 Ratio of estimates of rainfall depths at adjacent durations for ARI of 5 years *prior to smoothing*.

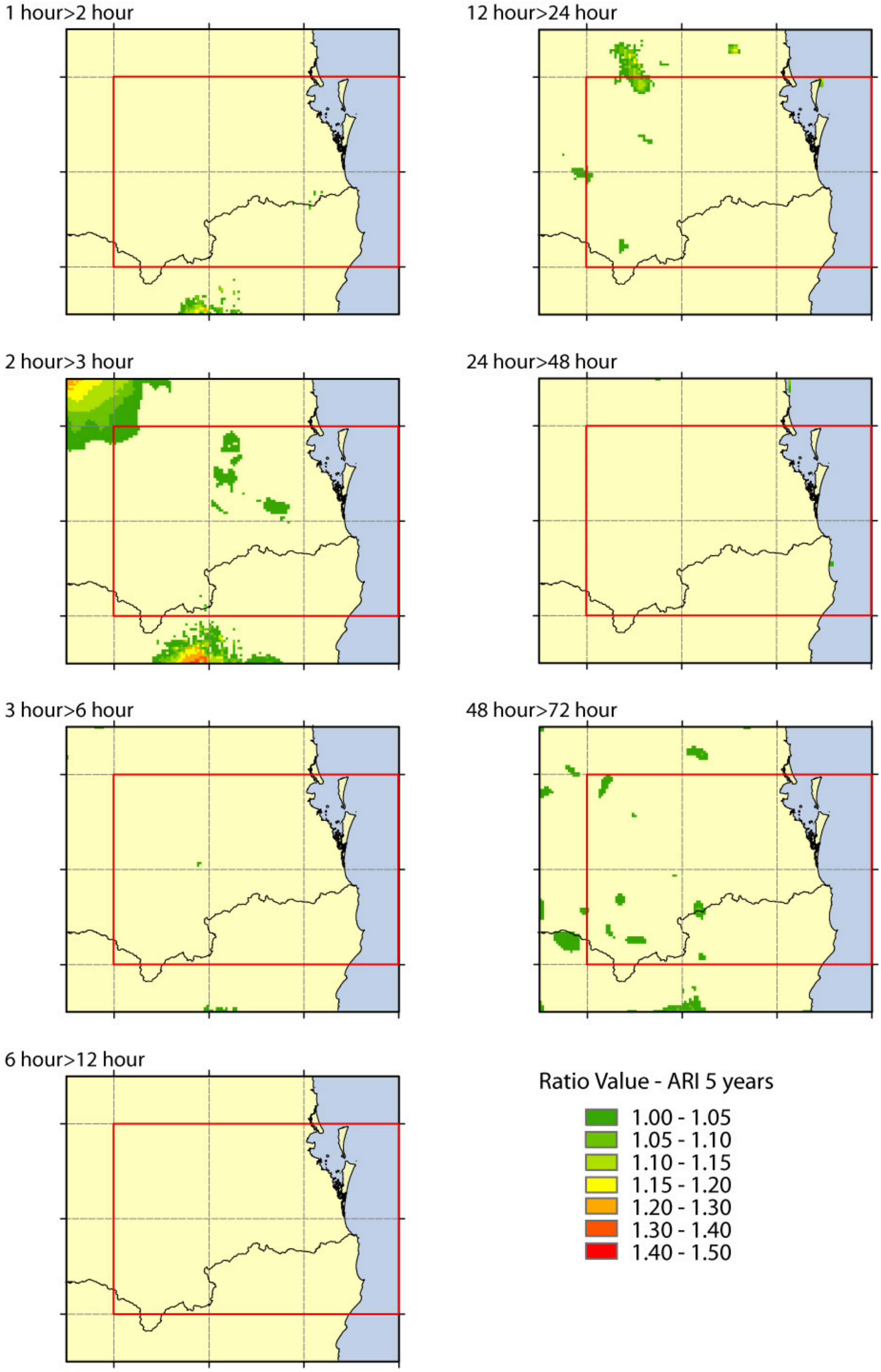


Figure 62 Ratio of estimates of rainfall depths at adjacent durations for ARI of 5 years *after smoothing*.

6.1.1 Effect of Smoothing on Spatial Extent of Inconsistencies

The spatial extent of inconsistencies (in percent of pilot study area, excluding the buffer zone) and magnitude of inconsistencies across the range of durations and ARI were assessed for both the original and smoothed sets of grids on the basis of ratios of depths for adjacent durations; for example, the ratio of depths at 1 and 2 hours.

Figure 63 shows the spatial extent of inconsistencies in estimated design rainfall depths across durations for both the original grids and the smoothed grids. The left panels show the results by ARI while the right panels show them by duration. From Figure 63, it can be seen that the maximum value exceeded 6% for the original grids for an ARI of 1 year, between 12 and 24 hours. However, for some durations, no inconsistencies were found across the whole range of ARIs for both the original and the smoothed grids (for example between the 6 and 12-h duration).

For the original grids, the extent of inconsistencies was generally largest for 2h/3h, followed by 12h/24h. The overall spatial extent of inconsistencies was greatly reduced after smoothing and the improvement was particularly apparent at 12h/24h. Most of the inconsistencies were removed through smoothing, with the following exceptions:

- Smoothing did not substantially reduce the spatial extent of inconsistencies at 48h/72h.
- Although inconsistencies were reduced for 2/3h, significant inconsistencies remain.
- A small number of inconsistencies were introduced for the highest ARI (50 and 100 years) at 3h/6h.

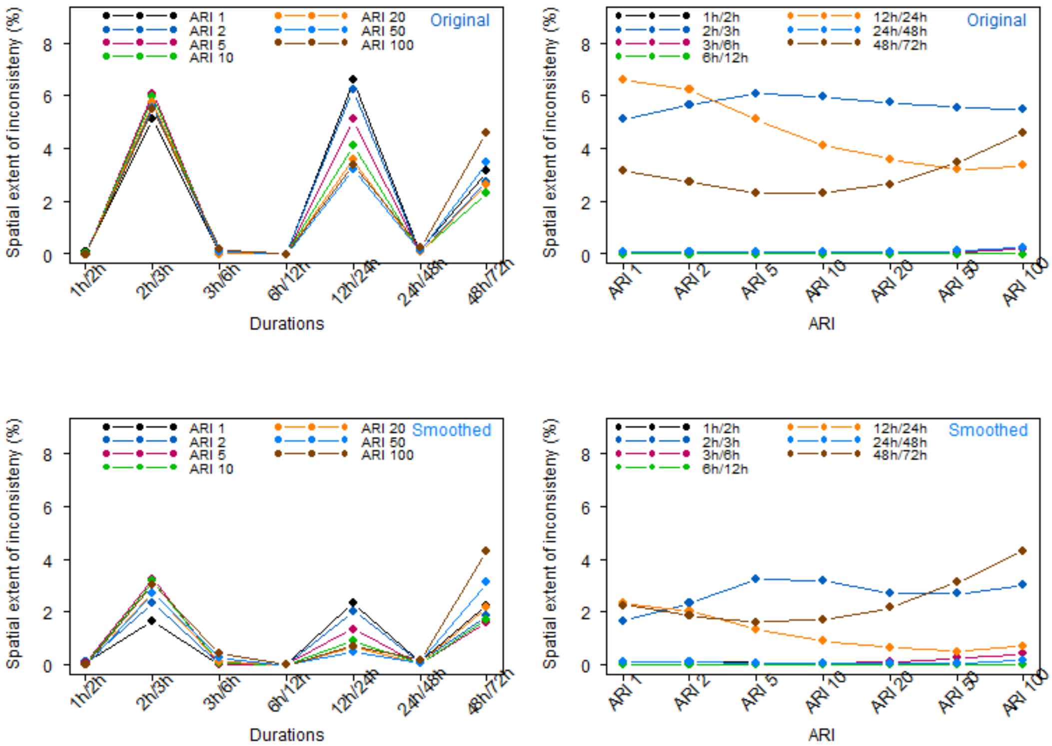


Figure 63 Spatial extent of inconsistencies across durations.

6.1.2 Smoothing and Effect on Magnitude of Inconsistencies

In addition to reducing the spatial extent of inconsistencies, the magnitudes of the inconsistencies were reduced through the smoothing process. Figure 64 and Figure 65 show the magnitudes of inconsistencies before and after smoothing. (Note that different scales have been used for the y-axes.) Since there were no inconsistencies for 6h/12h, this pair of durations is not included in the figures.

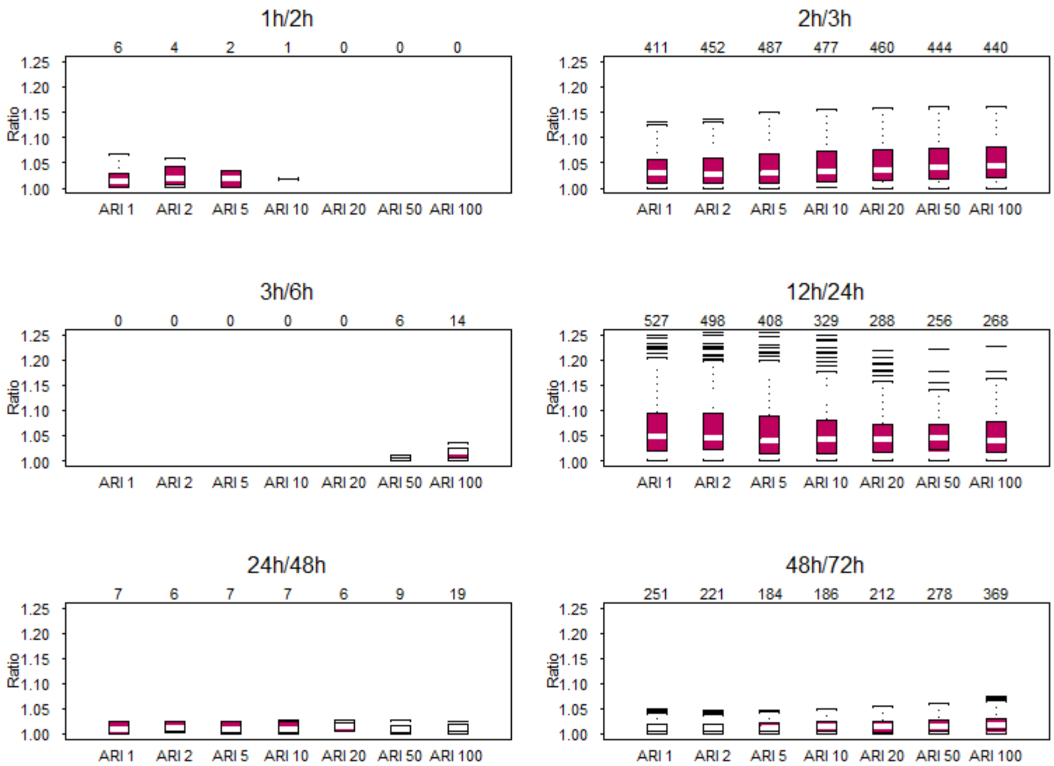


Figure 64 Ratio of rainfall depths where the ratio for adjacent durations is 1.0 or above (inconsistency across durations) for the *original grids*. The numbers below the title for each of the panels are counts of inconsistencies.

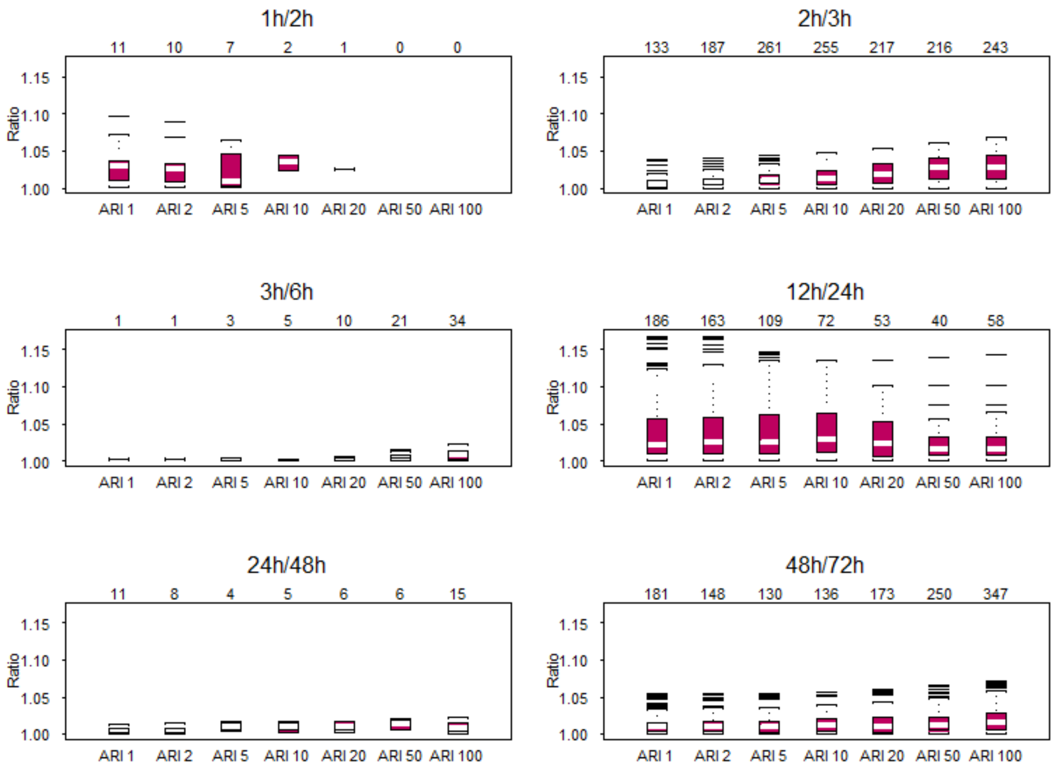


Figure 65 Ratio of rainfall depths where the ratio for adjacent durations is 1.0 or above (inconsistency across durations) for the *smoothed grids*.

Recommendation

Given that the smoothing helped address a large number of inconsistencies in a methodical approach, it is recommended that the data be smoothed, the remaining inconsistencies be corrected, and the corrected data be resmoothed.

Inconsistencies for the 48 to 72 hour duration are most likely due to the fact that rainfall depths for events are similar for the 48 and 72 hour durations. These inconsistencies can easily be removed by increasing the 72 hour depths as required (without introducing conflicts since this is the longest duration).

6.2 Inconsistencies in Intensity Estimates across Durations

Previously, inconsistencies had been assessed on the basis of estimates of rainfall depths. No inconsistencies across ARI had been found and inconsistencies across durations had been significantly reduced after smoothing. Inconsistencies with regards to rainfall depths were defined as

$$D_i / D_{i+1} \geq 1$$

ensuring accumulations at longer durations were larger than at shorter durations. Similarly, it is necessary to ensure that intensities at shorter durations exceed intensities at longer durations. The threshold for inconsistencies was either 2 or 1.5, depending on the durations involved. For instance when assessing intensities at 1 and 2 hours, it is necessary to satisfy $I_{1h} \geq I_{2h}$ which is equivalent to

$$I_{1h} * 60 = D_{1h} \geq D_{2h}/2 = (I_{2h} * 120) / 2 \quad \text{and therefore} \quad D_{1h}/D_{2h} \geq 2.$$

The two checks for consistency (rainfall depths and intensity respectively) are complementary. While comparisons of depths ensure rainfall depths at the longer duration are high enough (to exceed those at shorter durations), comparisons of intensities ensure depths at the longer durations are not too large (causing higher intensities at longer durations than at shorter durations).

With regards to intensity, estimates of rainfall depths would be considered inconsistent where $D_{1h} / D_{2h} < 2$. Table 11 below lists the thresholds against which the intensities were tested, according to duration. No inconsistencies across durations with respect to intensities were found within the pilot study area.

Table 11 Testing for inconsistencies in intensity estimates

Duration i	Duration i+1	Threshold
1h	2h	2
2h	3h	1.5
3h	6h	2
6h	12h	2
12h	24h	2
24h	48h	2
48h	72h	1.5

6.3 Removal of Inconsistencies in Smoothed Grids

Given that smoothing estimates of design rainfall depths across durations helped correct a large number of inconsistencies, it was decided that the remaining inconsistencies would be corrected after smoothing.

An inconsistency was detected where

$$\frac{\text{rainfall depth at the shorter duration}}{\text{rainfall depth at longer duration}} \geq 1.0$$

Inconsistencies were corrected by increasing the depth at the longer duration to lower the ratio to 0.99.

Procedure

A prescribed sequence of steps was developed to remove inconsistencies. An experiment was undertaken to investigate to what extent raising design rainfall estimates at one duration, with the aim of removing inconsistencies, might lead to the introduction of inconsistencies at longer durations.

Inconsistencies at 2 to 3 hours were removed by increasing the depths at 3 hours. It was tested whether such an adjustment might have introduced inconsistencies for the 3h/6h ratios. This test was undertaken for the ARI where the largest adjustment was required but no pre-existing inconsistencies for 3h/6h had been identified. Based on estimates for an ARI of 50 years and by adjusting the ratio to 0.99, it was possible to remove the inconsistencies at 2h/3h without introducing inconsistencies at 3h/6h. This is demonstrated in Figure 66.

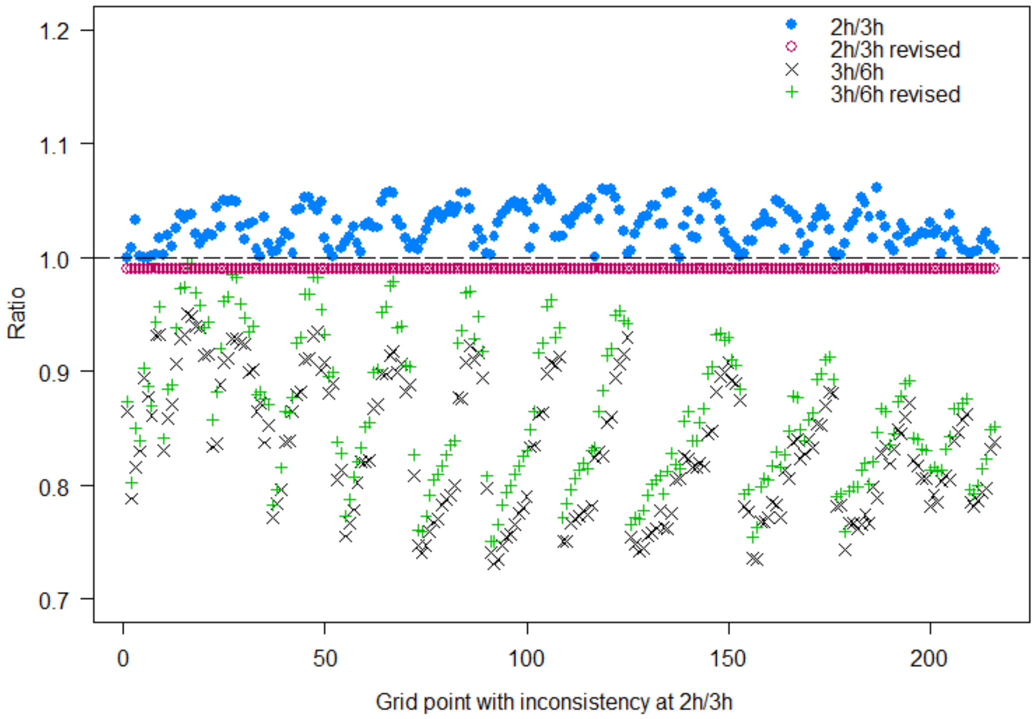


Figure 66 Ratios of design rainfall depths for grid points which exhibit inconsistency from 2h to 3h at ARI50.

The following procedure could be applied to derive grids of smoothed estimates that are consistent across durations:

a. *Remove inconsistencies* by increasing rainfall depth at the longer duration. Test whether this has created (additional) inconsistencies at the next higher duration.

b. *Map grids* (of quantile estimates) after removing inconsistencies to judge whether some spatial artefacts might have been introduced in correcting the inconsistencies.

c. *Smooth resulting grids* (across durations using polynomials) to derive new grids of quantile estimates.

d. *Map grids* (of quantile estimates) after inconsistencies have been removed and smoothing has been done.

e. *Check for inconsistencies* which might have been introduced through smoothing.

Sequence

The majority of inconsistencies occurred at 2h/3h, 12h/24h and 48h/72h. Starting at the shortest durations (2h/3h) and working up to 48h/72h was judged the safest approach since it allowed any inconsistencies created in the process to be addressed.

Preliminary grids of design rainfall estimates had been derived by multiplying grids of growth factors by grids of index rainfall (section 4). The resulting grids were smoothed across durations (section 5) and inconsistencies across durations were removed. Grids were then inspected to judge whether removing inconsistencies had introduced spatial artefacts.

Absolute differences between grids prior to and after removing inconsistencies were largest at the 24-h duration. These differences typically did not exceed 20 mm apart from a small number of grid points on Fraser Island as shown in Figure 67. Locations where inconsistencies needed to be addressed showed some similarities across the range of ARIs. Inspection of the resulting grids of design rainfall depths did not reveal any spatial artefacts introduced in the process of smoothing and removing inconsistencies.

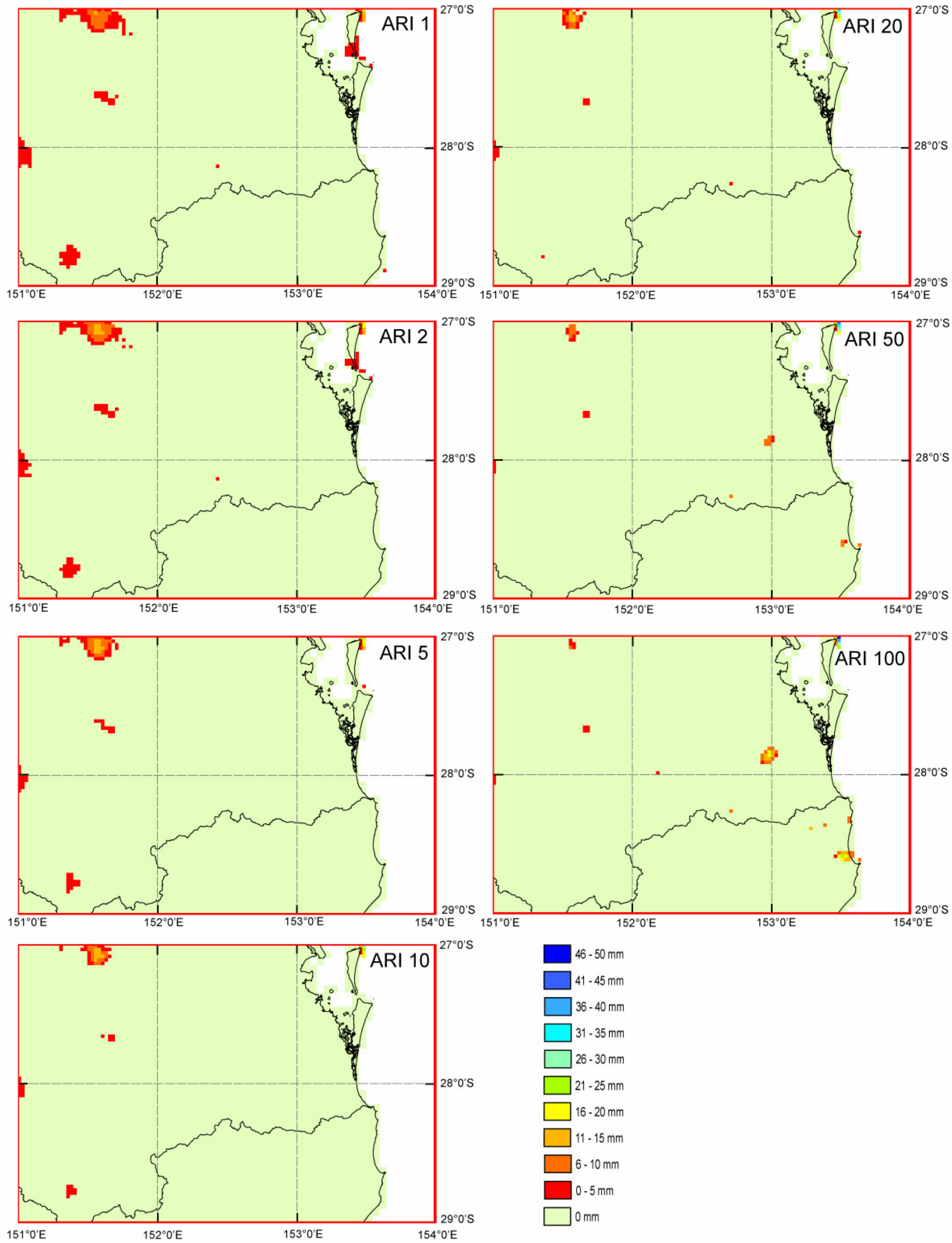


Figure 67 Differences (in mm) between estimates prior to and after iteratively removing inconsistencies across durations. Differences are shown for the 24-h duration.

7 COMPARISON WITH PREVIOUS ESTIMATES

Estimates derived in the pilot study were compared to current estimates, in particular those in ARR87. Such comparisons were considered essential for a number of reasons. Firstly, the magnitude of differences will be a useful guide with regards to the need to issue updated estimates. However, there are other reasons why an update might be sensible, even if the estimates themselves were to change only marginally. These reasons include:

- an extended range of durations and ARIs for which estimates are supplied,
- the ability to supply confidence intervals together with estimates,
- the ability to supply results through an appropriate medium and format (webpages, electronic files, ArcGIS ready) and
- defendable and documented statistical techniques together with up-to-date records.

There are two distinctly different reasons why ARR87 and pilot study estimates might differ: firstly, the methods used to derive estimates (including the gridding procedure) and secondly, the data used in the analyses. It was not possible to establish exactly what data (stations, record lengths) had been used to derive the ARR87 estimates. An overview of some of the important differences between ARR87 and the pilot study is given in Table 12. To emulate the data availability for the development of ARR87, only data up to 1983 were used when undertaking comparisons.

Table 12 Summary of differences in data and methods used in the development of ARR87 and the pilot study

	ARR87	Pilot study
Data	~ up to 1983	Up to 2006
Frequency analysis	Moments, Log-Pearson Type III	L-moments, Generalised Extreme Value distribution
Daily to subdaily durations	Principal Component Analysis	Partial Least Squares Regression
Regionalisation	Fixed, non-overlapping geographical regions	Flexible regions
Mapping	Subjective (meteorological analysis)	Objective (thin-plate spline smoothing)

7.1 Daily Durations

7.1.1 Comparison of Estimates at Daily Durations for Points

At an earlier stage of the pilot study (March 2007), estimates at two locations were assessed across a number of different methods. Some details on this assessment can be found in Xuereb et al (2006). The two stations selected for investigation were Rocky Point Sugar Mill (gauge number 40319) and Mount Tamborine (gauge number 40197). At the first of these two sites, pilot study estimates were consistently higher than estimates derived from the three other methods. Comparison of pilot study estimates against estimates from three other methods showed that pilot study estimates for this site were closest to ARR87 estimates (Figure 68).

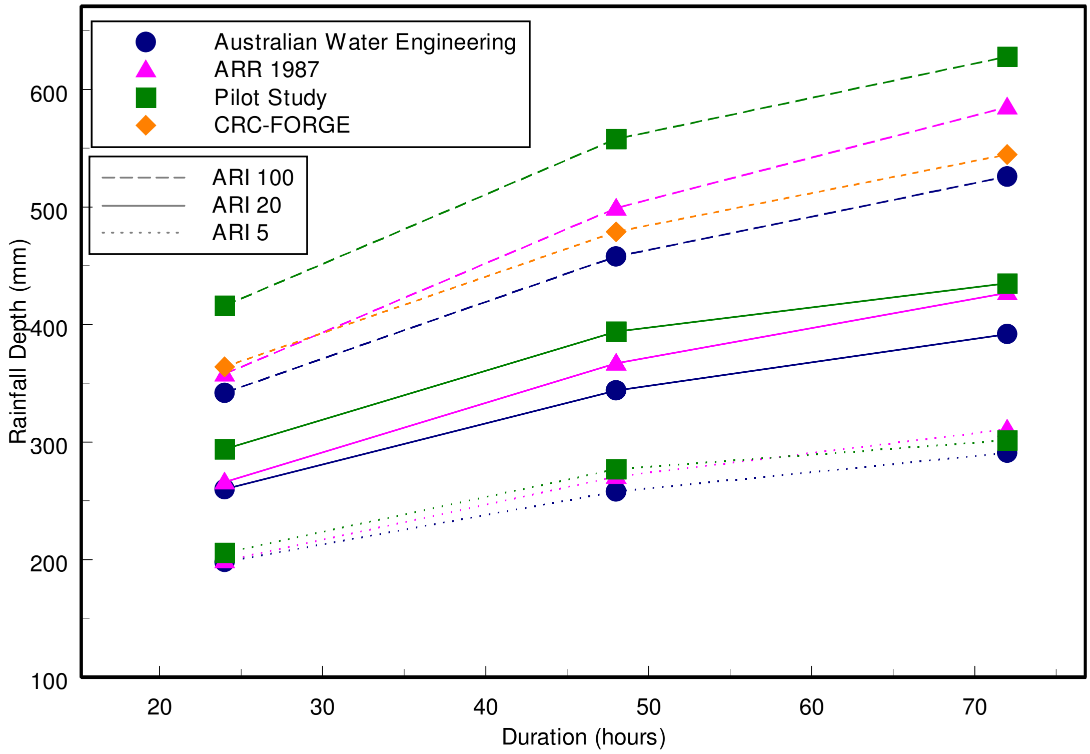


Figure 68 Comparison of 48-hour rainfall depths for Rocky Point Sugar Mill (gauge number 40319) derived using four different methods.

At Mount Tamborine estimates were considerably higher than ARR87 but very close to estimates from two other techniques (Figure 69). Documentation exists that point to the fact that the ARR87 estimates for Mount Tamborine are known to be doubtful (Helen Pearce, pers. comm.). (For the 72-hour duration this station was not included in maps used for deriving ARR87 estimates).

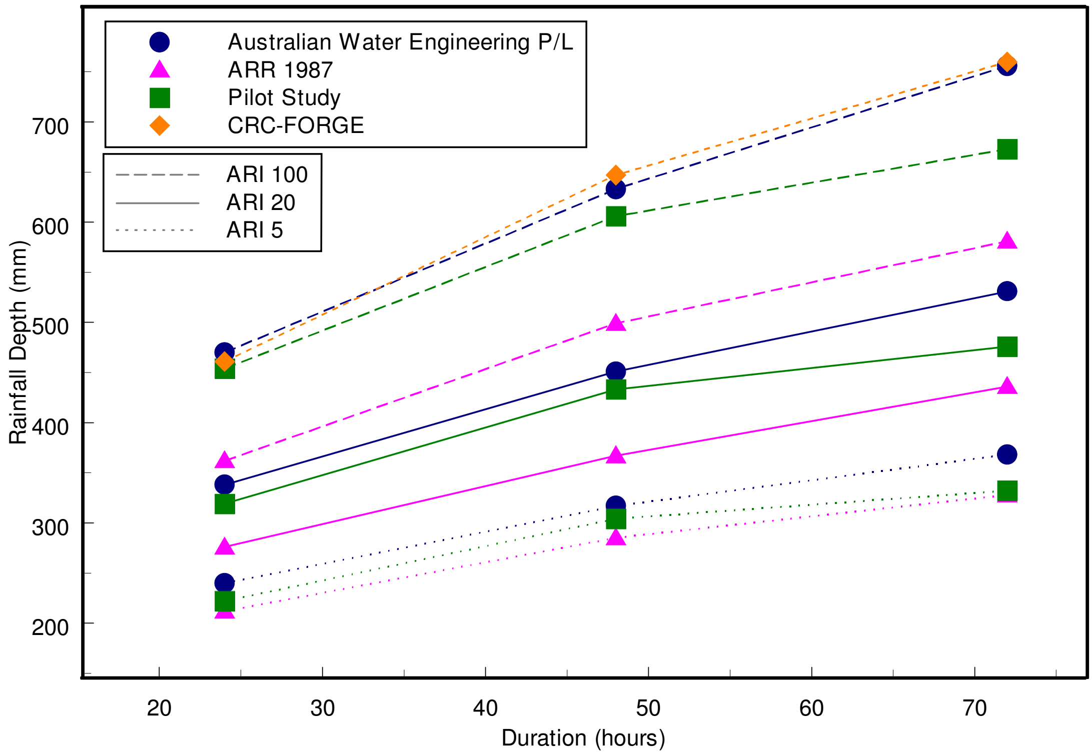


Figure 69 Comparison of 48-hour rainfall depths for Mount Tamborine (gauge number 40197) derived using four different methods.

7.1.2 Comparison of Grids of Estimates at Daily Durations

Comparisons in this section were not based on the final sets of grids. Nonetheless, these comparisons are presented here to show how pilot study estimates compare to estimates derived using two other methods: firstly, ARR87 and secondly, CRCFORGE.

Regional design rainfall estimates had been mapped using the software package ANUSPLIN, which employs the technique of thin-plate spline smoothing (Hutchinson 1998). A manually-drawn digitised 48-hour 20-year master chart was used as an independent variable (together with mean annual rainfall, latitude and longitude) to derive surfaces for durations of 24, 48 and 72 hours and ARIs of 2, 5, 10, 20, 50 and 100 years. The manually-drawn 48-hour 20-year master chart used meteorological experience to interpolate rainfall in areas where no data were available.

Generally, the estimates derived were in good agreement with previously derived estimates. Figure 70 shows percentage differences between pilot study estimates and previously derived estimates for the 48-h duration and an ARI of 100 years. Over large areas, differences in estimates were less than 10%. For a number of areas, the new estimates were lower when compared against ARR87 estimates but not when compared to the more recent CRCFORGE estimates (Nandakumar et al. 1997, Hargraves 2004). Overall there was a tendency for the new estimates to be somewhat higher than previous estimates, including for areas with good station coverage.

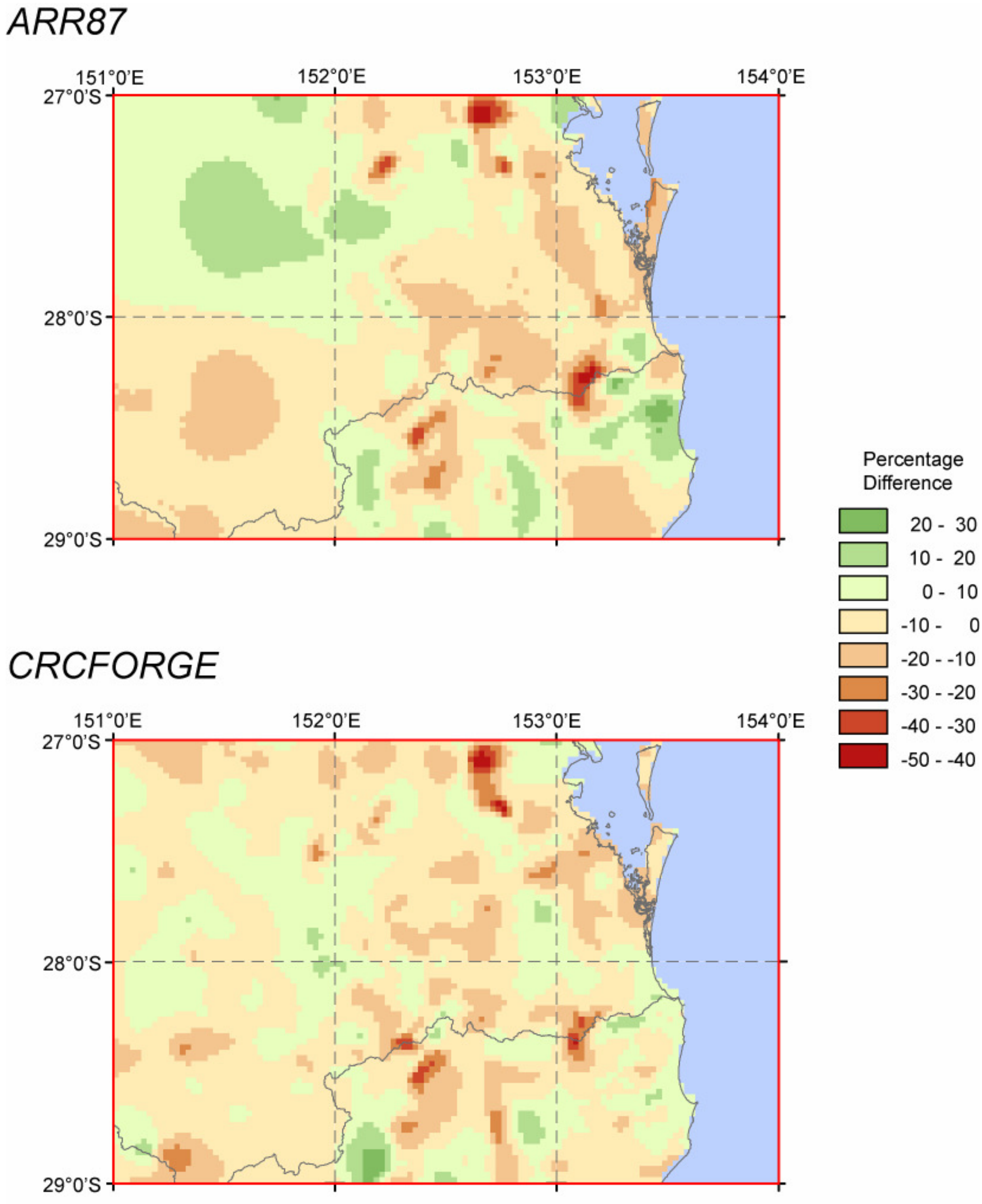


Figure 70 Percentage difference between design rainfall estimates derived for the pilot study, and ARR87 and CRCFORGE respectively.

7.1.3 Effects of Method and Data

A more up-to-date set of grids was used to assess the effect of different methods and data availability on estimates of rainfall depths. The grids used differed in two ways from the final grids prepared in the pilot study:

- For the final set of grids, polynomials were fitted to rainfall depths across subdaily and daily durations (in an attempt to smooth and reduce magnitude and extent of inconsistencies). However, for the comparison, grids had to be derived for a data cut-off in 1983. Preparing such grids for subdaily durations would have required a disproportional amount of time and effort. Since subdaily grids were not available for a cut-off in 1983, it was not possible to derive appropriate smoothed versions of the daily grids.
- For the final pilot study estimates, inconsistencies were removed after smoothing across durations. For the comparisons, inconsistencies between rainfall depths at daily durations (24, 48 and 72 hours) were not corrected. However, the number of inconsistencies (even for unsmoothed grids) between the 24 and 48 hour duration was known to be small (Figure 64) and inconsistencies between the 48 and 72 hour durations were to be removed by increasing the rainfall depth at the 72-hour duration.

For the above reasons, it appeared justifiable to use grids, which had been neither smoothed nor adjusted for inconsistencies, to derive an assessment of how differences in method and data affect estimates. The top panel in Figure 71 (marked ‘A’) shows the approximate percentage difference between pilot study estimates and ARR87 estimates that is due to the different methods while the bottom panel (marked ‘B’) shows the percentage difference due to the different data availability. It is apparent that differences due to methods tend to be larger than those due to data. Green shades in the top panel indicate the pilot study estimates are higher than the ARR87 estimates. Green shades in the bottom panel indicate estimates based on the longer records (up to 2006) are higher than those based on data up to 1983.

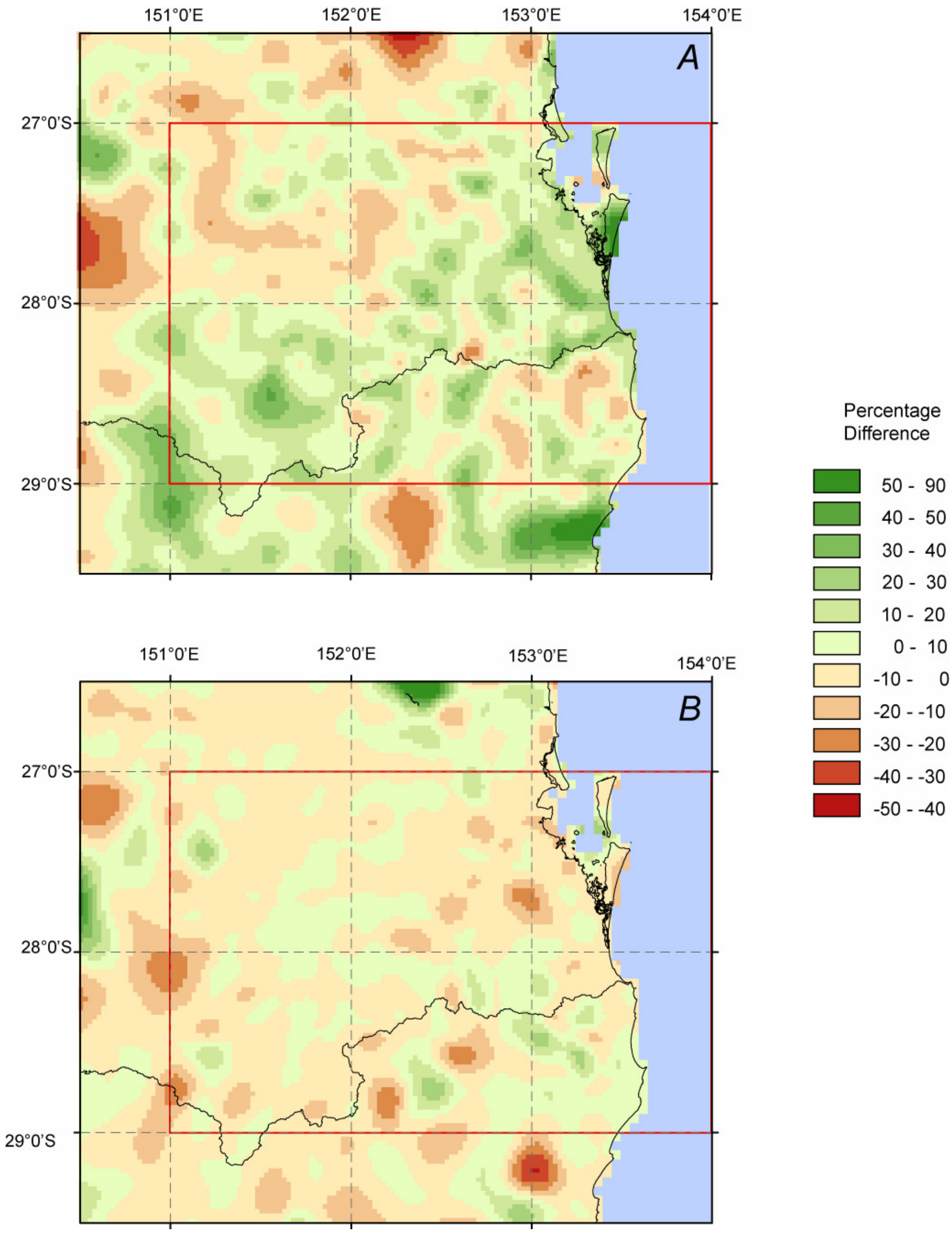


Figure 71 Comparison of pilot study estimates and ARR87 to assess the relative effects of method and data (for 24 hour, ARI 100 years).

7.2 Comparisons for Subdaily Durations

An initial comparison of pilot study estimates and estimates from ARR87 had been undertaken based on site estimates (see section 2.2, Predicting at subdaily durations). The effect of additional data used in the pilot study might be more notable at subdaily durations due to the following reasons:

- For the shorter durations, station density was much lower than for the daily durations and meteorological judgement had been used in developing isopleths for ARR87.
- In addition to the differences in methodology for the daily durations, a different technique was used to infer information about the subdaily from the daily durations.

It has not yet been assessed how sensitive quantile estimates at subdaily durations are to differences in data and methods respectively. However, comparing pilot study estimates with ARR87 might be considered one step in validating pilot study estimates. Where there are significant differences between the two estimates, attempts should be undertaken to use meteorological reasoning to verify the correctness of pilot study estimates.

Percentage differences between ARR87 and pilot study estimates for an ARI of 20 years are shown in Figure 72. Maps are provided to allow a qualitative assessment of differences in estimates. Green (orange) shades indicate areas where pilot study estimates were lower (higher) than the ARR87 estimates.

At sub-daily durations, areas with higher (lower) estimates were typically large and contiguous. For the coastal region, and particularly for intermediate durations (6 and 12 hours), there was a tendency for pilot study estimates to exceed ARR87 estimates. For durations from 1 to 3 hours, there were large contiguous areas in the ‘hinterland’ where pilot study estimates were lower than ARR87 estimates. For durations of 24 hours and above, areas with positive and negative differences were typically much smaller.

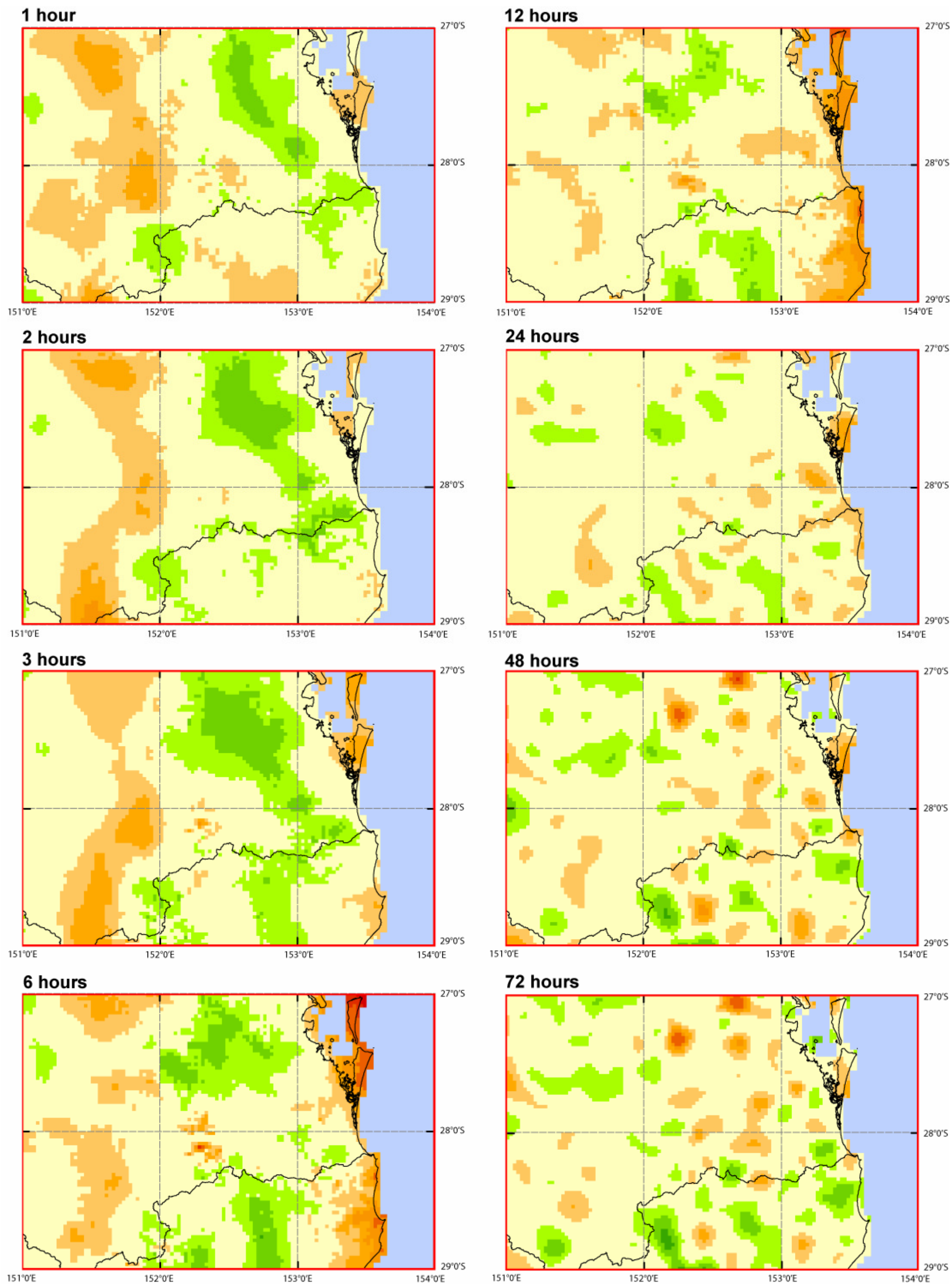


Figure 72 Indicative maps of percentage difference between design rainfall estimates derived in the pilot study and ARR87 for an ARI of 20 years.

8 ASSESSING UNCERTAINTY

When providing estimates of design rainfall estimates, it is essential to quantify the uncertainty associated with these estimates. The following inputs are considered to contribute to the overall uncertainty:

- data collection and ingestion related errors (instrumentation, digitising)
- sampling error
- model uncertainty (e.g. distributional assumption)
- mapping errors (which could be described using grids of standard errors)
- climate aspects (climate variability and climate change)

Internationally, most of the recently revised approaches allow for the supply of design rainfall estimates together with uncertainty estimates. The sophistication of methods applied in assessing uncertainties varies considerably, from standard methods (Thompson, 2002) to more complex methods addressing interrelationship between errors in estimating index rainfall and regional growth factors (Kjeldsen and Jones, 2004), and methods taking into account choice of distribution (Lin, 2004).

8.1 Validity of Assumption of a Stationary Climate

The classical frequency approach (as used for developing the methods in the pilot study) is based on the assumptions of independence and stationarity. There are major concerns as to whether these assumptions will still be justified under climate change.

Assessing changes based on observations

Changes in four rainfall indices

- maximum 5-day rainfall total,
- heavy rainfall days (defined as the number of days with rainfall totals above 10 mm),
- proportion of annual rainfall from very wet days and
- annual rainfall totals

were investigated using linear trends for the period 1951 to 2005 (Jakob et al 2005b). Some stations located in or near the pilot study area exhibited significant decrease in the number of days with heavy rainfall and the maximum 5-day rainfall (Figure 73).

A very simple assessment was undertaken to assess how these changes might have affected design rainfall estimates. At-site estimates of design rainfall for ARIs of 2, 5, 10, 20, 50 and 100 years and the 24-hour duration were derived based on two 50-year and two 25-year periods respectively, to explore the extent to which these estimates might depend on the period of record. When deriving estimates based on 50 years of data (1901 to 1950 and 1951 to 2000), estimates from the earlier period tended to be lower than for the later period (although these are not shown). For the two 25-year periods (1951 to 1975 and 1976 to 2000), estimates from the earlier period tended to be higher than those from the later period. Percentage differences for an ARI of 100 years are shown in Figure 74. It is possible that design rainfalls in mountainous areas could have increased while decreasing elsewhere. This assessment also shows that the effects are by no means negligible.

Khaliq et al (2006) discuss a number of relevant methods to address climate change. To address non-stationarity, different extremal models (annual maxima/'r largest' and GEV, partial duration series and POT model), time-varying parameters/moments; pooling; local likelihood approaches and quantile regression approaches could be considered. With respect to regional frequency analysis, the definition of homogeneous regions needs to take potential effects of non-stationarity into account.

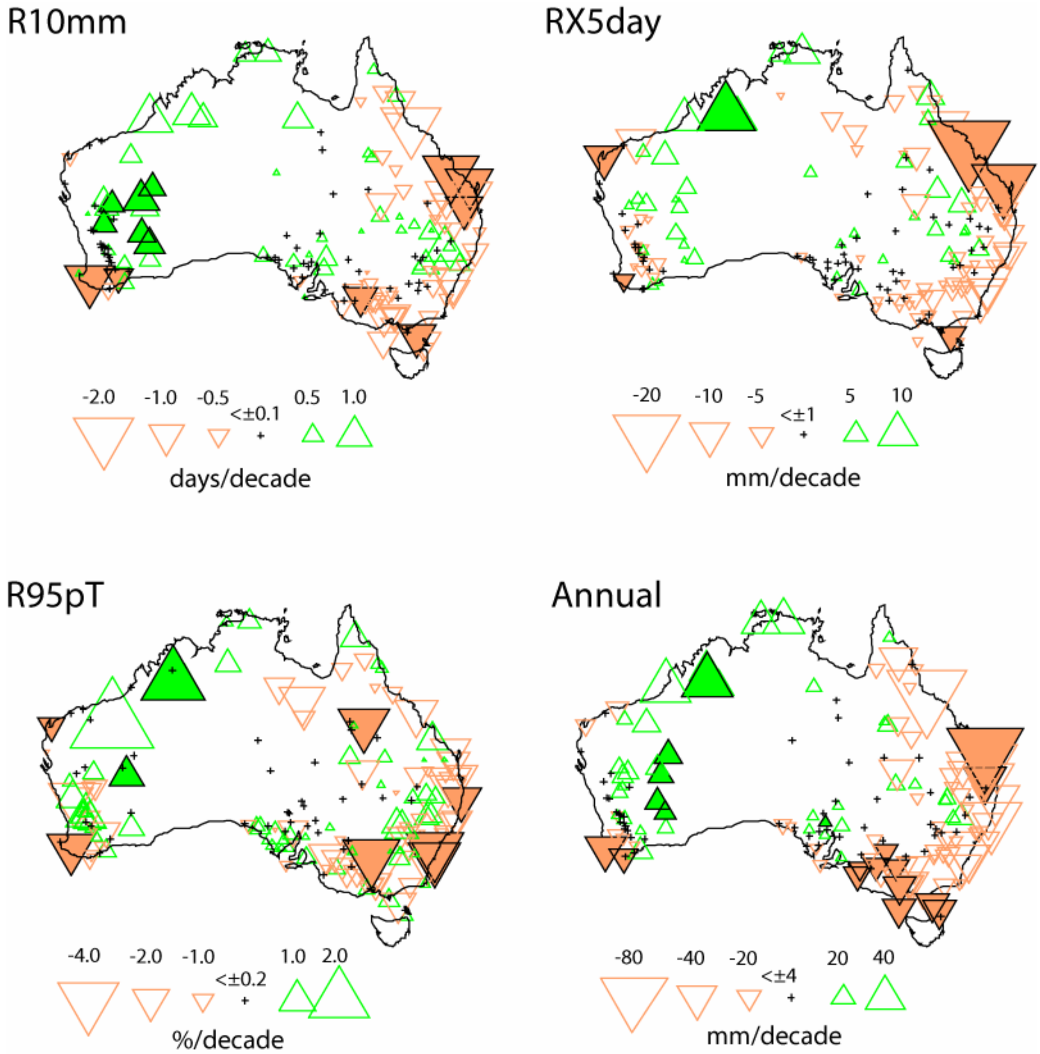


Figure 73 Annual trends in three rainfall indices (*R10mm*, *RX5day* and *R95pT*), and annual total rainfall, for the period 1951-2005. Trends significant at the 0.05 level are shown as solid symbols with a black outline. (Source: Jakob et al 2008b)

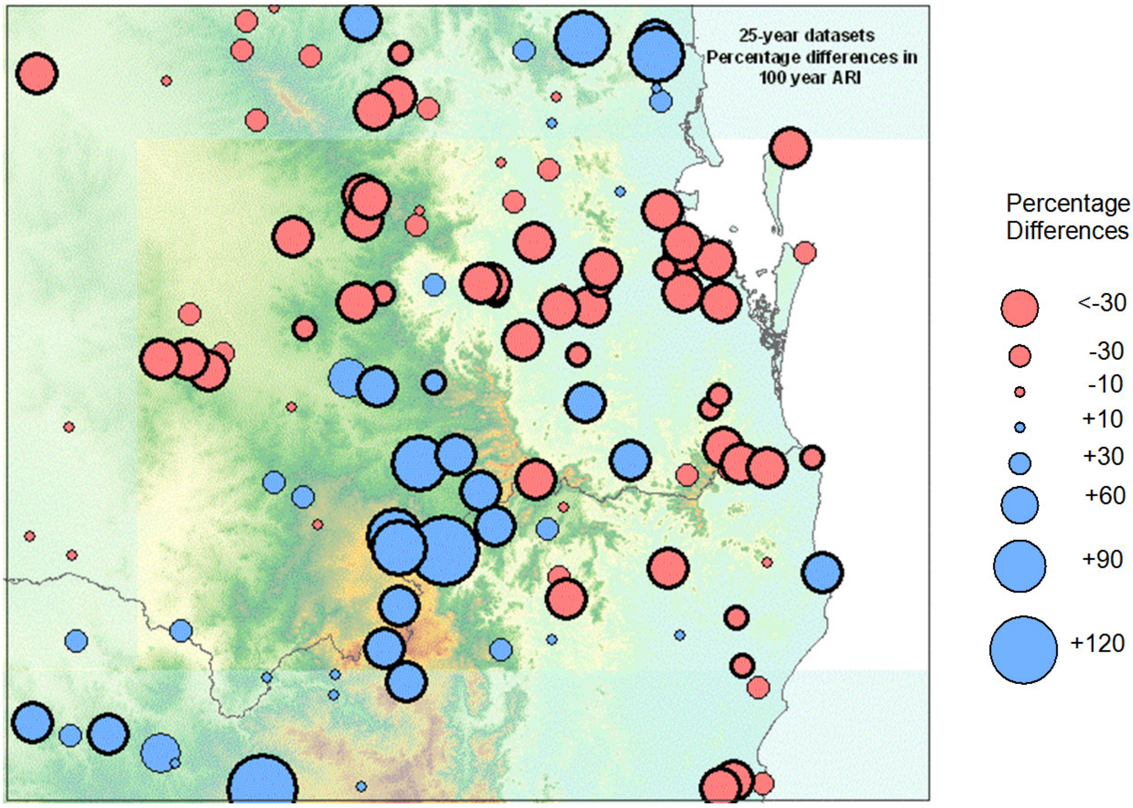


Figure 74 Percentage differences in design rainfall estimates for two 25-years periods (1951-1975 and 1976-2000) for the 24-hour duration and an ARI of 100 years. Solid black outline indicates differences are statistically significant. (Source: Jakob et al 2008b)

Using climate model output

Observed and projected changes in intensity of precipitation are discussed in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change:

> Basic theory, climate model simulations and empirical evidence all confirm that warmer climates, owing to increased water vapour, lead to more intense precipitation events even when the total annual precipitation is reduced slightly, and with prospects for even stronger events when the overall precipitation amounts increase.
> (IPCC, 2007)

According to a recently released report (CSIRO and Bureau of Meteorology, 2007), increases in daily precipitation intensity (defined as rain per rain-day) are likely for large parts of Australia. Extreme daily precipitation (defined as the highest 1%) will tend to increase in the north and decrease in the south but these changes will vary with season.

There are two issues that need to be considered when interpreting climate model output with respect to design rainfall:

- For higher percentiles the models' precipitation intensity is likely to be much weaker than observed but the synoptic behaviour (500 hPa) is generally reproduced well (Gutowski et al, 2008).

- Robustness of estimates varies with season and geographic location. Depending on the signal-to-noise ratio and the required significance, two strategies could be employed: the ensemble size could be increased and spatial pooling (regionalisation) could be applied (Kendon et al 2008).

8.2 Confidence Intervals

Uncertainty estimates will have to be provided at each grid point for all combinations of standard ARI and durations. Although not implemented in the current study, consideration was given to possible approaches that could be adopted for the estimation of uncertainty.

An automated procedure will be required to provide the full set of uncertainty estimates. As an alternative to a non-parametric bootstrapping approach, a maximum likelihood approach or a Bayesian approach could be considered. Within a Bayesian framework it is possible to take into account sampling errors as well as model assumptions and parameter uncertainty to develop a joint assessment of uncertainty in the final estimates. For instance, Kwon et al (2008) describe the use of a hierarchical Bayesian approach for ‘climate informed’ flood frequency analysis. A maximum likelihood approach would allow incorporating effects of temporal and/or spatial non-stationarities and other covariates (Khaliq et al 2006).

Using a bootstrap approach, it may not be straightforward to assess the joint uncertainty. Although sampling and mapping errors could be assessed using rigorous statistical methods, for other sources of uncertainty it may be more appropriate to quantify the effect to provide guidance rather than an exact estimate.

Standard techniques could be used to assess the sampling uncertainty when deriving uncertainty estimates for a single site. Confidence intervals are derived using a resampling approach based on the series of annual maxima. Related scripts have been written as part of the pilot study. Such confidence intervals essentially represent sampling uncertainty but ignore other sources (such as choice of technique [L-moments] or distribution [GEV]).

With additional assumptions, this part could be extended to cover regional estimates (at gauged locations). Estimation of confidence intervals could be based on the extremal series for the sites that are included in the region (ignoring predictions for now, addressed later). As before, resampling could then be used to derive *sets* of at-site L-moments. For each of these sets, regional L-moments could be calculated. Confidence intervals could then be derived as for the at-site estimates above. This procedure implies the assumption that the regionalisation (circles/modified circles) is known, including appropriate region size and correct weights for the site L-moments. (This approach does not yet cover uncertainty in estimating the index rainfall.)

For subdaily estimates, direct and predicted estimates are combined. To fully address the problem, two separate estimates for confidence intervals (one for direct estimates and one for predictions) would have to be derived. These would then have to be combined in some meaningful way (taking into account the weights chosen for direct/predictions when calculating the regional L-moments). To estimate confidence intervals for the predictions (which are based on daily AM) a resampling approach to calculate a number of L-moment estimates could be developed. These estimates could then be fed into the prediction model to derive a number of predictions (of sub-daily L-moments) which could then be combined with the resampled direct L-moment estimates.

Since this approach has not been tested it is not known how easy it would be to implement. The fact that uncertainty is involved in predicting the subdaily L-moments has so far been ignored, only the sampling error has been taken into account. It might be sensible to assess the respective magnitudes of prediction and sampling errors, possibly by investigating a few well-chosen cases first.

9 TERMINOLOGY

When deciding on terminology and scope for use in a revision of IFD estimates, it might be sensible to consider not only ARR87 but also current best practice based on recently revised methods (see Table 13).

Table 13 Design rainfall estimation in Australia and for recently revised methods

Country	***Year***	***Durations***	***Frequency***	***Extremes***	***Depth/ Intensity***	***Comments***
Australia	1987	5 min - 72 hours	ARI 1 to 100 years	Annual maxima	Intensity	Online (point, for standard durations and ARI)
New Zealand	2002	10 min - 72 hours	ARI 2 to 150 years	Annual maxima	Depth	Software
UK	1999	1 hour - 8 days	Return periods 2 to 1000 years	Annual maxima	Depth	Software (point/ catchment, event rarity), FORGEX method
US	On-going (by state)	5 min - 60 days	AEP 2 to 1000 years (for annual maxima), ARI for PDS	PDS and Annual maxima	Depth and Intensity	Online; Seasonality, Grids of confidence intervals; possible reduction in frequency range supplied

10 REFERENCES AND RELATED PUBLICATIONS

Alila, Y. (1999), A hierarchical approach for the regionalization of precipitation annual maxima in Canada, Journal of Geophysical Research vol. 104, no. 36, pp. 31645-31655.

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T. ,Yekta, M., Riley, D. (2004a), Precipitation-Frequency Atlas of the United States, NOAA Atlas 14, Volume 1 Version 4.0: Semiarid Southwest (Arizona,Southeast California, Nevada, New Mexico, Utah), U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland, 2004 (revised 2006).

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T. ,Yekta, M., Riley, D. (2004b), Precipitation-Frequency Atlas of the United States, NOAA Atlas 14, Volume 2 Version 3.0 (Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, NewJersey, North Carolina, Ohio, Pennsylvania,South Carolina, Tennessee, Virginia, West Virginia), U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland, 2004.

Boughton, W. and Jakob, D. (2008), Adjustment factors for restricted rainfall, Australian Journal of Water Resources, Vol. 12, No. 2.

CSIRO and the Bureau of Meteorology (2007), Climate Change in Australia: Technical report 2007, 148 pp. www.climatechangeinaustralia.gov.au

Daly, C., Taylor, G. and Gibson, W. (1997), The PRISM Approach to Mapping Precipitation and Temperature, 10th Conf. on Applied Climatology, American Meteorology Society, 10-12, Reno, Nevada. http://www.ocs.orst.edu/pub/prism/docs/appclim97-prismapproach-daly.pdf

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L. and Pasteris, P. (2002), A knowledge-based approach to statistical mapping of climate, Climate research, Vol22, pp 99 - 113.

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.

Gutowski, W. J., Willis, S. S., Patton, J. C., Schwedler, B. R. J., Arritt, R. W. and Takle, E. S. (2008) Changes in extreme, cold-season synoptic precipitation events under global warming. Geophysical research Letters, 35, L20710.

Haan, C. T. (2002), Statistical Methods in Hydrology, Second edition, Iowa State Press.

Hargraves, G.W. (2004) Final Report - Extreme Rainfall Estimation Project, CRCFORGE and (CRC)ARF Techniques, Queensland and Border Locations, Development and Application, NR&M Internal Document.

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. (2007), ANUSPLIN version 4.37 User Guide, The Australian National University, Centre for Resources and Environmental Studies, Canberra.

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.

IPCC (2007), Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (Eds), Cambridge, United Kingdom and New York, USA, 996 pp.

Jakob, D., Taylor, B., Xuereb, K. (2005a), A Pilot Study to Explore Methods for Deriving Design Rainfalls for Australia – Part 1, Hydrology Report Series, HRS Report No. 10.

Jakob, D., Taylor, B. and Xuereb (2005b), A Pilot Study to Explore Methods for Deriving Design Rainfalls for Australia, Proceedings of the 29th Hydrology and Water Resources Symposium, 21-23 February 2005, Canberra.

Jakob, D., Xuereb, K. and Taylor, B. (2007), Revision of Design Rainfalls over Australia: A Pilot Study, Australian Journal of Water Resources, Vol.11, No. 2.

Jakob, D., Meighen, J., Taylor, B. and Xuereb, K. (2008a), Methods for deriving design rainfall estimates at sub-daily durations. Conference proceedings of Water Down Under 2008, 14-17 April 2008, Adelaide, Australia, 11 pp.

Jakob, D., Smalley, R., Meighen, J., Xuereb, K. and Taylor, B. (2008b), Climate change and probable maximum precipitation, Final report to the Commonwealth Department of Climate Change and the Queensland Department of Natural Resources and Water.

Jeffrey, S. J., Carter, J. O., Moodie, K. B. and Beswick, A. R. (2001), Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environmental Modelling & Software, 16, 309-330.

Jones, D. A., Wang, W. and Fawcett, R., (2007). Climate data for the Australian Water Availability Project, Final Milestone Report.

Kendon, E. J., Rowell, D. P., Jones, R. G., Buonomo, E. (2008), Robustness of future changes in local precipitation extremes, Journal of Climate, Vol21, pp 4280-4297.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J.-C., Gachon, P. and Bobée, B. (2006), Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, Journal of Hydrology, 329, pp 534-552.

Kjeldsen, T. R. and Jones, D. A. (2004), Sampling variance of flood quantiles from the generalised logistic distribution estimated using the method of L-moments, Hydrology and Earth System Sciences, 8(2), 183-190.

Kjeldsen, T. R. and Jones, D. A. (2006), Prediction uncertainty in a median-based index flood method using L moments, Water Resources Research, Vol 42.

Kwon, H., Brown, C. and Lall, U. (2008), Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modelling, Geophysical Research Letters, Vol 35.

Lavery, B., Joung, G., and Nicholls, N. (1992). A high–quality historical rainfall data set for Australia. Australian Meteorological Magazine. 40, 33–39.

Lin, B., Bonnin, G., Todd, D., Parzybok, T., Yekta, M. 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), Beijing, China, June 27-30, 2004.

Nandakumar, N., Weinmann, P. E., Mein, R. G., Nathan, R. J. (1997), Estimation of Extreme Rainfalls for Victoria using the CRCFORGE Method, CRCCH.

Menabde, M., Seed, A., and Pegram, D. (1999), A simple scaling model for extreme rainfall, Water Resources Research, 35(1), 335-340.

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

Nguyen, V. and Nguyen, T. (2008), A spatial-temporal downscaling approach to construction of intensity duration frequency relations in consideration of GCM-based climate change scenarios, Expert Symposium on Climate Change: Modelling, Impacts & Adaptations, Singapore.

Srikanthan, R., Peel, M., McMahon, T. and Jakob, D. (2008), Application of the generalized extreme value distribution to Australian peak discharge data, accepted for publication.

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.

Viney, N. R. and Bates, B. C. (2004), It never rains on Sunday: The prevalence of untagged multi-day rainfall accumulations in the Australian high-quality data set, Int. J. Climatol. 24: 1171–1192.

Wallis, J. R., Schaefer, M. G., Barker, B. L. and Taylor, G. H. (2007), Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for Washington State, Hydrol. Earth Syst. Sci., 11(10), 415-442, 2007

Xuereb, K., Taylor, B. and Jakob, D. (2006), Revision of Design Rainfalls over Australia: A Pilot Study, Proceedings 30th Hydrology and Water Resources Symposium, 4 - 7 December 2006, Hobart.

11 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 (AM series) — a series that contains only the event with the largest magnitude that occurred in each year.

ArcGIS / **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 AM series: 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.

Bayesian — A Bayesian approach can be used as an alternative to the classic 'frequentist' approach to estimating parameters of an underlying distribution see Haan (2002).

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.

Direct estimates and **predictions of L-moments —** These terms are used in reference to L-moments calculated at subdaily durations. L-moments were calculated on the basis of annual maxima where subdaily data are available (direct estimates); alternatively they can be inferred from information at longer durations (predictions).

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 (as opposed to **sliding** 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.

Gumbel reduced variate (y_G) — For extreme value plots, the frequency axis is often appropriately rescaled using a 'reduced variate'. The choice of reduced variate depends on distribution fitted; the Gumbel reduced variate is appropriate for use with the GEV distribution. Using a Gumbel reduced variate scale, the 2-parameter Gumbel distribution would plot as a straight line. $y_G = -\log(-\log(F))$

HYDSYS format — A specific format of electronic files used with HYDSTRA (formerly HYDSYS) software (http://www.kisters.de/english/html/au/homepage.html).

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. median annual maximum 24-hour rainfall) at that site, used in regionalisation.

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 (L-CV, L-skewness and L-kurtosis) — 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. Originally records were taken using a Dines Pluviograph but, starting in 1997, increasingly from tipping bucket raingauges (TBRG).

POT model — Peaks-Over-Threshold model. This refers to modeling threshold exceedences. If the assumptions for a POT model are valid, then certain distributions can be used to model occurrence and magnitude of exceedence.

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.

Residuals — The difference between the observed and predicted values; the sum of the residuals can be used to characterize the fitting error of a model.

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 (as opposed to **fixed** values in a rainfall time series) — Rainfall recorded essentially continuously in time (e.g. 1 minute, 6 minutes intervals).

12 ACRONYMS

ADAM — Australian Data Archive for Meteorology

AEP — Annual Exceedence Probability

ALERT — Automated Local Evaluation in Real-Time

AM — Annual Maxima

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

ARI — Average Recurrence Interval

ARR, ARR87 — Australian Rainfall and Runoff, 1987.

AWS — Automatic Weather Station

CDIRS — Computerised Design IFD Rainfall System

CRCFORGE — Cooperative Research Centre - FOcussed Rainfall Growth Estimation

CSIRO — Commonwealth Scientific and Industrial Research Organisation

DEM — Digital Elevation Model

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

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

GCV — Generalised Cross Validation

GEV — General Extreme Value (distribution)

GIS — Geographic Information System

GLS — Generalised Least Squares

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

HRS10 — Hydrology Report Series, Number 10.

IEAust — Engineers Australia (formerly The Institution of Engineers Australia)

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

IPCC — Intergovernmental Panel on Climate Change

IQR — Interquartile Range

MAR — Mean Annual Rainfall

NCC — National Climate Centre (Australian Bureau of Meteorology)

NOAA — National Oceanic and Atmospheric Administration

PCA — Principal Component Analysis

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

TBRG — Tipping Bucket Raingauge

APPENDIX

A1 Rainfall near Comparison Report

Rainfall Near Comparison Report

Precipitation (in 10th of mm) near a station

Station 40074 ERNEST JUNCTION RAIL STN
Year / Month 192701
No Neighbours 20
QC F68 Only N

Rainfall QC | Near 3 Mths | Daily Totals | Daily Details | Near Graph | Sites DB | Prev Mth | Next Mth

Station	Ht	Name	Dist Km	Ob Qual	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
40074	22	ERNEST JUNCTION RAIL	0.0	0	0	0	0	0	0	0	767	0	0	0	1295	0	0	0	0	0	0	0	0	0	0	0	0	0	3810	0	0	0	0	0	0
40160	25	NERANG GILSTON RD	5.7	0	0	0	79	0	229	0	0	0	1377	152	264	23	0	0	0	0	33	102	813	894	0	20	457	787	325	752	180	551	170	0	0
40190	17	SOUTHPORT RIDGEWAY A	5.7	0	0	0	0	0	295	0	0	0	2042	114	249	8	0	0	0	0	0	38	538	432	0	0	521	653	152	345	178	99	234	0	0
40602	21	GILSTON STATE SCHOOL	8.9	0	0	165	0	0	538	0	0	0	1003	0	302	0	0	0	0	0	0	0	1029	381	0	81	589	897	267	909	1133	140	0	0	0
40166	9	OXENFORD (OBERON WAY	9.0	0	0	180	0	0	264	0	0	0	810	102	262	5	0	0	0	0	33	84	803	546	23	20	427	643	363	627	213	564	274	0	0
40197	515	MT TAMBORINE FERN ST	15.2	0	0	0	0	3	267	0	0	28	404	61	343	5	8	0	0	0	15	84	813	917	91	10	754	658	696	1626	765	422	30	0	0
40042	100	CANUNGRA FINCH ROAD	18.8	0	0	0	0	0	0	0	0	0	264	76	208	0	0	0	0	0	0	0	505	853	51	36	650	503	417	1067	554	236	25	0	0
40196	12	TALLEBUDGERA GUINEAS	19.6	0	0	0	0	0	104	0	0	48	0	1229	269	0	0	0	0	0	13	46	284	660	13	0	0	1422	361	1041	295	102	234	0	0
40524	177	LITTLE NERANG DAM	20.7	0	0	18	18	8	396	0	0	30	579	201	193	20	13	0	0	0	36	124	465	1059	30	18	572	615	607	1440	587	800	234	0	0
40504		BROMFLEET	24.1	0																															
40000	48	ABBOTSFORD	24.6	0	0	0	38	140	0	0	0	0	124	53	157	0	0	0	0	0	0	505	625	0	0	0	432	472	216	518	338	135	30	0	0
40015	520	BEECHMONT BINNA BURR	25.4	0	3	0	432	18	130	3	0	3	170	157	208	15	15	0	0	0	23	119	732	1168	46	23	678	691	706	1448	973	645	140	0	0
40150	61	MUNDOOLIN	26.1	0	0	0	0	20	41	0	0	0	127	30	102	5	0	0	0	0	0	0	610	914	20	0	528	437	127	462	419	89	53	0	0
40599		CAMBERRA	26.1	0																															
40538	76	TABRAGALBA	26.8	0																															
40615	100	ROTTINGTON	27.8	0																															
40600	137	GLEN ELLEN	29.7	0	0	0	0	117	0	3	0	51	1074	160	650	3	0	0	13	0	0	64	439	1143	25	5	1316	1262	678	1633	886	531	348	0	0
40192	806	SPRINGBROOK FORESTRY	29.7	0	0	0	8	15	584	5	13	76	752	467	1052	28	23	0	0	0	61	173	536	1120	112	94	1501	1450	1138	2718	1732	1755	366	41	0
58067	330	TOMEWIN (BORDER GATE	30.6	0	0	0	0	5	97	5	0	64	1181	117	450	0	0	0	0	0	0	33	353	1036	15	8	775	1067	495	1544	826	389	277	13	0
58092		WARNLEIGH	31.6	0	0	0	0	104	20	0	0	71	709	0	249	10	0	0	0	0	23	249	279	826	15	8	737	953	269	823	381	221	183	0	0

A2 Mapping

A2.1 Mapping Subdaily Index Rainfall - Trials

- These trials used the 12-hour data initially, as this was considered the most reliable duration; as they are closer to the data rich 24-hour duration.
- All the data were for a 2 year ARI.
- The trials were all run on the annual maximum series (medians) rather than the partial duration series and the ultimate grids need to be multiplied by the conversion factor (1.11 for 2 year data).
- The Mean Annual Rainfall (MAR) grid was provided by the National Climate Centre, Bureau of Meteorology. The extent is 150.5°E,154.0°E/-26.5°S,-29.5°S and the grid spacing is 0.025°.
- The elevation grid was constructed from the Geoscience Australia 9 second DEM. The original grid was 'aggregated' within ArcMap to a courser grid (grid spacing 0.05°) to dampen the effect small scale extreme changes when used in splining. Note that whilst the information is at 0.05 spacing, the actual grid has points at 0.025 spacing to match all the other grids. It has the same extent and spacing as the MAR grid.

<u>Trials A to F</u>
These trials mainly use elevation and mean annual rainfall (MAR) as independent variables or independent covariates in either or both of SPLINA and LAPGRID. As the MAR is considered to use elevation as an input as well, these grids all reflected the elevation model to an unrealistic extent and will not be used.

However, the sequence C1, C2 and C3 can be used to suggest the best way of gridding the direct estimates whilst incorporating the predictions. The set of predictions used was that based on a subset of sites; this having been settled on as the superior version in a previous investigation.
C1 Creates a grid of the predictions
C2 Uses the C1 grid values as independent covariate
C3 Uses the C1 grid values as independent variable

Examination of the statistics in the SPLINA log and comparisons displayed in a spreadsheet showed a slight advantage in using the predications as independent covariates.

<u>Trials G to L</u>
This series of trials tried to simplify the inputs to the grids in order for the rainfall data to be mapped without too much influence from the other inputs. Grids G and L show the 'raw' splining of the predictions and the direct estimates respectively. Grids H to K all use the predictions (grid G) as independent covariates and combinations of the elevation and MAR as independent variables.

After examination of the different versions it was decided to repeat the exercise for data at the 1-hour duration i.e. mapping 1-hour versions of maps G, H and K.

<u>Trials M to O</u>
These reproduce maps G, H and K but use 1-hour instead of 12-hour data. One object of the mapping was to test if the reasonable results obtained for 12-hour data in Trial K was reproducible at 1 hour: it was.

Trials P and Q

All the previous trials had used order 3 when splining and other orders ought to be tested. These trials both used the input configurations as for Trial K (12 hours) and Trial O (1 hour) and only changed the spline order. It was found that although order 2 is the most usual according to the ANUSPLIN notes and the on-screen interface ‘help’, this order was too low for the chosen input configuration. So only the higher order spline (4) was tested.

Splining with order 4 produced over-fitted mapping with many spurious areas around the edges of the grids.

Trials R to U

Another aspect of splining is the choice of a ‘smoothing directive’. All the previous trials used the default option of minimising the generalised cross validation (GCV), as this is the method recommended in the ANUSPLIN notes and the only one used in its worked examples.

However in talks with Craig Thompson from the National Institute of Water and Atmospheric Research (New Zealand), he said that they had used, amongst other differences, the fixed signal option as the ‘smoothing directive’. By setting the signal to a figure, the signal to noise ratio is effectively set as the signal plus the error (noise) must add to the number of knots used by the splining program which is constant for a particular input dataset.

Trials R and T tested the 1- and 12-hour data respectively with the signal ‘fixed’ at 44. The input data for this mapping gives 66 knots so, by setting the signal at 44, the error must be 22 which gives a signal to noise ratio of 2:1.

Trials S and U tested the 1- and 12-hour data respectively with the signal ‘fixed’ at 52. Again the input data for this mapping gives 66 knots so, by setting the signal at 52, the error must be 14 which gives a signal to noise ratio of roughly 4:1.

Values of index rainfall were extracted from the grid of Trials K, S and U (12-hours) at the sites of the Direct Estimates and compared in a spreadsheet. Similarly this was done for the 1-hour Trials O, R and T.

It was found that using the ‘fixed signal’ option gave closer matches to the Direct Estimates whist not producing over-spined or unrealistic grids. The 4:1 signal-to-noise versions were slightly better than the 2:1 ratio versions.

A2.2 Mapping Subdaily Index Rainfall - Sample Inputs and Logs

Sample input data, ANUSPLIN command and log files

Sample input data - 12hour duration, 2year ARI

Station, Lon, Lat, Elevation, Mean Annual Rainfall, Prediction, Direct Estimate

```
40000,153.1,-27.95,76.73,1006,78.82748,-99
40004,152.7111,-27.6297,19.93,930,76.20392,71.75
40006,152.9167,-28.2833,238.71,1138,76.05771,-99
40007,152.9,-26.8,130.12,1945,137.23537,-99
40008,152.7,-27.4167,81.75,1078,88.53243,-99
40012,152.8,-28.25,184.41,1021,75.3138,-99
40014,153.0131,-28.0206,99.17,951,70.0132,62.35
40015,153.1903,-28.1467,383.14,1478,105.7227,-99
40016,153.1833,-27.7167,25.55,1266,94.98388,-99
40017,152.9764,-26.8564,22.18,1638,115.0046,-99
40019,152.15,-26.9,374.8,908,69.55608,69.6
40020,152.1006,-26.8844,436.93,888,68.16417,-99
40024,152.6919,-27.9925,122.17,883,69.75624,-99
40027,153.1658,-27.0892,0.01,1380,100.55919,-99
40030,152.4775,-28.2367,866.49,1629,98.32229,-99
40031,153.05,-26.6867,31.25,1846,125.99624,-99
40032,153.0667,-26.6833,31.25,1846,133.65383,-99
40035,153.0089,-27.1414,3.39,1309,97.07607,-99
40037,152.5,-27.15,206.19,936,68.826,-99
40038,152.9517,-27.085,8.77,1371,95.65591,-99
40039,153.1383,-26.8083,4.24,1716,120.52619,-99
40040,153.15,-26.8017,-9999,1614,106.49753,-99
40041,152.8664,-27.3961,59.98,1201,83.14673,-99
40042,153.1664,-28.0142,160.15,1495,92.84064,-99
40043,153.4661,-27.0314,10.02,1559,92.99316,-99
.
.
.
```

Sample SPLINA command file

```
INDEX RAIN 12h 2y
0  -99.0
4
1
0
0
150.50  154.00  0  5  0.00
-29.50  -26.50  0  5  0.00
0.00  3000.00  1  1  0.00
1000.00
0.00  5000.00  3  0  0.00
1.00  0.00
0.00  1000.00  0  0  0.00
2
3
1
0
1
3
71.00
arrr\dir_sub_pred_elev_mar_12h.csv
1000
6

arrr\median_12h.sur

arrr\median_12h.cov
```

Sample LAPGRD command file

```
arrr\median_12h.sur
1
1
1
arrr\median_12h.cov
2
1000.0
1
1
150.50  154.00  0.025
2
-29.50  -26.50  0.025
2
arrr\mask.asc
2
arrr\dem.asc
2
arrr\MAR_025.asc
2
arrr\preds_12h.asc
2
-999.0
arrr\median_12h.grd
(10f9.3)
2
-999.0
arrr\median_errors_12h.grd
(10f9.3)
```

Sample SPLINA log file

```
SPLINA  VERSION 4.37  16/07/07
COPYRIGHT AUSTRALIAN NATIONAL UNIVERSITY

TITLE OF FITTED SURFACES (60 CHARS):
-----------------------------------
INDEX RAIN 12h 2y

SURFACE VALUE UNITS CODE AND MISSING DATA VALUE:
-----------------------------------------------
  0 - UNDEFINED
  1 - METRES
  2 - FEET
  3 - KILOMETRES
  4 - MILES
  5 - DEGREES
  6 - RADIANS
  7 - MILLIMETRES
  8 - MEGAJOULES
    0  -99.00

INDEPENDENT VARIABLES
---------------------
NUMBER OF INDEPENDENT SPLINE VARIABLES (0 TO 10):
    4

NUMBER OF INDEPENDENT COVARIATES (0 TO  6):
    1

NUMBER OF SURFACE SPLINE VARIABLES (0 TO  5):
    0

NUMBER OF SURFACE COVARIATES (0 TO  5):
    0

TRANSFORMATION CODES            REFERENCE UNIT CODES
--------------------            --------------------
  0 - NO TRANSFORMATION           0 - UNDEFINED
  1 - X/A                         1 - METRES
  2 - X*A                         2 - FEET
  3 - A*LOG(X + B)                3 - KILOMETRES
  4 - (X/B)**A                    4 - MILES
```

```
  5 - A*EXP(X/B)                  5 - DEGREES
  6 - A*TANH(X/B)                 6 - RADIANS
  7 - ANISOTROPY ANGLE            7 - MILLIMETRES
  8 - ANISOTROPY FACTOR           8 - MEGAJOULES

LOWER & UPPER LIMITS, TRANSF CODE, REF UNIT, MARGIN(S) FOR VARIABLE  1:
   150.500       154.000       0         5

LOWER & UPPER LIMITS, TRANSF CODE, REF UNIT, MARGIN(S) FOR VARIABLE  2:
  -29.5000      -26.5000       0         5

LOWER & UPPER LIMITS, TRANSF CODE, REF UNIT, MARGIN(S) FOR VARIABLE  3:
   0.00000       3000.00       1         1
ENTER  1 TRANSFORMATION COEFFICIENT(S):
   1000.00

LOWER & UPPER LIMITS, TRANSF CODE, REF UNIT, MARGIN(S) FOR VARIABLE  4:
   0.00000       5000.00       3         0
ENTER  2 TRANSFORMATION COEFFICIENT(S):
   1.00000       0.00000

LOWER & UPPER LIMITS, TRANSF CODE, REF UNIT, MARGIN(S) FOR VARIABLE  5:
   0.00000       1000.00       0         0

SURFACE DIRECTIVES
------------------
DEPENDENT VARIABLE TRANSFORMATION:
  0 - NO TRANSFORMATION
  1 - NATURAL LOGARITHM
  2 - SQUARE ROOT
  5 - OCCURRENCE
    2

ORDER OF SPLINE (AT LEAST 3):
    3

NUMBER OF SURFACES (AT LEAST 1):
    1

NUMBER OF RELATIVE VARIANCES (0 OR 1):
    0

OPTIMIZATION DIRECTIVE (NORMALLY 1):
  0 - COMMON SMOOTHING PARAMETER FOR ALL SURFACES
  1 - COMMON SMOOTHING DIRECTIVE FOR ALL SURFACES
  2 - DIFFERENT SMOOTHING DIRECTIVE FOR EACH SURFACE
    1

SMOOTHING DIRECTIVE (NORMALLY 1):
  0 - FIXED SMOOTHING PARAMETER FOR EACH SURFACE
  1 - MINIMIZE GCV FOR EACH SURFACE
  2 - MINIMIZE TRUE MEAN SQUARE ERROR FOR EACH SURFACE
  3 - FIXED SIGNAL FOR EACH SURFACE
  4 - MINIMIZE GML FOR EACH SURFACE
    3

INPUT SIGNAL:
   71.00000

DATA FILE NAME:
--------------
arrr\dir_sub_pred_elev_mar_12h.csv

MAXIMUM NUMBER OF DATA POINTS (AT LEAST 16):
    1000

NO. OF CHARACTERS IN SITE LABEL (0 TO 20):
    6

DATA FORMAT (SITE LABEL,  5 INDEP VARS,  1 SURFACES,  0 REL VARIANCES):


INPUT BAD DATA FLAG FILE NAME (BLANK IF NOT REQUIRED):


OUTPUT BAD DATA FLAG FILE NAME (BLANK IF NOT REQUIRED):
```

```
OUTPUT LARGE RESIDUAL FILE NAME (BLANK IF NOT REQUIRED):


OUTPUT OPTIMIZATION PARAMETERS FILE NAME (BLANK IF NOT REQUIRED):


OUTPUT SURFACE COEFFICIENTS FILE NAME (BLANK IF NOT REQUIRED):
arrr\median_12h.sur

OUTPUT DATA LIST FILE NAME (BLANK IF NOT REQUIRED):


OUTPUT ERROR COVARIANCE FILE NAME (BLANK IF NOT REQUIRED):
arrr\median_12h.cov

VALIDATION DATA FILE NAME (BLANK IF NOT REQUIRED):
-------------------------------------------------


DATA SUMMARY
------------
NUMBER OF DATA POINTS READ     =    645
NUMBER OF POINTS WITHIN LIMITS =    644
POINTS    30 40053  AND    31 40054  COINCIDE

 SURF  MEAN RELATIVE VARIANCE  ROOT MEAN REL VAR
    1              1.00              1.00

NUMBER OF KNOTS      =    88

SURFACE STATISTICS
------------------
SURF       RHO    NPTS   KNOTS     ERROR    SIGNAL    SURF     MEAN     STD DEV
   1 0.185E-04      88      88      17.0      71.0       1   81.651       24.39

SURF       GCV       MSR       VAR     SURF   RTGCV     RTMSR     RTVAR
   1     0.294     0.110E-01 0.568E-01    1   0.542     0.105     0.238

SURF       GML       MSE       VAR     SURF   RTGML     RTMSE     RTVAR
   1     0.292     0.458E-01 0.347E-01    1   0.540     0.214     0.186

SURF   COVARIATES AND STANDARD ERRORS
   1    0.273E-01    0.103E-01

APPROXIMATE UNTRANSFORMED STATISTICS
------------------------------------
SURF       GCV       MSR       VAR     SURF   RTGCV     RTMSR     RTVAR
   1      96.2      3.58      18.6        1    9.81      1.89      4.31

SURF       MSE     RTMSE
   1      15.0      3.87

RANKED ROOT MEAN SQUARE RESIDUALS
---------------------------------
     1    -60  40106    0.299
     2   -288  40854    0.266
     3   -215  40386    0.257
     4   -143  40222    0.255
     5   -225  40406    0.243
     6   -515  41512    0.236
     7   -121  40189    0.226
     8   -583  58013    0.212
     9   -511  41467    0.188
    10   -395  41140    0.187
    11   -286  40769    0.173
    12    -11  40019    0.172
    13   -201  40318    0.151
    14   -136  40214    0.146
    15   -273  40609    0.143
    16   -642  58158    0.134
    17   -174  40265    0.129
    18   -631  58129    0.127
    19    -57  40101    0.117
    20    -44  40076    0.115
    21     -2  40004    0.107
    22   -513  41497    0.104
```

23	-272	40608	0.102
24	-112	40180	0.100
25	-266	40553	0.984E-01
26	-326	41044	0.959E-01
27	-7	40014	0.958E-01
28	-237	40454	0.897E-01
29	-516	41522	0.894E-01
30	-304	41018	0.878E-01
31	-268	40584	0.875E-01
32	-99	40160	0.850E-01
33	-37	40063	0.832E-01
34	-611	58072	0.831E-01
35	-279	40659	0.793E-01
36	-626	58113	0.720E-01
37	-239	40459	0.664E-01
38	-144	40223	0.614E-01
39	-126	40197	0.580E-01
40	-285	40715	0.574E-01
41	-337	41056	0.553E-01
42	-644	58192	0.542E-01
43	-506	41445	0.532E-01
44	-632	58131	0.516E-01
45	-282	40677	0.495E-01
46	-110	40178	0.493E-01
47	-36	40062	0.464E-01
48	-628	58116	0.457E-01
49	-58	40102	0.384E-01
50	-262	40537	0.364E-01
51	-160	40241	0.363E-01
52	-369	41100	0.357E-01
53	-183	40282	0.353E-01
54	-178	40270	0.346E-01
55	-625	58109	0.338E-01
56	-287	40849	0.332E-01
57	-623	58099	0.286E-01
58	-212	40382	0.283E-01
59	-253	40496	0.282E-01
60	-76	40133	0.264E-01
61	-53	40094	0.260E-01
62	-341	41060	0.258E-01
63	-385	41122	0.247E-01
64	-65	40112	0.238E-01
65	-601	58044	0.230E-01
66	-240	40460	0.227E-01
67	-198	40312	0.224E-01
68	-78	40135	0.223E-01
69	-540	56022	0.215E-01
70	-271	40606	0.209E-01
71	-394	41139	0.197E-01
72	-241	40461	0.190E-01
73	-590	58026	0.188E-01
74	-123	40192	0.187E-01
75	-614	58081	0.181E-01
76	-504	41430	0.147E-01
77	-238	40458	0.128E-01
78	-195	40308	0.117E-01
79	-281	40674	0.111E-01
80	-572	57095	0.968E-02
81	-554	56202	0.930E-02
82	-47	40082	0.875E-02
83	-552	56059	0.871E-02
84	-480	41359	0.815E-02
85	-528	54036	0.640E-02
86	-510	41457	0.584E-02
87	-409	41175	0.541E-02
88	-536	54104	0.284E-02
89	84	40141	0.00
90	88	40147	0.00
91	416	41192	0.00
92	91	40150	0.00
93	548	56044	0.00
94	92	40153	0.00
95	380	41116	0.00
96	93	40154	0.00
97	448	41261	0.00
98	94	40155	0.00
99	512	41469	0.00

100	95	40156	0.00		

RANKED RESIDUALS FOR SURFACE 1

1	-60	40106	0.299	71.9	77.1
2	-288	40854	0.266	81.5	86.4
3	-215	40386	-0.257	84.0	79.4
4	-143	40222	0.255	87.8	92.6
5	-225	40406	-0.243	92.6	88.0
6	-515	41512	0.236	46.2	49.5
7	-121	40189	0.226	60.7	64.3
8	-583	58013	-0.212	130.9	126.1
9	-511	41467	-0.188	75.2	72.0
10	-395	41140	0.187	55.9	58.7
11	-286	40769	-0.173	84.3	81.2
12	-11	40019	0.172	69.6	72.5
13	-201	40318	-0.151	77.8	75.2
14	-136	40214	-0.146	93.1	90.3
15	-273	40609	0.143	103.8	106.8
16	-642	58158	0.134	113.3	116.1
17	-174	40265	-0.129	103.4	100.8
18	-631	58129	0.127	90.2	92.6
19	-57	40101	-0.117	77.4	75.3
20	-44	40076	-0.115	78.3	76.3
21	-2	40004	0.107	71.8	73.6
22	-513	41497	-0.104	65.6	63.9
23	-272	40608	-0.102	108.6	106.5
24	-112	40180	-0.100	118.7	116.5
25	-266	40553	0.984E-01	105.1	107.1
26	-326	41044	-0.959E-01	56.0	54.6
27	-7	40014	0.958E-01	62.4	63.9
28	-237	40454	-0.897E-01	82.3	80.6
29	-516	41522	-0.894E-01	62.5	61.1
30	-304	41018	0.878E-01	48.1	49.3
31	-268	40584	-0.875E-01	106.5	104.7
32	-99	40160	0.850E-01	104.5	106.3
33	-37	40063	-0.832E-01	107.9	106.2
34	-611	58072	-0.831E-01	138.5	136.6
35	-279	40659	0.793E-01	79.7	81.1
36	-626	58113	-0.720E-01	109.5	108.0
37	-239	40459	0.664E-01	85.8	87.0
38	-144	40223	-0.614E-01	104.7	103.4
39	-126	40197	-0.580E-01	108.8	107.6
40	-285	40715	0.574E-01	86.5	87.5
41	-337	41056	0.553E-01	47.5	48.3
42	-644	58192	0.542E-01	66.3	67.2
43	-506	41445	0.532E-01	49.5	50.2
44	-632	58131	0.516E-01	129.9	131.1
45	-282	40677	0.495E-01	64.1	64.9
46	-110	40178	-0.493E-01	65.2	64.4
47	-36	40062	-0.464E-01	117.5	116.5
48	-628	58116	0.457E-01	78.5	79.4
49	-58	40102	0.384E-01	81.0	81.6
50	-262	40537	0.364E-01	112.9	113.7
51	-160	40241	-0.363E-01	91.5	90.8
52	-369	41100	-0.357E-01	54.7	54.2
53	-183	40282	-0.353E-01	114.9	114.1
54	-178	40270	-0.346E-01	82.3	81.6
55	-625	58109	-0.338E-01	97.7	97.0
56	-287	40849	0.332E-01	102.5	103.2
57	-623	58099	-0.286E-01	83.3	82.8
58	-212	40382	0.283E-01	63.4	63.9
59	-253	40496	-0.282E-01	111.3	110.7
60	-76	40133	-0.264E-01	72.5	72.1
61	-53	40094	-0.260E-01	68.3	67.9
62	-341	41060	-0.258E-01	58.4	58.0
63	-385	41122	0.247E-01	45.3	45.6
64	-65	40112	0.238E-01	67.1	67.5
65	-601	58044	-0.230E-01	93.9	93.5
66	-240	40460	0.227E-01	97.8	98.2
67	-198	40312	-0.224E-01	84.9	84.5
68	-78	40135	-0.223E-01	65.7	65.3
69	-540	56022	-0.215E-01	52.6	52.3
70	-271	40606	0.209E-01	108.5	108.9
71	-394	41139	-0.197E-01	53.6	53.3
72	-241	40461	0.190E-01	89.4	89.8
73	-590	58026	-0.188E-01	67.0	66.7
74	-123	40192	0.187E-01	159.0	159.5

```
    75   -614  58081   -0.181E-01         68.5     68.2
    76   -504  41430    0.147E-01         53.3     53.5
    77   -238  40458    0.128E-01         96.1     96.4
    78   -195  40308    0.117E-01        111.3    111.5
    79   -281  40674   -0.111E-01         65.2     65.0
    80   -572  57095   -0.968E-02         65.5     65.3
    81   -554  56202    0.930E-02         53.6     53.7
    82    -47  40082   -0.875E-02         58.3     58.2
    83   -552  56059   -0.871E-02         51.6     51.5
    84   -480  41359   -0.815E-02         57.4     57.3
    85   -528  54036    0.640E-02         56.0     56.1
    86   -510  41457    0.584E-02         52.3     52.4
    87   -409  41175   -0.541E-02         53.6     53.5
    88   -536  54104    0.284E-02         53.4     53.4
    89     84  40141     109.              0.0     96.3
    90     88  40147     109.              0.0    103.6
    91    416  41192     106.              0.0     54.6
    92     91  40150     108.              0.0     78.4
    93    548  56044     108.              0.0     75.6
    94     92  40153     106.              0.0     50.4
    95    380  41116     106.              0.0     51.1
    96     93  40154     107.              0.0     57.7
    97    448  41261     107.              0.0     59.8
    98     94  40155     107.              0.0     66.4
    99    512  41469     107.              0.0     59.6
   100     95  40156     107.              0.0     66.2

PROGRAM SPLINA  VERSION 4.37  DATE 28/08/2008  TIME 15.59.36
```

Sample LAPGRD log file

```
LAPGRD  VERSION  4.37  11/07/07
COPYRIGHT AUSTRALIAN NATIONAL UNIVERSITY

SURFACE FILE NAME:
-----------------
arrr\median_12h.sur

SURFACE TITLE = INDEX RAIN 12h 2y
SURFACE UNITS =

NUMBER OF SURFACES  =    1
ORDER OF DERIVATIVE =    3
NUMBER OF KNOTS     =   88

DEPENDENT VARIABLE TRANSFORMATION   =   2

NUMBER OF SPLINE INDEPENDENT VARIABLES =   4
NUMBER OF INDEPENDENT COVARIATES       =   1

VAR   LOWER LIMIT   UPPER LIMIT  TRANSF     UNITS   MARGINS
  1   150.500       154.000           0     deg     0.000       0.000
  2  -29.5000      -26.5000           0     deg     0.000       0.000
  3   0.00000       3000.00           1     m       0.000       0.000
   TRANSFORMATION CONSTANT  =   1000.00
  4   0.00000       5000.00           3             0.000       0.000
   TRANSFORMATION CONSTANTS =   1.00000       0.00000
  5   0.00000       1000.00           0             0.000       0.000

SURFACE NUMBERS (0 T0  1):
1

TRANSFORM SURFACE VALUES (0-NO, 1-YES):
    1

TYPE OF SURFACE CALCULATION (0-1):
  0 - SUMMARY STATISTICS ONLY
  1 - ALL SURFACE VALUES
    1

ERROR COVARIANCE FILE (BLANK IF NO ERRORS REQUIRED):
---------------------------------------------------
arrr\median_12h.cov
```

```
TYPE OF ERROR CALCULATION (0-4):
  0 - STANDARD ERROR OF AVERAGE ONLY
  1 - MODEL STANDARD ERRORS
  2 - PREDICTION STANDARD ERRORS
  3 - 95% MODEL CONFIDENCE INTERVALS
  4 - 95% PREDICTION CONFIDENCE INTERVALS
    2

MAXIMUM STANDARD ERROR (BLANK OR 1 STANDARD ERROR):
1000.0

GRID SPECIFICATIONS
-------------------
POSITION OPTION (0 - AT CELL CORNERS, 1 - AT CELL CENTRES):
    1

INDEX OF FIRST GRID VARIABLE (NORMALLY 1):
    1

LOWER LIMIT, UPPER LIMIT AND SPACING OF FIRST GRID VARIABLE:
   150.50000000000       154.00000000000      0.2500000E-01

INDEX OF SECOND GRID VARIABLE (NORMALLY 2):
    2

LOWER LIMIT, UPPER LIMIT AND SPACING OF SECOND GRID VARIABLE:
  -29.500000000000      -26.500000000000      0.2500000E-01

NUMBER OF COLUMNS =   140
NUMBER OF  ROWS   =   120

INPUT INDEPENDENT VARIABLE GRIDS
--------------------------------
MODE OF MASK GRID (0-3):
  0 - MASK GRID NOT SUPPLIED
  1 - GENERIC MASK GRID
  2 - ARC/INFO MASK GRID
  3 - IDRISI MASK IMAGE
    2
INPUT GRID FILE NAME:
arrr\mask.asc

MODE OF 3RD INDEPENDENT VARIABLE (0-3):
  0 - USER SUPPLIED CONSTANT
  1 - GENERIC INDEPENDENT VARIABLE GRID
  2 - ARC/INFO INDEPENDENT VARIABLE GRID
  3 - IDRISI INDEPENDENT VARIABLE IMAGE
    2
INPUT GRID FILE NAME:
arrr\dem.asc

MODE OF 4TH INDEPENDENT VARIABLE (0-3):
  0 - USER SUPPLIED CONSTANT
  1 - GENERIC INDEPENDENT VARIABLE GRID
  2 - ARC/INFO INDEPENDENT VARIABLE GRID
  3 - IDRISI INDEPENDENT VARIABLE IMAGE
    2
INPUT GRID FILE NAME:
arrr\MAR_025.asc

MODE OF 5TH INDEPENDENT VARIABLE (0-3):
  0 - USER SUPPLIED CONSTANT
  1 - GENERIC INDEPENDENT VARIABLE GRID
  2 - ARC/INFO INDEPENDENT VARIABLE GRID
  3 - IDRISI INDEPENDENT VARIABLE IMAGE
    2
INPUT GRID FILE NAME:
arrr\preds_12h.asc

OUTPUT SURFACE GRIDS
--------------------
MODE OF OUTPUT GRIDS (0-3):
  0 - X,Y,Z FORMAT
  1 - GENERIC GRID BY ROWS
  2 - ARC/INFO GRID
  3 - IDRISI IMAGE
    2
```

```
SPECIAL VALUE FOR OUTPUT GRIDS:
  -999.000

NAME OF OUTPUT GRID FILE FOR SURFACE 1:
arrr\median_12h.grd

OUTPUT ARC/INFO GRID FORMAT (BLANK FOR BINARY):
(10f9.3)

OUTPUT ERROR GRIDS
------------------
MODE OF OUTPUT GRIDS (0-3):
  0 - X,Y,Z FORMAT
  1 - GENERIC GRID BY ROWS
  2 - ARC/INFO GRID
  3 - IDRISI IMAGE
    2

SPECIAL VALUE FOR OUTPUT GRIDS:
  -999.000

NAME OF OUTPUT GRID FILE FOR SURFACE 1:
arrr\median_errors_12h.grd

OUTPUT ARC/INFO GRID FORMAT (BLANK FOR BINARY):
(10f9.3)

LAPGRD SUMMARY STATISTICS
-------------------------
OUTPUT SURFACE AND ERROR GRIDS FOR SURFACE   1
NUMBER OF CELLS =   14076      MEAN ERROR     =     6.3
MINIMUM ERROR   =     3.9      MAXIMUM ERROR  =     21.
MINIMUM VALUE   =    34.0      MAXIMUM VALUE  =    155.
MEAN SURF VALUE =    69.2      STANDARD ERROR =     1.0

PROGRAM LAPGRD  VERSION  4.37  DATE 29/08/2008  TIME 10.00.34
```

A2.3 Mapping Regional GEV Parameters (k and ß) for Subdaily Durations

Sample input data and ANUSPLIN command files

Sample input data – 6 hour duration, 0.4 prediction weighting

ß
Station, Longitude, Latitude, Beta_6h_Pnt4

```
40000,153.1,-27.95,0.2923
40004,152.711,-27.63,0.31002
40006,152.9167,-28.2833,0.31606
40007,152.9,-26.8,0.34839
40008,152.7,-27.4167,0.34227
40012,152.8,-28.25,0.29908
40014,153.013,-28.021,0.29951
40015,153.1903,-28.1467,0.32149
40016,153.1833,-27.7167,0.32624
40017,152.9764,-26.8564,0.35296
40019,152.15,-26.9,0.30937
40020,152.1006,-26.8844,0.31058
40024,152.6919,-27.9925,0.29501
40027,153.1658,-27.0892,0.34507
40030,152.4775,-28.2367,0.29982
40031,153.05,-26.6867,0.34839
40032,153.0667,-26.6833,0.348
40035,153.0089,-27.1414,0.34829
40037,152.5,-27.15,0.32853
.
.
```

k

Station, Longitude, Latitude, k_6h_Pnt4

```
40000,153.1,-27.95,-0.10061
40004,152.711,-27.63,-0.09793
40006,152.9167,-28.2833,-0.14092
40007,152.9,-26.8,-0.16422
40008,152.7,-27.4167,-0.03316
40012,152.8,-28.25,-0.15137
40014,153.013,-28.021,-0.07698
40015,153.1903,-28.1467,-0.09748
40016,153.1833,-27.7167,-0.19829
40017,152.9764,-26.8564,-0.16364
40019,152.15,-26.9,0.07072
40020,152.1006,-26.8844,0.10676
40024,152.6919,-27.9925,-0.14644
40027,153.1658,-27.0892,-0.0864
40030,152.4775,-28.2367,-0.07125
40031,153.05,-26.6867,-0.16422
40032,153.0667,-26.6833,-0.18069
40035,153.0089,-27.1414,-0.07848
40037,152.5,-27.15,-0.00054
.
.
.
```

Sample SPLINA command file (K and ß used same parameters)

```
Mapping k -6 hour 0.4 weighting
0
2
0
0
0
150.50  154.00  0  5  0.00
-29.50  -26.50  0  5  0.000
0
2
1
0
1
3
525.00
arrr\k_6h_pnt4.csv
1000
6

arrr\k_6h_pnt4.sur

arrr\k_6h_pnt4.cov
```

Sample LAPGRD command file (K and ß used same parameters)

```
arrr\k_6h_pnt4.sur
1
1
arrr\k_6h_pnt4.cov
0
1.0
1
1
150.50  154.00  0.025
2
-29.50  -26.50  0.025
2
arrr\mask.asc
2
-9.0
arrr\k_6h_pnt4.grd
(10f9.5)
```

A2.4 Mapping Daily Index Rainfall - Sample Inputs and Logs

Sample input data and ANUSPLIN command files

Sample input data- Median Rainfall 24-, 48-, 72-hour duration, 2year ARI

Station, Lon, Lat, Median_24h, Median_48h, Median_72h

```
40000,153.1000,-27.9500,99.36,134.75,138.57
40004,152.7111,-27.6297,95.45,124.65,129.9
40006,152.9167,-28.2833,95.68,130.15,140.28
40007,152.9000,-26.8000,187.22,254.3,301.42
40008,152.7000,-27.4167,113.96,145.19,156.54
40012,152.8000,-28.2500,94.65,123.76,138.57
40014,153.0131,-28.0206,86.14,113.33,128.72
40015,153.1903,-28.1467,140.3,199.19,217.26
40016,153.1833,-27.7167,123.57,160.62,170.56
40017,152.9764,-26.8564,153.41,212.9,257.34
40019,152.1500,-26.9000,85.33,111.89,123.05
40020,152.1006,-26.8844,83.26,97.79,106.47
40022,153.0833,-26.6000,167.55,202.46,222.88
40024,152.6919,-27.9925,86.08,108.28,117.65
40027,153.1658,-27.0892,131.45,173.05,187.79
40030,152.4775,-28.2367,129.95,170.61,196.56
40031,153.0500,-26.6867,169.86,217.5,250.54
40032,153.0667,-26.6833,181.47,238.26,301.31
40033,152.6000,-28.1000,91.42,124.1,131.29
40034,153.4500,-28.0833,151.92,226.38,264.4
.
.
.
```

Sample SPLINA command file (these are for 3 durations i.e. 3 surfaces)

```
Mapping Daily Index Rainfall
0  -99.0
2
0
0
0
150.50  154.00  0  5  0.00
-29.50  -26.50  0  5  0.00
2
3
3
0
0
3
506.00
arrr\Index_Rainfall_24_48_72_to_2006.csv
1000
6

arrr\to_2006.res
arrr\to_2006.opt
arrr\to_2006.sur
arrr\to_2006.lst
arrr\to_2006.cov
```

Sample LAPGRD command file (these are for 3 durations i.e. 3 surfaces)

```
arrr\to_2006.sur
0
1
1
arrr\to_2006.cov
0

1
1
150.50  154.00  0.025
```

```
2
-29.50  -26.50  0.025
2
arrr\mask.asc
2
-9.0
arrr\to_2006_24h.grd
arrr\to_2006_48h.grd
arrr\to_2006_72h.grd
(10f8.2)
```

A2.5 Mapping Regional GEV Parameters (k and ß) for Daily Durations

Sample input data and ANUSPLIN command files

Sample input data – 24-, 48-, 72-hour duration

ß

Station, Lon, Lat, Beta_24h, Beta_48h, Beta_72h

```
40000,153.1,-27.95,0.36004835,0.370544023,0.378361509
40004,152.7111,-27.6297,0.327985099,0.338371101,0.349731932
40006,152.9167,-28.2833,0.337079383,0.351098506,0.361765637
40007,152.9,-26.8,0.410041069,0.409891367,0.40544769
40008,152.7,-27.4167,0.366367779,0.382004477,0.39298101
40012,152.8,-28.25,0.327667443,0.344030164,0.352884273
40014,153.0131,-28.0206,0.356143716,0.366757778,0.3784175
40015,153.1903,-28.1467,0.404859596,0.33510941,0.347086208
40016,153.1833,-27.7167,0.360859484,0.381229433,0.376515468
40017,152.9764,-26.8564,0.408297777,0.412295628,0.408152259
40019,152.15,-26.9,0.320105216,0.325310537,0.323108981
40020,152.1006,-26.8844,0.315119532,0.318711192,0.31472561
40022,153.0833,-26.6,0.384464928,0.378802034,0.381841704
40024,152.6919,-27.9925,0.301666381,0.324516864,0.326858989
40027,153.1658,-27.0892,0.375040922,0.38452975,0.392252085
40030,152.4775,-28.2367,0.355026507,0.370272564,0.371693024
40031,153.05,-26.6867,0.395300655,0.387087305,0.393648805
40032,153.0667,-26.6833,0.392139332,0.382778133,0.388923358
40033,152.6,-28.1,0.309019347,0.331360531,0.327275586
40034,153.45,-28.0833,0.37422052,0.373973241,0.380727907
.
.
.
```

k

Station, Lon, Lat, k_24h, k_48h, k_72h

```
40000,153.1,-27.95,-0.103874344,-0.080283866,-0.095843802
40004,152.7111,-27.6297,-0.082835795,-0.11231523,-0.122190273
40006,152.9167,-28.2833,-0.095098557,-0.081184926,-0.040529375
40007,152.9,-26.8,-0.088827243,-0.093010356,-0.107432869
40008,152.7,-27.4167,-0.078330145,-0.107877227,-0.121013805
40012,152.8,-28.25,-0.081635301,-0.044357497,-0.011347512
40014,153.0131,-28.0206,-0.112610736,-0.094353028,-0.089126334
40015,153.1903,-28.1467,-0.073663401,-0.044816374,-0.063236152
40016,153.1833,-27.7167,-0.122337282,-0.118069506,-0.155277158
40017,152.9764,-26.8564,-0.092264029,-0.09539669,-0.11231523
40019,152.15,-26.9,-0.03100361,-0.0784805,-0.105357854
40020,152.1006,-26.8844,-0.041755179,-0.0865826,-0.113496986
40022,153.0833,-26.6,-0.067927702,-0.114973171,-0.115120727
40024,152.6919,-27.9925,-0.108765639,-0.150357484,-0.177117714
40027,153.1658,-27.0892,-0.094949474,-0.098673138,-0.101795528
40030,152.4775,-28.2367,-0.168239183,-0.12659567,-0.11231523
40031,153.05,-26.6867,-0.049704457,-0.087031735,-0.090919919
40032,153.0667,-26.6833,-0.047872841,-0.091517415,-0.097184526
40033,152.6,-28.1,-0.114382832,-0.149777869,-0.175831456
40034,153.45,-28.0833,-0.080584265,-0.034388957,-0.025142573
.
.
.
```

Sample SPLINA command file - K and ß used same parameters
- these are for 3 durations i.e. 3 surfaces

```
SPLINA Surface:beta  24,48,72 hours (to 2006)
0
2
0
0
0
150.5  154.0  0  5  0.0
-29.5  -26.5  0  5  0.0
0
2
3
0
0
3
561.0
arrr\beta_24_48_72_to_2006.csv
1000
6

arrr\beta.res
arrr\beta.opt
arrr\beta.sur
arrr\beta.lst
arrr\beta.cov
```

Sample LAPGRD command file - K and ß used same parameters
- these are for 3 durations i.e. 3 surfaces

```
arrr\beta.sur
0
1
arrr\beta.cov
0

1
1
150.5  154.0  0.025
2
-29.5  -26.5  0.025
2
arrr\mask.asc
2
-9.0
arrr\beta_to_2006_24h.grd
arrr\beta_to_2006_48h.grd
arrr\beta_to_2006_72h.grd
(10f9.5)
```

A2.6 Calculating Growth Factors from GEV Parameters in ArcGIS

1 + ([beta] / [k]) * (Pow(Ln(2), [k]) - Pow(- 1*Ln(F), [k]))

Where k ≠ 0

and F=(T-1)/T **non**-exceedence probability parameter
so F= 0.8 for 5y ARI,
0.9 for 10y ARI
0.95 for 20y ARI
0.98 for 50y ARI
0.99 for 100y ARI

<u>Examples:</u>
1 + ([beta_12hr] / [K_12hours]) * (Pow(Ln(2), [K_12hours]) - Pow(- 1 * Ln(0.8), [K_12hours])) ....
for 12h 5y

1 + ([beta_6hr] / [K_6hours]) * (Pow(Ln(2), [K_6hours]) - Pow(- 1 * Ln(0.98), [K_6hours]))
for 6h 50y

Where k = 0

1 + [beta] * (Ln(Ln(2)) - Ln(- 1*Ln(F)))

<u>Examples:</u>
1 + [beta_6hr] * (Ln(Ln(2)) - Ln(- 1 * Ln(0.98))) ....
for 6h 50y

1 + [beta_12hr] * (Ln(Ln(2)) - Ln(- 1 * Ln(0.8))) ....
for 12h 5y

The expression to put into the Raster Calculator uses the ‘Con’ function (Conditional evaluation): ***Con(<condition>, <if true expression>, <if false expression>)***.

<u>Examples:</u>
con([K_6hours] <> 0, 1 + ([beta_6hr] / [K_6hours]) * (Pow(Ln(2), [K_6hours]) - Pow(- 1 * Ln(0.98), [K_6hours])), 1 + [beta_6hr] * (Ln(Ln(2)) - Ln(- 1 * Ln(0.98)))) ....
for 6h 50y

con([K_12hours] <> 0, 1 + ([beta_12hr] / [K_12hours]) * (Pow(Ln(2), [K_12hours]) - Pow(- 1 * Ln(0.8), [K_12hours])), 1 + [beta_12hr] * (Ln(Ln(2)) - Ln(- 1 * Ln(0.8)))) ....
for 12h 5y

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
HRS11	A Pilot Study to Explore Methods for Deriving Design Rainfalls for Australia - Part 2	January 2009