**Natural Hazards
and Earth
System Sciences**

Non-stationarity in daily and sub-daily intense rainfall – Part 2: Regional assessment for sites in south-east Australia

D. Jakob[1,2]**, D. J. Karoly**[2]**, and A. Seed**[3]

[1]Australian Bureau of Meteorology, Melbourne, Australia
[2]School of Earth Sciences, University of Melbourne, Melbourne, Australia

Abstract. Using data for a common period (1976–2005) for a set of 31 sites located in south-east Australia, variations in frequency and magnitude of intense rainfall events across durations from 6 min to 72 h were assessed. This study was driven by a need to clarify how variations in climate might affect intense rainfall and the potential for flooding. Sub-daily durations are of particular interest for urban applications. Worldwide, few such observation-based studies exist, which is mainly due to limitations in data.

Analysis of seasonality in frequency and magnitude of events revealed considerable variation across the set of sites, implying different dominating rainfall-producing mechanisms and/or interactions with local topography. Both these factors are relevant when assessing the potential effects of climate variations on intense rainfall events. The set of sites was therefore split into groups ("north cluster" and "south cluster") according to the characteristics of intense rainfall events. There is a strong polarisation in the nature of changes found for the north cluster and south cluster. While sites in the north cluster typically exhibit decrease in frequency of events, particularly in autumn and at durations of 1 h and longer; sites in the south cluster experience an increase in frequency of events, particularly for summer and sub-hourly durations.

Non-stationarity found in historical records has the potential to significantly affect design rainfall estimates. An assessment of quantile estimates derived using a standard regionalisation technique and periods representative of record lengths available for practical applications show that such estimates may not be representative of long-term conditions, so alternative approaches need to be considered, particularly where short records are concerned. Additional rainfall information, in particular radar data, could be used for an in-depth spatial analysis of intense rainfall events.

1 Introduction

In a companion paper (Jakob et al., 2011) we outlined how design rainfall estimates are derived under the assumption of a stationary climate. Deriving design rainfall estimates under climate variability and climate change raises two questions: firstly, are climate variations driving changes in rainfall extremes that significantly affect design rainfall estimates; and secondly, what measures would be required to adapt to such changes? Mailhot and Duchesne (2010) discuss how to define design criteria under climate change. They state:

"[...] the question is: how can we define design criteria that take into consideration possible changes in the occurrence of intense rainfall events? In fact, the nonstationarity of climate conditions implies, in terms of definition of a design criterion, that (1) the expected lifetime of the infrastructure be explicitly considered when design criteria are defined; (2) the probability of exceeding the capacity and the associated risk, defined by the design criteria, will evolve with time; and (3) a statistical model describing the expected evolution of intense rainfall must be defined".

Due to the limited record lengths available (typically less than 30 yr), extrapolation to rare events introduces a high degree of uncertainty at sub-daily durations. Information from a group of sites is often pooled to allow derivation of design rainfall estimates for high return periods. The question of how to develop regionalisation techniques taking into account non-stationarity has been addressed by researchers but is not yet state-of-practice. Approaches for design rainfall estimation were recently revised for the UK, New Zealand, and the US (Institute of Hydrology, 1999; Thompson, 2002; Lin et al., 2004). Assessments of time series for non-stationarity

were undertaken but neither did the final analyses take non-stationarity into account nor did the choice of region address potential effects of non-stationarity. State-of-the-art approaches based on a hierarchical Bayesian framework have been developed (Lima and Lall, 2010) and Hanel et al., 2009 discuss a regionalisation approach for annual maxima of daily precipitation where the typical "index flood" approach is modified, so that all three parameters of the Generalised Extreme Value (GEV) distribution vary with time.

The purpose of this two-part study is to examine rainfall extremes for evidence of non-stationarity. For Part 1, analyses were undertaken for one Australian site with a long record (Sydney Observatory Hill, 85 yr) of rainfall measurements for durations between 6 min and 72 h. Using a set of 31 sites in the south-east of Australia for a concurrent period (1976–2005) allows the findings from analyses of a single long record to be set into a broader spatial context.

Part 2 explores how changes in frequency and magnitude of intense rainfall events vary with season and geographical location, and also whether such changes are consistent across groups of sites, giving greater confidence that changes identified as statistically significant are indeed climate-driven. Finally, we use a state-of-practice regionalisation approach to assess how non-stationarity in rainfall extremes may affect design rainfall estimates.

Data and methods are discussed in Sect. 2, and results are presented in Sect. 3. Summary and conclusions are covered in Sect. 4.

2 Data and methods

One of the major characteristics of Australian rainfall is a large degree of variability, spatially as well as across the seasons. Most stations in the study area do not experience a distinct dry or wet season. Average annual totals decrease with increasing distance from the coast. In south-east Australia, annual totals range from above 2000 mm to just above 200 mm. A large number of thunderstorms occurs in this region: 40 to 60 a year (Sturman and Tapper, 1996). Synoptic situations associated with intense rainfall events in this region include easterly dips, monsoon troughs, and tropical depressions.

2.1 Records of suitable lengths

A set of sites with "sufficiently complete" records for the period 1976 to 2005 was used to explore whether the results for Sydney Observatory Hill (presented in Part 1) are representative for a larger region.

Years are only accepted where less than 20 % of values are missing or accumulated, and at least 25 "acceptable" years are required for the 30-yr period. The location of sites is shown in Fig. 1 and details can be found in Table 1. The concurrent period (1976–2005) is too short to make definite
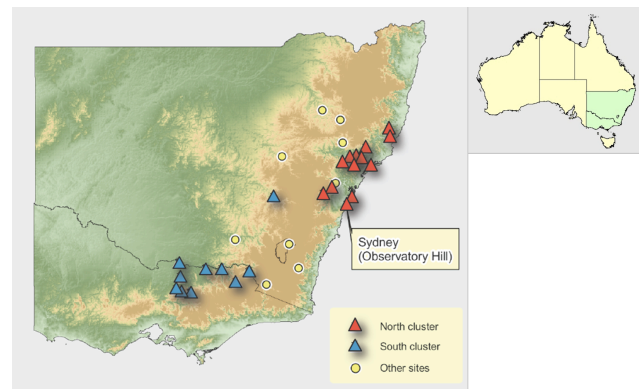
Fig. 1. Location of 31 sites with at least 25 yr of data for the period 1976 to 2005. For *Sydney Observatory Hill* the period 1921 to 2005 was analysed. (Symbols indicate cluster membership discussed in Sect. 3.2).

statements about long-term tendencies in changes in intense rainfall events. However, these analyses add valuable insights to the findings from the single long record presented in Part 1. Analysing both – a single site with a suitably long record and a group of sites for a concurrent period – has two major advantages: Firstly, findings for Sydney Observatory Hill can be put into context; and secondly, the spatial characteristics in changes in intense rainfall events across a range of durations can be explored.

2.2 Peaks over threshold

Design rainfall estimates are derived on the basis of extremal series, usually *annual maxima* (highest event in a given year) or *peaks over threshold* (POT). For the analyses presented in Part 1 and Part 2 of this paper, a peaks-over-threshold approach was chosen since it allows investigating changes in frequency as well as magnitude. The effects of climate variations are most readily detected for the highest thresholds (O'Gorman and Schneider, 2009). On the other hand, using very high thresholds may severely reduce sample sizes. The threshold most suited for a particular assessment will therefore depend on the type of analysis (parametric or non-parametric, seasonal or annual). Here, the threshold used for a set of analyses will be indicated using "POT3" to refer to a series with on average 3 events per year, etc. Seasonal series are derived by stratifying the annual series according to season.

2.3 Missing data and quality control

Missing data could potentially affect the analyses of changes in frequency and magnitude of events as well as quantile estimates. The potential effect of missing data was limited by screening sites (according to the percentage of missing and accumulated data) when selecting candidate sites.

Table 1. Station details for a set of sites with suitable data for the period 1976–2005. Last column indicates cluster membership (discussed in Sect. 3.2). "N" denotes "north cluster", "S" denotes "south cluster".

Station name	Station number	Latitude (° S)	Longitude (° E)	Elevation (m)	Cluster
Gunnedah Resource Centre	55024	31.03	150.27	307	
Gowrie North	55194	31.34	150.85	518	
Taree (Robertson St)	60030	31.90	152.45	21	N
Comboyne South	60080	31.63	152.44	670	N
Number One (Murrays Creek)	60106	31.66	152.06	120	
Williamtown RAAF	61078	32.79	151.84	9	N
Scone SCS	61089	32.06	150.93	216	
Chichester Dam	61151	32.24	151.68	194	N
Glendon Brook (Lilyvale)	61158	32.51	151.38	60	N
Colo Heights (Mountain Pines)	61211	33.36	150.71	320	
Pokolbin (Somerset)	61238	32.81	151.30	132	N
Paterson (Tocal AWS)	61250	32.63	151.59	30	N
Milbrodale (Hillsdale)	61309	32.69	150.97	120	N
Cowra Research Centre (Evans St)	63023	33.81	148.71	381	S
Katoomba (Murri St)	63039	33.71	150.31	1015	N
Kurrajong Heights (Bells Line Of Road)	63043	33.53	150.63	460	N
Wellington Research Centre	65035	32.51	148.97	390	
Sydney Airport AMO	66037	33.94	151.17	6	N
Sydney (Observatory Hill)	66062	33.86	151.21	39	N
Canberra Airport	70014	35.30	149.20	578	
Numeralla (Badja Composite)	70199	36.08	149.51	1040	
Ingebyra (Grosses Plains)	71042	36.60	148.47	1215	
Hume Reservoir	72023	36.10	147.03	184	S
Wagga Wagga AMO	72150	35.16	147.46	212	
Cobram (Goulburn Murray)	80109	35.91	145.64	113	S
Dookie Agricultural College	81013	36.37	145.70	185	S
Corryong (Parish Lane)	82011	36.20	147.90	314	S
Euroa	82016	36.75	145.57	175	S
Rutherglen Research	82039	36.10	146.51	175	S
Strathbogie	82042	36.85	145.73	502	S
Dartmouth Reservoir	82076	36.54	147.50	365	S
Lima South (Lake Nillahcootie)	82107	36.86	146.00	273	S

For daily (and longer) durations, techniques have been developed to infill missing data using data at neighbouring sites (Ladson, 2008). However, for durations as short as 6 min, such approaches are less suited, especially when investigating events that occur infrequently. Missing data might lead to spurious trends. For example, if data for a site were missing predominantly during the earlier part of the record, this might lead to an apparent increase in the frequency of events over time.

The problem of missing data was addressed using a combination of criteria:

– For the set of sites, a period of 30 yr (1976–2005) was used. At each station, years with 20 % or more of missing data were excluded from further analyses and sites were only included in the analyses if they had at least 25 yr with sufficient data for the 30-yr period. The

number of peaks to extract was adjusted to take the reduced record length into account. That is, for a POT3 series over a 30-yr period but with 5 yr of insufficient data, 75 rather than 90 peaks would be extracted.

– Only years with less than 20 % of missing data were included in the analyses. Rather than attempting to infill missing data or disaggregate rainfall accumulations, a simple approach was used to account for missing data when assessing changes in the frequency of events. This is done by adjusting the frequency for the period shortened due to missing data to the equivalent frequency for the complete period, e.g. for a season (or year) with 10 % of missing data, the frequency of events found for this season (year) would be multiplied by 1.11. This adjustment is based on the assumption that missing data occur randomly. It is less obvious how to adjust for

Table 2. Number of sites exhibiting significant increase or decrease in frequency/magnitude of events based on an average of 10 events per year (POT10) and a 90 % significance level. Numbers in bold indicate trends are collectively significant.

Duration		Frequency		Magnitude	
		Significant increase	Significant decrease	Significant increase	Significant decrease
short	6 min	**13**	0	3	0
	12 min	**8**	0	2	2
	18 min	**6**	0	**6**	3
	30 min	5	0	4	2
intermediate	1 h	2	0	0	2
	2 h	3	1	4	3
	3 h	2	1	4	4
	6 h	3	3	0	3
long	12 h	1	4	1	1
	24 h	1	2	0	3
	48 h	2	3	1	3
	72 h	3	2	1	4
Total		49	16	26	30

missing data when analysing changes in magnitude of events over time and therefore no adjustment has been undertaken. Quantile estimates for a set of sites were derived using a regional approach, minimising the effect of missing data.

The extracted POT series were examined for obvious outliers. Sites were screened using L-moment ratio diagrams (Hosking and Wallis, 1997), average magnitude of POT3 series, and seasonality index (Institute of Hydrology, 1999). Despite the fact that site 60106 (Number One, Murrays Creek) had sufficient data to be included in the set of sites, it was deemed not to be representative of conditions in the region and therefore excluded from further analyses. The site is located in a valley, sheltered from all directions except south, which may be the reason that this site exhibits seasonality (with regards to frequency and magnitude of events) markedly different from any of the other sites in the set.

3 Results

By using a set of sites with a common period, it is possible not only to explore how changes in frequency and magnitude of rainfall events vary with season and geographical location but also whether such changes are consistent across groups of sites, providing greater confidence that changes are not artefacts but are climate-related.

Sites were grouped according to characteristics of intense rainfall events, and changes in seasonal frequency and magnitude for these groups are discussed. A regionalisation approach was used to summarise these findings and expose differences between changes in quantile estimates for these

groups. To set the scene, examples for changes in frequency and magnitudes of events are discussed. The basis of these analyses are POT series abstracted for a set of 31 sites for the period 1976–2005.

For ease of reference, we will in the following refer to three sets of durations: short durations (sub-hourly durations), intermediate durations (1 h up to 6 h) and long durations (12 h and above).

3.1 Assessing changes in frequency and magnitude of events at sites

Changes in annual frequency and magnitude of events were assessed on the basis of POT10 series. A Mann-Kendall test was used to assess whether statistically significant changes in frequency and magnitude of events could be identified for the period 1976 to 2005. A Wilcoxon test was used to assess the statistical significance of changes in annual frequency and magnitude from an earlier to a later period (1976–1990 to 1991–2005).

In assessing the collective significance of a finite set of individual significance tests, the effects of finiteness of the sample and interdependence have to be accounted for. Livezey and Chen (1983) suggest a two-step approach to separately assess the effects of finiteness and interdependence.

Following this approach, the probability of passing $M \geq M_0$ out of N tests at a given significance level p was assessed. Here N denotes the number of sites (31) and M_0 denotes the number of sites passing the test for a given significance level (here 90 %). The cumulative probability that 5 or more out of 31 sites pass a test at the 90 % level is calculated using the binomial probability distribution as 19.3 %.
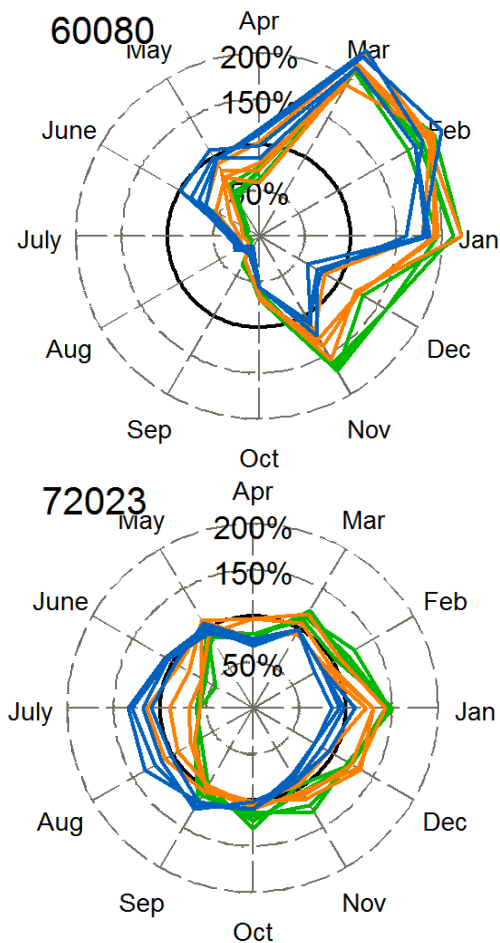
Fig. 2. Relative frequency of events (by calendar month) for 60080, Comboyne South (left) and 72023, Hume Reservoir (right). Colour denotes duration: green – short durations (6, 12, 18, and 30 min), orange – intermediate durations (1, 2, 3, and 6 h), blue – long durations (12, 24, 48, and 72 h).

The cumulative probability that 6 or more sites pass testing at the 90 % level is 8.3 %. Therefore, when taking into account the finiteness of the sample, at least 6 sites have to pass the test for the result to be collectively significant.

Resampling was undertaken to assess how interdependence affects the collective significance of results. Livezey and Chen (1983) specify that: (a) "First, the random component must be introduced in such a way as to retain the interdependence whose effects need to be evaluated." (b) "Second, probabilities of chance outcomes under the null hypothesis in the actual experiment must be matched in the Monte Carlo simulation." To maintain the interdependence between records at the 31 sites, the frequency of events was collectively resampled. The assessment of the effect of interdependence was undertaken using the steps outlined below:

1. Make a note of the frequency of POT events at each site and for each year.

2. Create a large number of random samples (1000) as permutations of the years 1976 to 2005 using a balanced resampling approach.

3. For each of these resamples, test for significant trends at each station and note the number of significant positive and negative trends at the 90 % significance level.

4. Use the 90th percentile of the number of sites identified as having statistically significant positive/negative trends (across all 1000 resamples) to derive an estimate of the number of sites that are identified by chance as having significant trends.

It was found that between 4 and 5 sites were identified by chance as having a statistically significant trend. Taking into account the effects of finiteness and interdependence, findings are judged collectively significant where 6 or more sites pass tests at the 90 % significance level.

### 3.1.1 Changes in annual frequency and magnitude of events

For the short (sub-hourly) durations, up to one third of the 31 sites exhibit significant increase in the *frequency* of events for the period 1976 to 2005 (see Table 2). No significant decrease is found for the short durations. For the intermediate and the long durations, a small number of sites do exhibit significant decrease in the frequency of events. Few sites show significant increase in frequency of events across the whole range of durations; one of these sites is Dartmouth Reservoir, the dam with the highest embankment in Australia.

Across all durations (6 min to 72 h) the number of sites with significant increase and decrease in *magnitude* is very similar (26 and 30, respectively). The overall pattern is similar when the significance level is increased to 95 %. However, for the increased significance level, twice as many significant decreases as increases (19 and 9, respectively) are identified. A number of sites show a peak in magnitudes around the 1990s (which ties in well with a number of very wet summers between 1991 and 1995) but no clear pattern emerges for the majority of sites over the short period considered here. Sites like Dartmouth Reservoir, where the magnitude of the ten highest-ranking events has consistently increased, are the exception.

3.1.2 Seasonality

When estimating design floods from design rainfalls, a number of factors need to be taken into consideration that are likely to vary with season, such as antecedent conditions (e.g. soil moisture, water table). It was investigated whether intense rainfall events for the group of sites exhibit seasonality and if so, whether seasonality varied with duration and geographical location.

The following analyses are based on POT10 series for the concurrent period 1976–2005. Two contrasting examples for
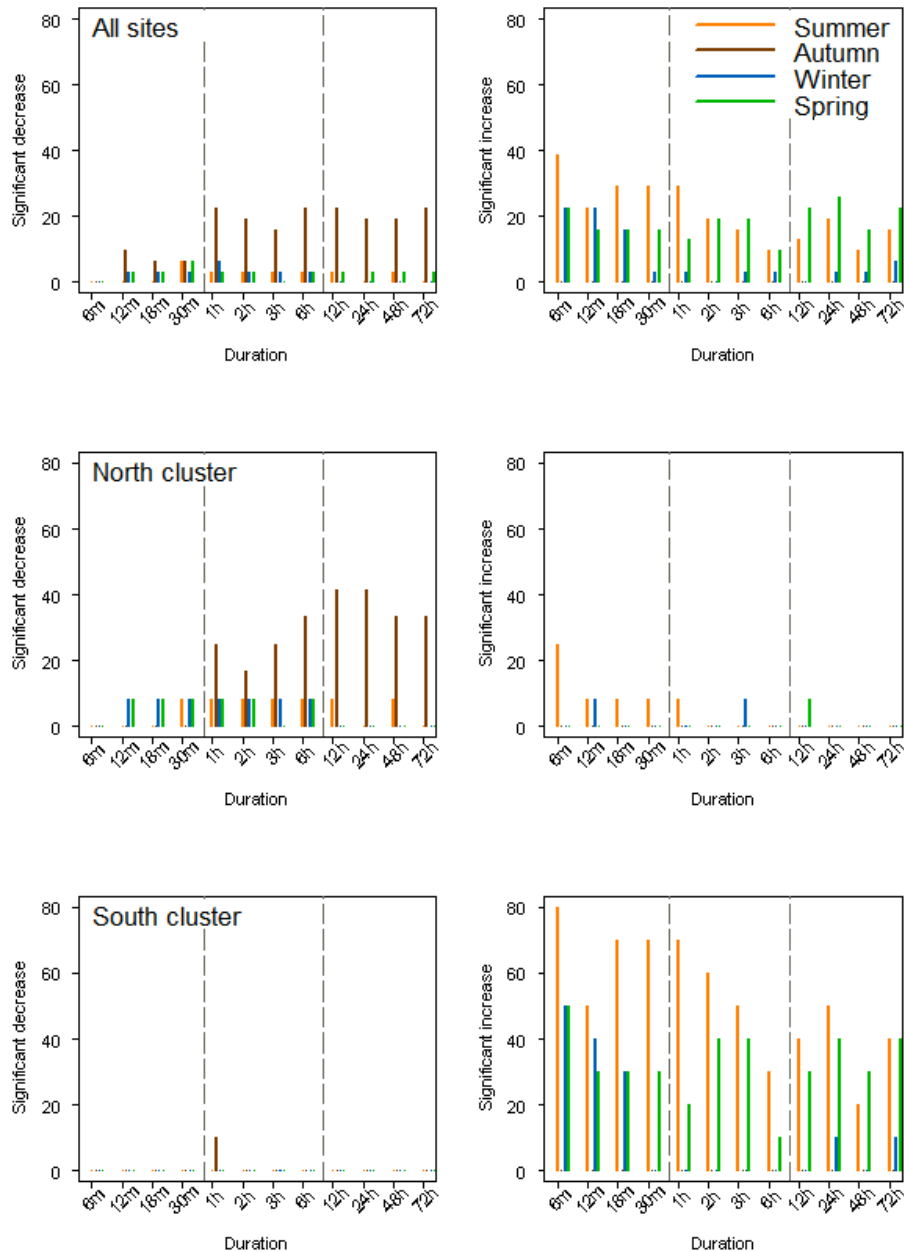
Fig. 3. Percentage of sites with significant changes in frequency from an earlier to a later period (1976–1990 and 1991–2005) based on 12 sites in the north cluster and 10 sites in the south cluster, using POT10 series.

seasonality in frequency of events are shown in Fig. 2. Regardless of duration, events at a site in the north of study area, Comboyne South (site 60080), occur much more frequently during the warmer months (November to March), with a peak for the months January to March. Below average frequencies are found for the cooler months and numbers are particularly low for July to September. For a site further south (Hume Reservoir, site 72023), seasonality is much weaker and depends on duration: most of the short duration peaks occur in summer and early autumn (December to March),

while most of the long-duration peaks occur during winter and early spring (June to September).

Most sites exhibit highest average peaks in summer, with lower magnitudes for the cooler months. Sites with significant orographic enhancement of rainfall exhibit a maximum in average magnitudes (especially for long durations) during winter (June to August), while sites further south and located in the rain shadow show decreased magnitudes for the cooler months when short and medium durations are considered but no discernable annual cycle is found for the longer durations.

Analysis of seasonality in frequency and magnitude of events reveals considerable variation across the set of sites, implying different dominating rainfall-producing mechanisms and/or interactions with local topography. Both these factors are relevant when assessing the potential effects of climate variations on intense rainfall events. The set of 31 sites was therefore split into groups according to characteristics of intense rainfall events.

3.2 Spatial assessment

Sites were grouped according to two measures:

1. the median magnitude of events based on the POT3 series for the set of 12 durations; and

2. a seasonality index (Institute of Hydrology, 1999) for the 12 durations, consisting of the day of the year at which events are most likely to occur and a measure of "concentration" indicating the spread around this day.

Two distinctly different groups of sites were identified (see Table 1): a group of 12 sites (including Sydney Observatory Hill) north of 34° S and on the windward side of the mountains (for easterly wind directions) and a group of 10 sites further south (roughly between 36 and 37° S) located in the rain shadow of the mountain ranges for easterly and southeasterly directions. These two groups will in the following be referred to as "north cluster" and "south cluster", respectively. Rainfall events for sites in the north cluster tend to occur mainly during the warmer months (for example Comboyne South, Fig. 2) and have larger average magnitudes than for sites in the south cluster, where events are spread more evenly throughout the year (for example Hume Reservoir). Differences in average frequency are largest for the shortest durations (about 20 % for durations between 6 and 18 min).

A regional approach allows summarising changes for each cluster and minimises the effect of missing data. This approach also enables us to consider changes at return periods exceeding the 15-yr periods used in the comparisons. Quantile estimates for a set of key return periods (2, 5, 10, 20 and 50 yr) were derived using an L-moments approach and assuming a Generalised Pareto Distribution (GPD). More sophisticated approaches have been developed, for instance a Bayesian approach (Ribatet et al., 2007). Such an approach will be considered for future work since it allows for uncertainty in estimating the "index rainfall" as well as the distributional parameters. A Wilcoxon test was used to assess the statistical significance of changes in frequency and magnitude from an earlier to a later period (1976–1990 to 1991–2005). Results for the number of statistically significant changes in the frequency of events are summarised in Fig. 3. Accounting for the finite sample (Livezey and Chen, 1983), changes in frequency are collectively significant for the full set of 31 sites if 20 % or more of sites exhibit significant changes. For the modified north cluster and the south cluster, the relevant threshold is 40 % of sites.

Table 3. Values for the homogeneity measure H_1 for south and north cluster (original and modified to improve homogeneity). Values above 2 ("definitely heterogeneous") are highlighted in bold.

Duration		South cluster	North cluster	Modified north cluster
short	6 min	−0.64	1.17	0.04
	12 min	−0.35	1.79	−0.3
	18 min	0.26	**2.33**	−40.61
	30 min	0.30	**2.63**	0.22
intermediate	1 h	0.64	**2.64**	0.47
	2 h	**2.01**	**2.28**	1.69
	3 h	1.83	1.92	1.43
	6 h	1.95	1.95	1.81
long	12 h	1.85	1.60	1.60
	24 h	1.21	1.11	1.25
	48 h	**2.02**	0.91	1.30
	72 h	1.73	0.68	0.50

Considering all sites, decrease in the number of events is confined almost exclusively to autumn and occurs mainly for the intermediate and long durations (at about 20 % of sites). Significant increase occurs mainly during summer and spring. Up to 40 % of sites (e.g. at 6 min) are showing an increase in the frequency of events for the summer months.

For the north cluster, significant decrease dominates, occurring mainly in autumn and for intermediate and long durations. Significant increase is found for about 10 % of locations for summer and short durations. For the south cluster, significant increase dominates and up to 80 % of sites show significant increase (summer, 6-min duration) while only one site shows significant decrease (autumn, 1-h duration). In addition to testing for changes in frequency from an earlier to a later period, changes across the whole 30-yr period were assessed, using a Mann-Kendall test. There is generally good agreement between these two assessments. However, the number of sites with significant decrease (particularly for the north cluster) is lower when trends over the full period are analysed.

In summary, there is a strong polarisation in the nature of changes found for the north and south clusters. While sites in the north typically exhibit decrease in frequency, particularly in autumn and at medium and long durations; sites in the south experience an increase in frequency of events, particularly for summer and short durations but also for spring (across the range of durations). A similar assessment was undertaken for the statistical significance of changes in magnitudes. Typically less than 20 % of sites exhibit significant changes for a given duration. The only apparent tendency is for an increase in magnitude particularly in spring at a small number of sites in the south cluster but across the range of durations.
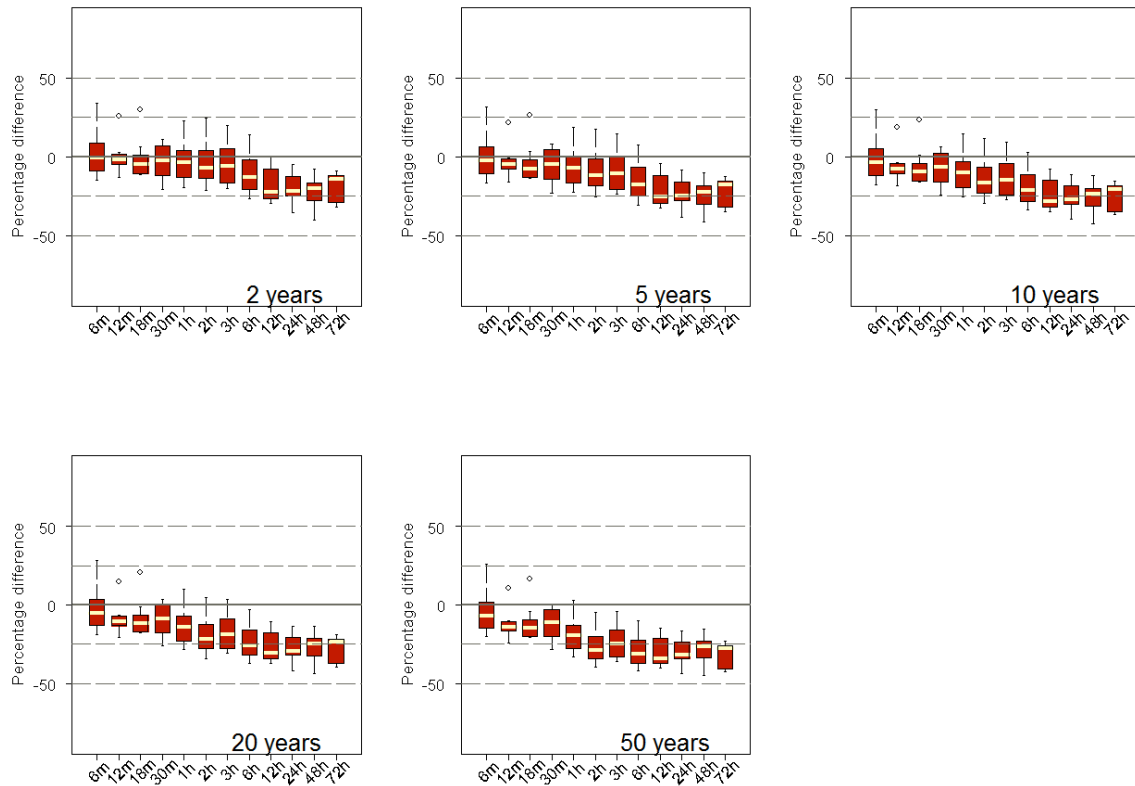
Fig. 4. Percentage difference between quantile estimates based on POT3 series for sites in the modified north cluster for five different return periods (2, 5, 10, 20, and 50 yr). Negative values indicate a decrease from earlier to later period.

Thompson (2006) studied the relationship of negative and positive Interdecadal Pacific Oscillation (IPO) phases with annual maximum rainfall in New Zealand for durations of 1, 12, and 24 h. No discernable effect was found for the 1-h duration, which was ascribed to convective processes dominating precipitation at this duration. However, for the 24-h duration, the IPO phase can have significant effect. While the north and east of the North Island experience an increase in the magnitude of annual maxima during the negative IPO phase, the opposite is found for the west and south of the South Island. For Sydney Observatory Hill, the period with above average quantile estimates for the long durations (24 h and above) roughly matches the period of the IPO negative phase from 1945 to 1977 (Fig. 4 in Part 1). The period considered for the group of sites (1976–2005) falls almost entirely within an IPO positive phase.

3.3 Regional frequency analysis

For Part 1, empirical quantile estimates were derived based on 10-yr sliding windows. This approach is less suited to a set of sites for a 30-yr period. Instead, changes in regional quantile estimates have been assessed for an earlier and a later period based on data for the north cluster and south cluster, respectively.

For the following, a threshold of on average 3 events per year (POT3) was chosen as a "sufficiently high threshold" to justify the assumption that the distribution function of excess values converges to a Generalised Pareto Distribution (GPD). However, this precludes us from assessing seasonal changes in quantile estimates. For practical applications, one is often faced with the problem that the available records are too short to derive design rainfall estimates for high return periods with the required level of confidence. A standard way to approach this issue is by applying concepts of *regional frequency analysis* (Hosking and Wallis, 1997). The fundamental assumption underlying this approach is that the frequency curves for sites in a region are identical apart from a scaling factor ("index flood" or here, "index rainfall").

The following describes the procedure for these assessments:

1. Based on the POT3 series, derive at-site L-moments (linear combination of moments) for the entire 30-yr period (1976–2005). Adjust the number of peaks to be extracted by taking into account the percentage of missing data.

2. For each site, derive the "index rainfall" as the median of the POT3 series.
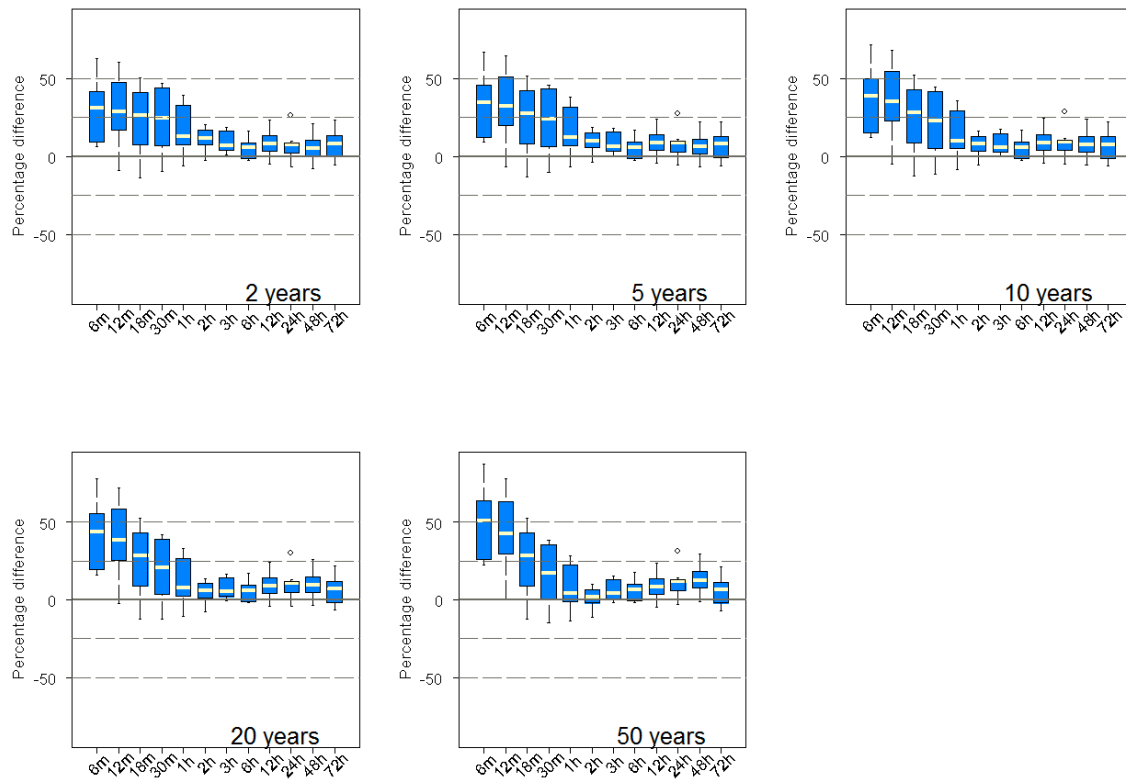
Fig. 5. As Fig. 4 but for the south cluster. Positive values indicate an increase from the earlier to the later period.

3. For each region, derive estimates of regional L-moments (as weighted average of site L-moments).

4. Derive the regional standardised frequency curves (growth curves) for each region and duration.

5. Derive quantile estimates for a set of key return periods (2, 5, 10, 20, and 50 yr) for each station by multiplying the at-site "index rainfall" by the relevant regional growth factors.

6. Repeat the procedure for two 15-yr periods (1976–1990 and 1991–2005) and calculate percentage differences.

7. Assess the significance of changes using a Wilcoxon rank sum test also known as Mann-Whitney test (Haan, 2002).

When deriving regional estimates, it should be assessed whether regions can be considered homogeneous. Such an assessment has been undertaken on the basis of tests developed by Hosking and Wallis (1997) assuming a 4-parameter kappa distribution. Here we use the measure H_1, which is based on L-CV and L-skewness. Based on the H_1 statistics, the north cluster was judged "definitely heterogeneous" for durations from 18 min to 2 h (Table 3). Regions can be modified (by removing or reassigning sites) to improve (statistical) homogeneity. This may imply a reduction in region size

as well as possibly different regions at different durations. To increase the homogeneity for the north cluster, 4 sites were removed. These sites were "least similar" to other sites in the north cluster (with respect to their L-moments).

Quantile estimates were derived for the north and south clusters for the set of 12 durations and 6 return periods. Differences in quantile estimates for two 15-yr periods (1976–1990 and 1991–2005) and against the full 30-yr period (1976–2005) were assessed for their statistical significance. The boxplots in Figs. 4 and 5 summarise the differences for the (modified) north cluster and south cluster, respectively, for a range of return periods. Filled boxes indicate the interquartile range; the dividing line represents the median. Values outside 1.5 times the interquartile range are judged outliers and are shown as open circles. To aid comparison between clusters and return periods, the same scale is used regardless of return period and cluster. Quantile estimates for the (modified) north cluster for the later period (1991–2005) are predominantly lower than for the earlier period (1976–1990). Differences in quantile estimates from the earlier to the later period tend to increase with duration and return period. The largest differences are therefore found for a return period of 50 yr and durations of 12 h and longer (a decrease of about 30 % compared to about 10 % for the 2-yr return period). In a similar vein, for the south cluster the increase in quantile estimates from the earlier to the later period is
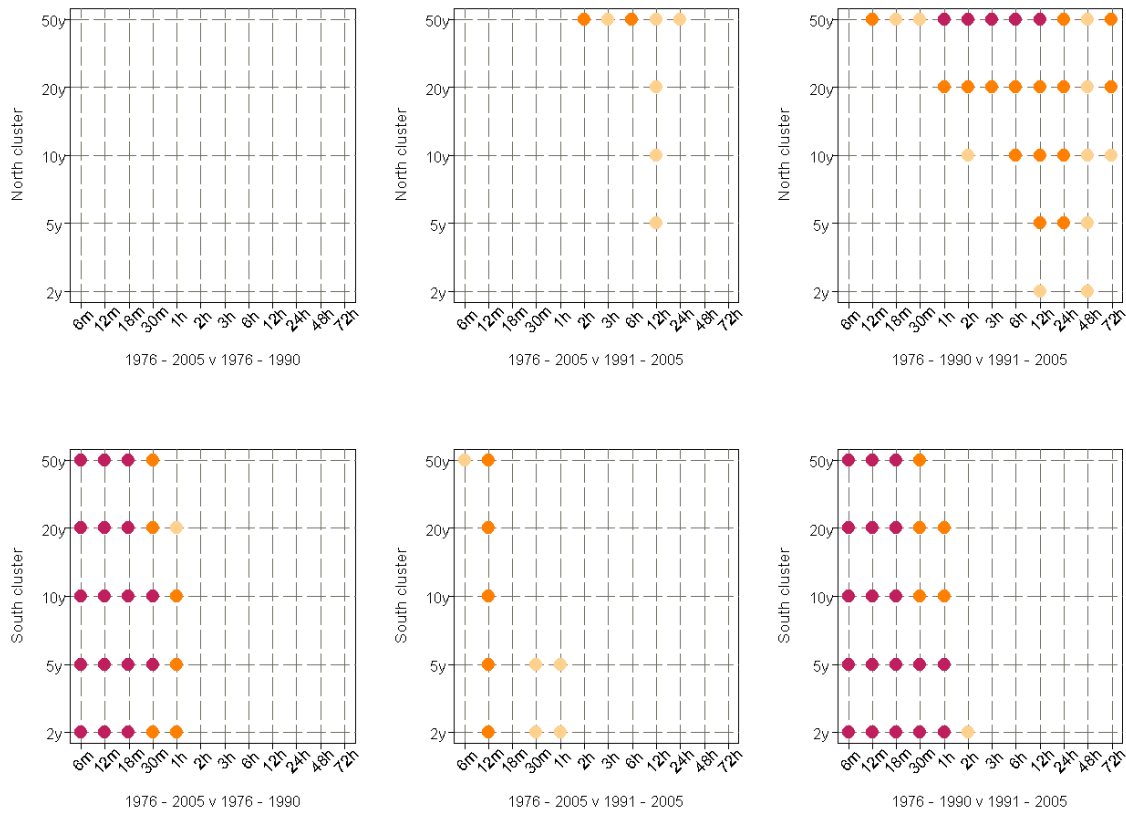
Fig. 6. Significance of difference in quantile estimates from regional analysis based on POT3 series for 12 durations (x-axis) and 5 return periods (y-axis). The top row shows the results for the modified north cluster (8 sites), bottom row for the south cluster (10 sites).

largest for the shortest durations and the highest return period (about 25 % increase for the 2-yr return period but reaching about 80 % for the 6-min duration for 50 yr).

Figure 6 gives an indication of the significance of differences in quantile estimates between periods. The significance of changes was assessed using a Wilcoxon rank sum test also know as Mann-Whitney test (Haan, 2002). Results for the (modified) north cluster are shown in the top row, for the south cluster in the bottom row. No significant differences in quantile estimates were found for the modified north cluster between the full period (1976–2005) and the earlier of the 15-yr periods (1976–1990), and the top left panel of Fig. 6 is therefore blank.

For the south cluster, comparison of quantile estimates based on the full 30-yr period with estimates derived on the basis of only the first half of the full period reveal that estimates for the shorter period are significantly different (lower) for the four shortest durations and across the range of return periods (Fig. 6, bottom right panel). For neither of the two clusters do estimates derived using only the last 15 yr typically differ significantly from those based on the full 30 yr. There are, however, significant differences between estimates derived from the two 15-yr periods. For the north cluster, this affects particularly durations of 6 h and above.

At the highest return period (50 yr) the majority of durations exhibit significant differences. For the south cluster, significant differences are found for durations up to 1 h and across the range of return periods.

The assessment was undertaken using standard regionalisation techniques and periods representative of record lengths available for practical applications. The analyses show that despite using a currently widely accepted approach, derived quantile estimates may not be representative of long-term conditions and alternative approaches need to be considered. For instance, Ribatet et al. (2007) describe a regional Bayesian POT model for flood frequency analysis and discuss the superior performance of such an approach, particularly where short records are concerned. The performance of such an approach for the data used in this study has yet to be assessed.

The distinct differences found between the north and south clusters highlight the need to identify regions suitable for the purpose of analysing changes in intense rainfall events. Period of record effects may be particularly notable for certain subsets of durations, e.g. for the north cluster at long durations and for the south cluster at short durations.

4 Summary

Variations in frequency and magnitude of intense rainfall events across durations from 6 min to 72 h were assessed using data from sites in the south-east of Australia. This study was driven by a need to clarify how variations in climate might affect intense rainfall and the potential for flooding. Sub-daily durations are of particular interest for urban applications. Worldwide, few such observation-based studies exist, which is mainly due to limitations in data. Here the problem was addressed by supplementing one long record (discussed in Part 1) by a set of shorter records for a 30-yr concurrent period.

Analysis of seasonality in frequency and magnitude of events reveals considerable variation across the set of sites, implying different dominating rainfall-producing mechanisms and/or interactions with local topography. Both these factors are relevant when assessing the potential effects of climate variations on intense rainfall events. The set of sites was therefore split into groups according to characteristics of intense rainfall events.

For the sub-hourly durations, up to one third of sites exhibit significant increase in the number of events. No significant decrease is found for the short durations. For the intermediate and the long durations, a small number of sites do exhibit a decrease in the frequency of events. Few sites show significant increase in frequency of events across the whole range of durations. Using the same level of significance as for the changes in frequency, for magnitudes a smaller number of significant changes are identified. Sites where the magnitude of the highest events (say the top 10) has consistently increased are the exception. A number of sites show a peak in magnitudes around the 1990s but no clear pattern emerges for the majority of sites. Considering all 31 sites, decrease in the number of events is confined almost exclusively to autumn and occurs mainly for the intermediate and long durations (at about 20 % of sites). Significant increase occurs mainly during summer and spring. Up to 40 % of sites (e.g. for the 6-min duration) are showing an increase in the frequency of events for the summer months. There is a strong polarisation in the nature of changes found for the north and south clusters. While sites in the north typically exhibit decrease in frequency, particularly in autumn and at medium and long durations; sites in the south experience an increase in frequency of events, particularly for summer and short durations but also for spring (across the range of durations).

Quantile estimates for the (modified) north cluster for the later period (1991–2005) are predominantly lower than for the earlier period (1976–1990). Differences tend to increase with duration and return period. The largest differences are therefore found for a return period of 50 yr and durations of 12 h and longer (a decrease of about 30 % compared to about 10 % for the 2-yr return period). For the south cluster, the increase in quantile estimates from the earlier to the later period is largest for the shortest durations and the highest re-turn period (about 25 % increase for the 2-yr return period but reaching about 80 % for the 6-min duration for 50 yr).

At least for some regions of Australia, non-stationarity found in historical records has the potential to significantly affect design rainfall estimates. Additional rainfall information, in particular radar data, could be used for an in-depth spatial analysis of rainfall events. How climate variations affect intense rainfall events (in terms of tendency and magnitude of changes) strongly depends on the duration and can vary across relatively small scales (hundreds of km). Further work should therefore focus on identifying appropriate techniques for estimating design rainfall taking into account non-stationarity. How non-stationarity is addressed in design rainfall estimation is likely to depend on the application (risks involved as well as lifespan of structures to be designed). Considerations may have to include suitable choices of period of record as well as record lengths, identifying regions of suitable size but, possibly more importantly, whether it is required and feasible to use state-of-the-art statistical approaches for practical applications.

References

Haan, C. T.: Statistical Methods in Hydrology, 2nd Edn., Iowa State Press, 496 pp., 2002.

Hanel, M., Buishand, T. A., and Ferro, A. T.: A nonstationary index flood model for precipitation extremes in transient regional climate model simulations, J. Geophys. Res., 114, D15107, doi:10.1029/2009JD011712, 2009.

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

Institute of Hydrology: Flood Estimation Handbook, Institute of Hydrology, Wallingford, 3, 178–180, 1999.

Jakob, D., Karoly, D. J., and Seed, A.: Non-stationarity in daily and sub-daily intense rainfall – Part 1: Sydney, Australia, Nat. Hazards Earth Syst. Sci., 11, 2263–2271, doi:10.5194/nhess-11-2263-2011, 2011.

Ladson, A.: Hydrology: An Australian Introduction, Oxford University Press, Melbourne, 326 pp., 2008.

Lima, C. H. R. and Lall, U.: Spatial scaling in a changing climate: A hierarchical Bayesian model for non-stationary multi-site annual maximum and monthly streamflow, J. Hydrol., 383, 307–318, doi:10.1016/j.jhydrol.2009.12.045, 2010.

Lin, B., Bonnin, G., Todd, D., Parzybok, T., Yekta, M., and Riley, D.: 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.

Livezey, R. E. and Chen, W. Y.: Statistical field significance and its determination by Monte Carlo techniques, Mon. Weather Rev., 111, 46–59, 1983.

Mailhot, A. and Duchesne, S.: Design Criteria of Urban Drainage Infrastructure under Climate Change, J. Water Res. Pl., 388(3–4), 173–185, doi:10.1061/ASCEWR.1943-5452.0000023, 2010.

O'Gorman, P. A. and Schneider, T.: Scaling of precipitation extremes over a wide range of climates simulated with an idealized GCM, J. Climate, 22, 5676–5685, 2009.

Ribatet, M., Sauquet, E., Gresillon, J.-M., and Ouarda, T.: A regional Bayesian POT model for flood frequency analysis, Stoch. Enviro. Res. Ris. Assess., 21, 327–339, doi:10.1007/s00477-006-0068-z, 2007.

Sturman, A. and Tapper, N.: The weather and climate of Australia and New Zealand, Oxford University Press, Melbourne, 476 pp., 1996.

Thompson, C. S.: The High Intensity Rainfall Design System: HIRDS. Proceedings International Conference on Flood Estimation, March, 2002, Berne, Switzerland. International Commission for the Hydrology of the Rhine Basin, CHR Report II-17, 273–282, 2002.

Thompson, C. S.: Decadal climate variability of extreme rainfalls in New Zealand, Weather Climate, 26, 3–20, 2006.