

Non-stationarity in daily and sub-daily intense rainfall – Part 1: Sydney, Australia

D. Jakob^{1,2}, D. J. Karoly², and A. Seed¹

¹Australian Bureau of Meteorology, Melbourne, Australia

²School of Earth Sciences, University of Melbourne, Melbourne, Australia

Received: 3 April 2011 – Revised: 24 June 2011 – Accepted: 8 July 2011 – Published: 19 August 2011

Abstract. 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. While there are still large discrepancies between precipitation data sets from observations and models, both show that there is a tendency for moist regions to become wetter and for dry regions to become drier. However, changes in extreme conditions may show the opposite sign to those in average conditions. Where changes in observed intense precipitation have been studied, this has typically been for daily durations or longer.

The purpose of this two-part study is to examine daily and sub-daily rainfall extremes for evidence of non-stationarity. Here the problem was addressed by supplementing one long record (Part 1) by a set of shorter records for a 30-yr concurrent period (Part 2). Variations in frequency and magnitude of rainfall extremes across durations from 6 min to 72 h were assessed using data from sites in the south-east of Australia. For the analyses presented in this paper, a peaks-over-threshold approach was chosen since it allows investigating changes in frequency as well as magnitude. Non-parametric approaches were used to assess changes in frequency, magnitude, and quantile estimates as well as the statistical significance of changes for one station (Sydney Observatory Hill) for the period 1921 to 2005. Deviations from the long-term average vary with season, duration, and threshold. The effects of climate variations are most readily detected for the highest thresholds. Deviations from the long-term average tend to be larger for frequencies than for magnitudes, and changes in frequency and magnitude may have opposite signs.

Investigations presented in this paper show that variations in frequency and magnitude of events at daily durations are a poor indicator of changes at sub-daily durations. Studies like the one presented here should be undertaken for other regions to allow the identification of regions with significant increase/decrease in intense rainfall, whether there are common features with regards to duration and season exhibiting most significant changes (which in turn could lead to establishing a theoretical framework), and assist in validation of projections of rainfall extremes.

1 Introduction

Building standards are used in the design of structures such as bridges, dams, and sewers. These standards are based on estimates of design values, for example buildings are designed to withstand a certain wind load, and bridges have to be able to withstand certain flood levels. Depending on the nature of the structure and the intended life span, estimates are required for durations ranging from a few minutes out to a number of days, and for fairly frequent events (say, events occurring on average once a year) to much rarer events (for instance, events occurring on average once in a hundred years). Current design rainfall estimates for Australia have been derived by fitting extreme value distributions under the assumption of a stationary climate. Over the last decade or so, research into the effects of climate variations on rainfall has grown dramatically, and more recently the focus has started to shift to rainfall extremes. The phrase “rainfall extreme” is used with different meanings in different communities. It can refer to the rarity, the magnitude, or the impact of an event. Hydrologists might refer to an event as extreme if it occurs once in a thousand years or less frequently. Here, we use “rainfall extreme” to indicate that events are part of an extremal series (like annual maxima or peaks-over-threshold)



Correspondence to: D. Jakob
(d.jakob@bom.gov.au)

suitable for extreme value analysis. Depending on threshold, such events would be expected to occur a few times a year.

Concern is starting to grow amongst engineers over whether the currently valid design rainfall estimates are affected by climate change. Changes in rainfall are amplified in streamflow, for Australia typically by a factor of 2 to 3 (Chiew, 2006), and it has already been shown that climate variability can affect flood risk on multi-decadal time scales (for instance through effects of the Interdecadal Pacific Oscillation on rainfall).

Temperature records indicate an increase in global temperature of about 0.7 °C over the last 100 yr (IPCC, 2007). Most of this warming has been attributed to greenhouse gas emissions from human activity (e.g. burning of fossil fuels). According to the Clausius-Clapeyron relation, for each 1 °C increase in global temperature, the precipitable water increases by about 7 %, but models show that the associated increase in *average* rainfall is typically considerably lower (about 2 % per 1 °C). However, Lenderink and van Meijgaard (2008) show that for a site in the Netherlands (de Bilt), for temperatures above 12 °C the increase in magnitude for intense rainfall events (99th percentile) at the 1-h duration could be as high as 14 % per 1 °C temperature increase. According to IPCC (2007), “Extremes of daily precipitation are very likely to increase, except possibly in areas of significant decrease in mean rainfall (southern Australia in winter and spring).” Using downscaled climate model runs, Christensen and Christensen (2002) show that even as summers become drier over parts of northern Europe, severe 5-day precipitation (99th percentile) might increase. This is echoed by a study by Boé et al. (2009) using downscaled projections from 14 CMIP3 models. According to that study, an increase in high-flow intensity could occur in parallel with an increase in low-flow frequency and a decrease in mean discharge. However, the authors also state that such changes “are strongly linked to changes in extreme precipitation, whose evolution at the regional scale is hard to predict”.

Bukovsky and Karoly (2011) employed the Weather Research and Forecasting (WRF) model as a nested regional climate model to dynamically downscale output from the National Center for Atmospheric Research’s (NCAR) Community Climate System Model (CCSM) version 3 to study frequencies and magnitudes of heavy, very heavy, and extreme precipitation (90th, 99.7th, and 99.9th percentile, respectively) over the Central US. Projections from the WRF for late 21st century warm season precipitation (May–August) show decreases in average precipitation, but an increase in the intensity of both heavy precipitation events and rain in general when it does fall. A decrease in the number of 6-h periods with rainfall accounts for the overall decrease in average precipitation. The WRF also shows an increase in the frequency of very heavy to extreme 6-h average events, but a decrease in the frequency of all events lighter than those. For Canada, Mailhot et al. (2007) found that changes at short durations exceeded those at longer durations. Based

on comparison of control run (1961 to 1990) and future climate simulation (2041 to 2070) from the Canadian Regional Climate Model (CRCM), return periods were found to have halved for the 2-h and 6-h durations, and to have decreased by one third for the 12-h and 24-h durations.

The classical approach to frequency analysis, however, is based on the assumption of stationarity. Khaliq et al. (2006) state: “The classical notions of “probability of exceedence” and “return period” are no longer valid under non-stationarity”. The objective of this paper is to assess whether the assumption of stationarity for rainfall frequency analysis is justifiable. To this end, it is investigated whether non-stationarity is detected for rainfall above certain thresholds and if so, how non-stationarity might affect design rainfall estimates. In order to develop approaches to account for potential effects of non-stationarity in estimating design rainfalls, it is necessary to identify the underlying drivers. As a preliminary step for such an analysis, changes in frequency of rainfall above a threshold are assessed as well as the magnitude of such events, and events are stratified by season. In this paper, we do not attempt to identify the causes for changes in frequency or magnitude of events over time, nor do we suggest solutions to account for potential effects of non-stationarity.

There is little information in the literature discussing potential changes in magnitude and frequency of rainfall extremes at sub-daily and in particular sub-hourly durations for Australia. This is partially due to lack of suitable observations and model data. Design rainfall estimates for short durations are particularly relevant for urban applications. Given the potential for significant economic loss and severe social implications, it appears prudent to explore changes in the frequency and magnitude of these rainfall events. The hydrological response to intense rainfall events strongly depends on antecedent conditions such as soil moisture. Assessing changes in these conditions is outside the scope of this paper.

In a strict statistical sense, stationarity implies that the statistical properties of a time series do not change over time (Coles, 2001). For most practical applications, the definition of stationarity can be relaxed to what is referred to as “wide-sense stationarity”, which is satisfied if neither mean nor autocorrelation change with time. The objective of this paper is to assess whether the assumption of stationarity for rainfall frequency analysis is justifiable. To this end, it is investigated whether non-stationarity is detected for rainfall above certain thresholds and if so, how non-stationarity might affect design rainfall estimates.

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, 1921–2005) of rainfall measurements for durations between 6 min and 72 h. Part 2 explores 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,

giving greater confidence that changes identified as statistically significant are indeed climate-driven.

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

Design rainfall estimates are derived on the basis of historical data, typically measured at rain gauges. Rainfall at sub-daily durations is recorded by pluviographs or *Tip-ping Bucket Raingauges* (TBRG). Pluviographs record rainfall (pen on paper) on paper charts, which are then digitised. It is commonly assumed that rainfall for durations of less than 6 min can not be extracted with sufficient accuracy from a pluviograph. This is why for the analyses presented here, a minimum duration of 6 min was selected. Pluviographs are now increasingly being replaced by TBRG. These instruments measure every 0.2 mm tip and therefore allow better resolution in time. For convenience, instruments that allow sub-daily measurements will in the following collectively be referred to as “pluviographs”.

The highest short duration rainfall depth observed at Sydney Observatory Hill occurred on 8 November 1984 (28.6 mm over 6 min and a daily total of 234.6 mm up to 9 a.m. on 9 November). This month was marked by intense rainfalls and thunderstorms, leading to severe flash flooding in the Sydney Metropolitan area. Damage was estimated at \$40 million (Bureau of Meteorology, 1984). The station is located in an urban setting, at an elevation of 39 m, and at 33.86° S and 151.20° E.

2.1 Records of suitable lengths

The word “climate” is usually used to describe conditions typical for a certain region with regards to a number of climate elements and climate variables. Statistical parameters can be used to describe mean state, range, and variability in addition to extremes. These parameters would usually be derived for a standard period, for instance for 30-yr periods defined by the World Meteorological Organisation (WMO), e.g. 1961–1990. With the increasing use of climate models and scenarios, shorter periods are now also considered. When assessing potential effects of variations in climate, changes can then be assessed against this baseline.

To adequately sample the effects of climate variability, long homogeneous series are required to study the effects of variations in climate on rainfall extremes. Few records of sub-daily rainfall measurements of sufficient length (say 100 yr) and quality exist worldwide. One such record is available for Sydney Observatory Hill, New South Wales, Australia (Bureau of Meteorology station number 066062). This record starts in 1913 but due to significant amounts of missing and accumulated data in some of the early years in

the record, only the period from 1921 to 2005 will be considered.

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 this paper, a peaks-over-threshold approach was chosen since it allows investigation of changes in frequency as well as magnitude. For sufficiently high thresholds, the distribution function of excess values converges to a Generalised Pareto Distribution (GPD). The corresponding time series for the occurrence of extreme events can be presented as a Poisson Process.

Lang et al. (1999) discuss three tests (based on the mean number of events, the mean threshold exceedance, and the dispersion index) to decide on appropriate thresholds u . Here the *dispersion* (ratio of variance of number of events per year to average number of events per year) is used to test whether events do occur randomly in time (rather than in clusters or uniformly distributed). Tests showed that the dispersion depends on the threshold selected and that for suitably high thresholds (less than about 3 events per year), the assumption of a point process is justified. For practical application, it is worth noting that the scale and shape parameter of the GPD can be related directly to the Generalised Extreme Value distribution (GEV) to describe the corresponding annual maximum series (Coles, 2001). If it were known how frequency and magnitude of events change over time, time-varying parameters could be fitted (Coles, 2001). Nogaj et al. (2006) have applied this approach to the analysis of temperature extremes.

The assumption of stationarity implies that firstly, events are independent and that secondly, they are sampled from the same underlying distribution. The first of these two assumptions is addressed by requiring a minimum separation in time when abstracting peaks to ensure independence. The algorithm used for the extraction of peaks identifies the largest event (for a given duration). The period of the event duration and 24 h prior to and after the event is flagged and excluded from further extractions. For events of 6-min duration, this equates to 2 days and 6 min but for a 24-h event, this would equal a 3-day period during which no other extreme can be extracted. Tests with longer and shorter separation times were undertaken but the resulting series of peaks-over-threshold showed little sensitivity to the exact choice of separation time. The choice of a 24-h separation time is also supported by findings of other authors: for 10-min rainfall Ntegeka and Willems (2008) used a 12-h “interevent time”, and for 3-s wind gusts Brabson and Palutikof (2000) use an 18-h separation time.

For intermediate durations in particular, the second assumption (same underlying distribution) is not always valid.

Approaches have been described that seek to address this issue (Willems, 2000; van den Brink and Können, 2008).

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.

An alternative to addressing non-stationarity using the classical approach is to apply a Bayesian approach (Kwon et al., 2008). This approach assumes prior knowledge (for instance about relevant climate indices). More recently, it has been suggested to undertake “pre-processing” of the full data set “*to model the non-stationarity in the body of the process*” (Eastoe and Tawn, 2009) before proceeding with standard extreme value analysis.

Before applying any of these three approaches (classical approach with time-varying parameters, Bayesian, pre-processing), it appears sensible to try and gain an understanding of processes that are relevant for variability in intense rainfall events at different durations and for different climates. The relationship between seasonal and annual Australian rainfall and circulation patterns such as El Niño – Southern Oscillation (ENSO) and the Southern Annular Mode (SAM) has been extensively studied. SAM and Sea Ice Extent (SIE) appear of larger relative importance for the southwest and southern coastal fringes of the continent. ENSO is the dominating influence for the eastern parts of the continent (Pezza et al., 2008). Using wavelet analysis, Westra and Sharma (2006) showed that after removing the effects of ENSO, significant variability in annual rainfall total (with a period of 13 yr) remained for parts of coastal New South Wales. Rainfall in the eastern seaboard (the region to the east of the Great Dividing Range where Sydney is located) does not have the same strong relationship with ENSO as for most of eastern Australia (Timbal, 2010).

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. 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 record was analysed only from 1921 to exclude early years with a high percentage of

missing data. For the remainder of the series, the percentage of missing data per year never exceeded 20%. Quality control (including inspection of metadata) was undertaken. These analyses revealed a suspiciously high value for an event in April 1927. Investigations showed that the paper chart had become soaked during the event, which made it difficult to correctly digitise the trace. Staff in the Bureau of Meteorology’s National Climate Centre have redigitised the chart and the rainfall depth at the shortest duration was found to be to about 1/3 of the originally digitised value. The corrected value was used for the subsequent analyses.

The data were assessed for inhomogeneities using statistical techniques in conjunction with available metadata. First, the program RHtest (Wang and Feng, 2010) was used to test the monthly data, then a special version of the program (RHtests_dlyPrp.r) was used to test for break points in log-transformed daily rainfall data. Initial tests did not identify change points in either of the series although metadata indicate a number of location and instrumentation changes. To test whether these changes might have led to breakpoints, these dates were added prior to further testing. Results indicated four change points, however none of these coincided with known metadata. It was decided against an adjustment to the series for the following three reasons:

- Identified change points were not substantiated by metadata.
- The suggested adjustments were two short jumps (1948 to 1950 and 1979 to 1983), ending at a slightly higher level than the starting level. However, the remaining longer segments would be adjusted using a slight downward trend. An alternative model is no adjustment.
- Undertaking adjustments for the analysis of extremes is problematic (even if using a quantile matching procedure) and should only be undertaken if there are well-founded arguments for undertaking such an adjustment.

3 Results

The results presented in this section are based on pluviograph data available at Sydney Observatory Hill for the period 1921 to 2005. Non-parametric techniques were used, avoiding the need for assumptions about the underlying distribution(s). Changes in frequency and magnitude of rainfall events and empirical quantile estimates are explored by comparing averages for 10-yr sliding windows to long-term averages. The term “event” is used rather loosely. In fact, extreme series were abstracted for a set of standard durations, regardless of the duration of the meteorological event.

3.1 Changes in frequency and magnitude

The average frequency and magnitude for 10-yr sliding windows (shifted by one year at a time) were compared to

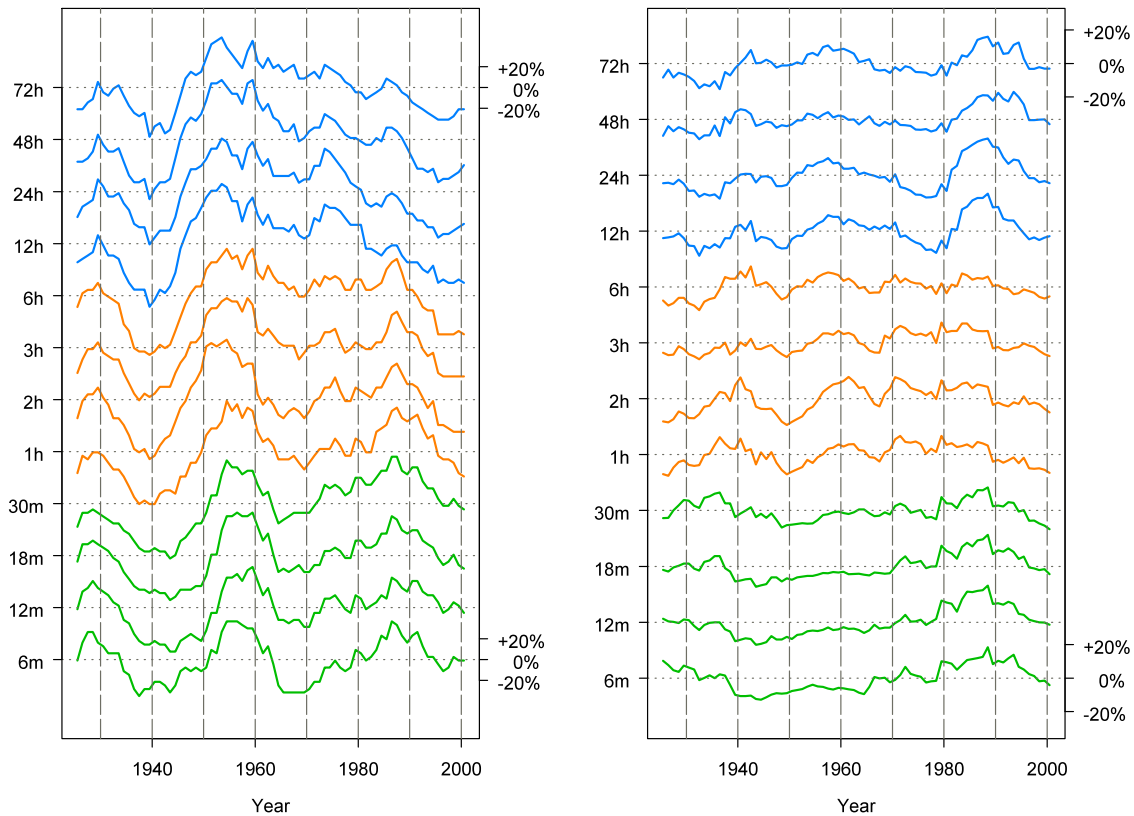


Fig. 1. Percentage difference in frequency (left panel) and magnitude (right panel) for 10-yr sliding windows against the long-term average for POT3. Green, yellow, and blue lines indicate “short”, “intermediate”, and “long” durations, respectively.

the long-term average (for the period 1921 to 2005), see Fig. 1. Changes depend strongly on the *threshold* and are typically more apparent at higher thresholds. This finding is in line with statements that changes in extremes are likely to be more pronounced than changes in annual totals (IPCC, 2007).

Based on changes in frequency and magnitude and for ease of reference, one can define *sets of durations* that exhibit similar changes: short durations (sub-hourly durations), intermediate durations (hourly up to 6 h), and long durations (12 h and above).

Deviations from the long-term average tend to be larger for frequencies than for magnitudes and changes in frequency and magnitude may have opposite signs. For example for long durations for the period 1980 to 2000, a tendency to fewer events is accompanied by an increase in the magnitude of events (Fig. 1).

There are distinct differences in the change in frequency for short and long durations, respectively: with regards to changes in *frequency*, short durations exhibit some oscillation; while for the long durations, the frequency of events appears to have decreased over the three most recent decades. Short durations show an extended period (about three decades starting around 1940) with below

average *magnitudes*, while at long durations magnitudes show a period with well above average magnitudes (late 1980s).

To assess the statistical significance of percentage differences in frequency and magnitude from the long-term mean, confidence intervals were constructed using block-bootstrapping (Prudhomme et al., 2003). To allow for dependence (for instance related to seasonality), data were resampled in 1-yr blocks. For selected durations and for the highest threshold (POT3), percentage differences for frequency and magnitude together with the 90 % confidence intervals are shown in Figs. 2 and 3.

For most durations, the *frequency* of events was significantly reduced for a period around 1940. Periods with significantly higher number of events tend to be shorter and are restricted to intermediate and long durations. Despite the fact that the frequency of events for the long durations has generally decreased since the mid-1970s, only the 24-h duration exhibits significantly below average frequency for the late 1990s. Compared to changes in frequency, significant differences in *magnitudes* are less consistent across duration. Significantly above average magnitudes are found for the late 1980s but only for two durations (12 min and 24 h). For the 12-min duration, magnitudes for the 1940s are judged

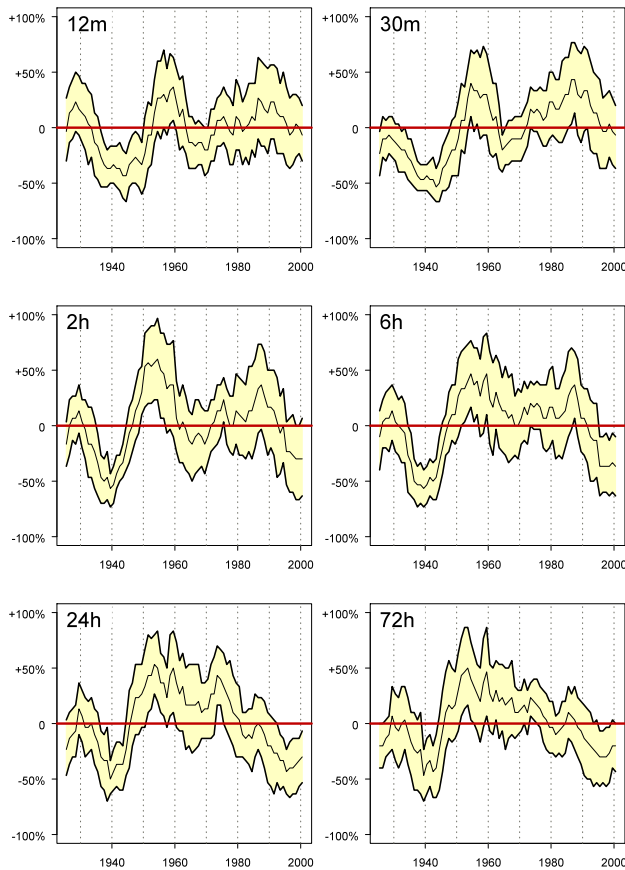


Fig. 2. 90 % confidence intervals for the percentage differences between frequency for 10-yr sliding windows and the long-term average for POT3. For clarity only 6 durations are shown.

significantly lower than the long-term average. This may be related to the fact that Australia experienced a major drought between 1939 and 1945, which was particularly severe in coastal regions of New South Wales (Foley, 1957). These results indicate that it may not be appropriate to infer changes in frequency and magnitude of short duration intense rainfall events from those found for daily durations.

Relative deviations of *seasonal* frequencies from the long-term average are typically larger than variations in magnitude. In some cases, seasonal changes in frequency (or magnitude) cancel each other out. The seasonality of events is covered in more detail in Part 2.

3.2 Changes in average quantile estimates

Both changes in frequency and magnitude of events have the potential to affect quantile estimates and therefore estimates of design rainfall. The following analyses are based on empirically derived quantile estimates for 10-yr sliding windows and the full record.

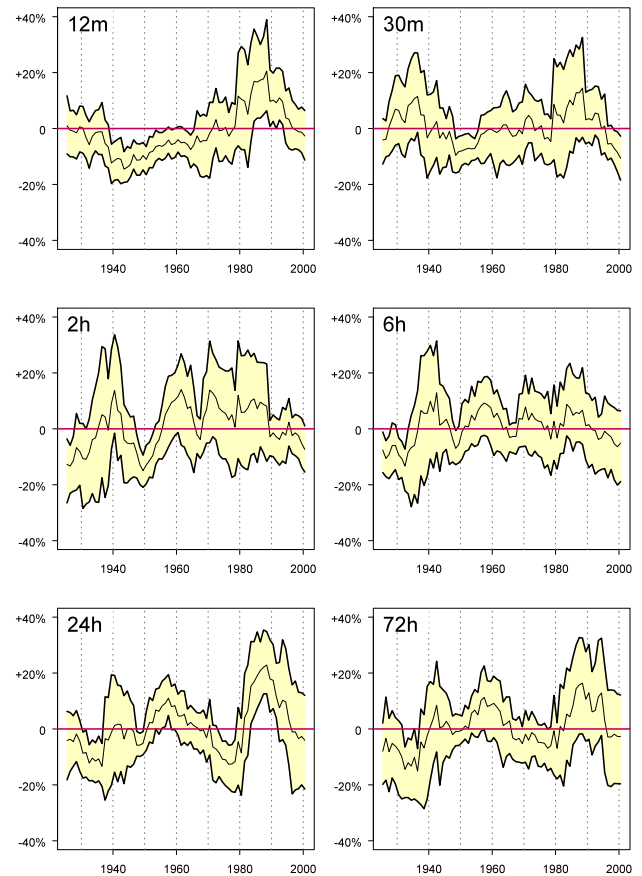


Fig. 3. As Fig. 2 but for magnitudes.

Empirical quantile estimates were derived for the 10-yr sliding windows and the full record respectively. Averages of ratios of quantile estimates were calculated for percentiles between the 50th and the 90th percentile (at increments of 10 %). These average ratios are a suitable measure for characterising changes in quantile estimates for frequent events (occurring once a year or more often) since visual examination of plots typically showed that quantile estimates for this range are either consistently lower (or higher) than for the long-term average. While a non-parametric approach does not require distributional assumptions, it does, on the other hand, not allow inference about changes to rarer events.

Deviations in seasonal quantile estimates from the long-term average are shown in Fig. 4. For seasonal analyses, the annual POT20 series are stratified by season. If events were occurring uniformly throughout the year, this would result in on average 5 events per season. Conventional seasons are used, i.e. (Southern Hemisphere) summer refers to December, January, and February. For clarity, results are presented for only 6 of the 12 durations. Deviations in seasonal quantile estimates can reach up to 60 % (for spring at the 72-h duration). Deviations for quantile estimates based on years (rather than seasons) are much lower and do not exceed

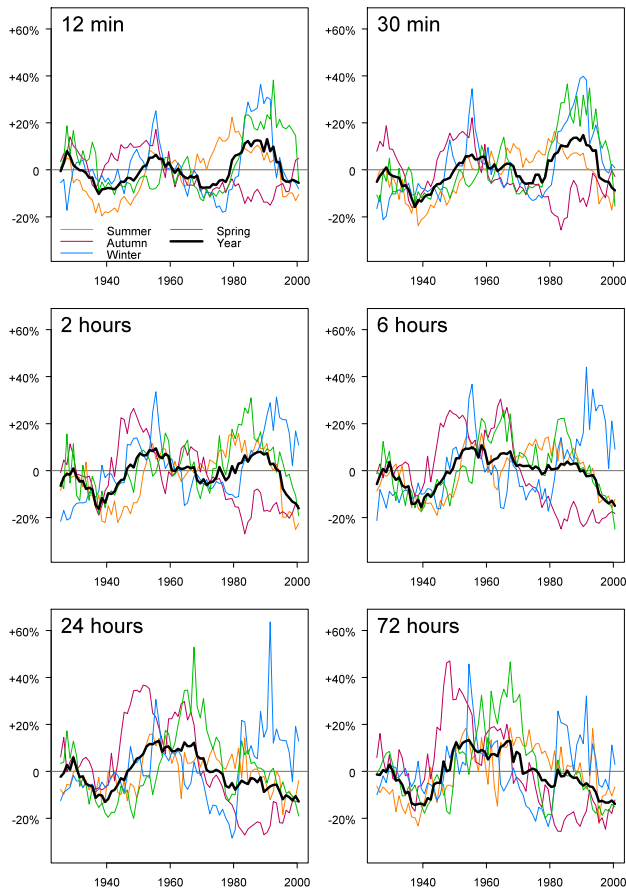


Fig. 4. Average deviation in quantile estimates for 10-yr sliding windows against long-term average for 6 durations based on POT20. Quantile estimates are derived empirically. Deviations are based on an average between the 50th and 90th percentile.

20 %. Seasonal changes occur “out of phase” with each other and therefore variations partially compensate each other. For instance, for a period in the late 1970s, quantile estimates are unusually high for summer but unusually low for winter. Considering the last three decades, one common feature emerges across the range of durations: quantile estimates for autumn have stayed well below the long-term average.

There are distinct differences between the behaviour at short and long durations: for the long durations (12 h and above), average quantile estimates have decreased since about 1960. But for the short durations (up to 30 min), deviations of quantile estimates from the long-term average oscillate rather than exhibit a monotonic increase or decrease. Changes in seasonal quantile estimates are small compared to the longer durations, but over recent decades these changes have been “in phase” (with the exception of autumn).

As for changes in frequency and magnitude, the significance of deviations of quantile estimates from the long-term average was assessed on the basis of 90 % confidence intervals. Only for the highest threshold (POT3) are quantile

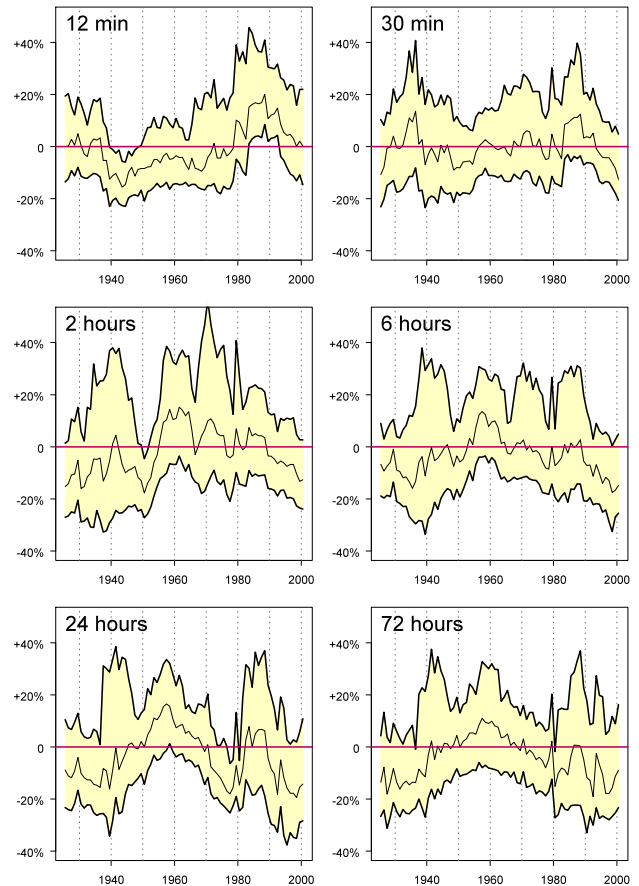


Fig. 5. 90 % confidence intervals for percentage differences between quantile estimates for 10-yr sliding windows and the long-term average based on POT3.

estimates judged to significantly deviate from the long-term average over multi-year periods (Fig. 5), and only the 12-min duration exhibits decade-long periods with significantly decreased quantile estimates (1940s) but significantly increased quantile estimates (from the mid-1980s). This means that for deriving estimates at short durations, period of record might affect estimates. Since most Australian pluviograph records are shorter than 15 yr, this would be a highly relevant finding and this question will be revisited in Part 2 of this paper when we analyse a set of sites. To account for autocorrelation when assessing *significance of trends* in quantile estimates, a modified version of the non-parametric Mann-Kendall test (Hamed and Rao, 1998) was used. The presence of positive autocorrelation leads to underestimation of the variance S . Hamed and Rao (1998) describe an empirical approximation to calculate S in the presence of significant autocorrelation. There is no loss in power when the modified version of the test is used, but assessment of the empirical significance of the test showed it to be less accurate in the presence of large autocorrelation and for small number of observations, here we use 76 10-yr sliding windows. The advantage over the

parametric t-test (based on linear regression) is that one is not restricted to linear trends and that the non-parametric test is less affected by outliers since it is based on ranks rather than differences.

Modified Mann-Kendall tests were applied for two periods: the entire period from 1921–2005 and then again for a shorter period starting in 1976. Considering the entire period, no significant trends were found. For the period from 1976, significant (negative) trends are found for only two durations (6 and 72 h) at a significance level of 10 %.

4 Summary

Variations in frequency and magnitude of rainfall extremes 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.

While there are still large discrepancies between precipitation data sets from observations (including satellites) and models (including reanalysis), both observations and models show that there is a tendency for moist regions to become wetter and for dry regions to become drier (John et al., 2009). However, changes in extreme conditions may show the opposite sign to those in average conditions. Where changes in observed intense precipitation have been studied, this has typically been for daily durations or longer. Investigations presented in this paper, however, show that variations in frequency and magnitude of events at daily durations are a poor indicator of changes at sub-daily durations. Studies like the one presented here, should be undertaken for other regions to allow the identification of regions with significant increase/decrease in intense rainfall; whether there are common features with regards to duration and season exhibiting most significant changes (which in turn could lead to establishing a theoretical framework) and assist in validation of projections of extreme rainfall events.

For the analyses presented in this paper, a peaks-over-threshold approach was chosen since it allows investigating changes in frequency as well as magnitude. Non-parametric approaches were used to assess changes in frequency, magnitude and quantile estimates as well as the statistical significance of changes for one station (Sydney Observatory Hill) for the period 1921 to 2005. Deviations from the long-term average vary with season, duration and threshold. The effects of climate variations are most readily detected for the highest thresholds. Deviations from the long-term average tend to be larger for frequencies than for magnitudes and changes in frequency and magnitude may have opposite signs.

Based on changes in frequency and magnitude, sets of durations were defined: short durations (sub-hourly durations), intermediate durations (hourly up to 6 h) and long durations

(12 hours and above). There are distinct differences between the behaviour at short and long durations: for the long durations (12 h and above), average quantile estimates have decreased since about 1960. But for the short durations (up to 30 min), deviations of quantile estimates from the long-term average oscillate rather than exhibiting a monotonic increase or decrease. Changes in seasonal quantile estimates are small compared to the longer durations but over recent decades these changes have been “in phase” (with the exception of autumn). This means that for deriving estimates at short durations, period of record might affect estimates. This is a highly relevant finding, since most Australian pluviograph records are shorter than 15 yr. It appears plausible that different physical mechanisms force changes for short and long durations, respectively and it may therefore not be appropriate to infer changes in frequency and magnitude of short duration intense rainfall events from those found for daily durations. Using a range of durations and three different thresholds, few significant deviations in frequency and magnitude from the long-term average were found. However, quantile estimates derived for Sydney Observatory Hill for the period 1976 to 2005 show significant decrease for the 6 and 72-h durations.

At least for some regions of Australia, non-stationarity found in historical records has the potential to significantly affect design rainfall estimates. How climate variations affect intense rainfall events (in terms of tendency and magnitude of changes) strongly depends on the duration. Knowledge of how events at daily durations are affected are by themselves a poor guide to inferring changes at sub-hourly durations. Further work should therefore focus on gaining a more thorough understanding of climate driven changes in intense rainfall events. This could be addressed by analysing long records at a small number of sites in Australia. Further analyses might include investigating relevant climate variables (e.g. temperature and precipitable water content associated with rainfall events) as well as climate indices.

Edited by: R. Crockett

Reviewed by: K. Pandzic and another anonymous referee

References

- Brabson, B. B. and Palutikof, J. P.: Tests of the Generalized Pareto Distribution for Predicting Extreme Wind Speeds, *J. Appl. Meteorol.*, 39, 1627–1640, 2000.
- Bukovsky, M. S. and Karoly, D. J.: A regional modeling study of climate change impacts on warm-season precipitation in the central US, 24, 1985–2002, doi:10.1175/2010JCLI3447.1, 2011.
- Bureau of Meteorology: Monthly Weather Review for New South Wales, November 1984.
- Boé, J., Terray, L., Martin, E., and Habets, F.: Projected changes in components of the hydrological cycle in French river basins

- during the 21st century, *Water Resour. Res.*, 45, W08426, doi:10.1029/2008WR007437, 2009.
- Chiew, F. H. S.: Estimation of rainfall elasticity of streamflow in Australia, *Hydrol. Sci. J.*, 51, 613–625, 2006.
- Coles, S.: An introduction to the Statistical Modelling of Extreme Values, Springer Series in Statistics, London, 2001.
- Christensen, J. H. and Christensen, O. B.: Severe summertime flooding in Europe, *Nature*, 421, 805–806, doi:10.1038/421805a, 20 February 2002.
- Eastoe, E. F. and Tawn, J. A.: Modelling non-stationary extremes with application to surface level ozone, *J. R. Stat. Soc.*, 58, Part 1, 24–45, 2009.
- Foley, J. C.: Droughts in Australia: a review of records from earliest years of settlement to 1955, *Aust. Bur. Met. Bull.*, 43, p. 281, 1957.
- Hamed, K. H. and Rao, A. R.: A modified Mann-Kendall trend test for autocorrelated data, *J. Hydrol.*, 204, 182–196, 1998.
- IPCC, Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge, United Kingdom and New York, USA, 996 pp., 2007.
- John, V. O., Allan, R. P., and Soden, B. J.: How robust are observed and simulated precipitation responses to tropical ocean warming?, *Geophys. Res. Lett.*, 36, L14702, doi:10.1029/2009GL038276, 2009.
- Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J.-C., Gachon, P., and Bobée, B.: Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, *J. Hydrol.*, 329, 534–552, 2006.
- Kwon, H., Brown, C., and Lall, U.: Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modelling, *Geophys. Res. Lett.*, 35, L05404, doi:10.1029/2007GL032220, 2008.
- Ladson, A.: Hydrology: An Australian Introduction, Oxford University Press, Melbourne, 2008.
- Lang, M., Ouarda, T., and Bobée, B.: Towards operational guidelines for over-threshold modeling, *J. Hydrol.*, 225, 103–1171, 1999.
- Lenderink, G. and van Meijgaard, E.: Increase in hourly precipitation extremes beyond expectations from temperature changes, *Nature*, 1, 511–514, doi:10.1038/ngeo262, August 2008.
- Mailhot, A., Duchesne, S., Caya, D., and Talbot, G.: Assessment of future change in intensity-duration-frequency (IDF) curves for Southern Quebec using the Canadian Regional Climate Model (CRCM), *J. Hydrol.*, 347, 197–210, 2007.
- Nogaj, M., Yiou, P., Parey, S., Malek, F., and Naveau, P.: Amplitude and frequency of temperature extremes over the North Atlantic region, *Geophys. Res. Lett.*, 33, L10801, doi:10.1029/2005GL024251, 2006.
- Ntegeka, V. and Willems, P.: Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-yr time series of 10 min rainfall intensities at Uccle, Belgium, *Water Resour. Res.*, 44, 1–15, doi:10.1029/2007WR006471, 2008.
- 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.
- Pezza, A. B., Durrant, T., and Simmonds, I.: Southern Hemisphere Synoptic Behavior in Extreme Phases of SAM, ENSO, Sea Ice Extent, and Southern Australia Rainfall, *J. Climate*, 21, 5566–5584, 2008.
- Prudhomme, C., Jakob, D., and Svensson, C.: Uncertainty and climate change impact on the flood regime of small UK catchments, *J. Hydrol.*, 277, 1–23, 2003.
- Timbal, B.: The climate of the Eastern Seaboard of Australia: A challenging entity now and for future projections, *IOP Conf. Ser., Earth Environ. Sci.* 11, 012013, doi:10.1088/1755-1315/11/1/012013, 2010.
- van den Brink, H. W. and Können, G. P.: The statistical distribution of meteorological outliers, *Geophys. Res. Lett.*, 35, L23702, doi:10.1029/2008GL035967, 2008.
- Wang, X. L. and Yang, F.: RHtestV3 User manual, Environment Canada, available at: <http://cccma.seos.uvic.ca/ETCCDMI/software.shtml> (last access: 24 June 2011), 2010.
- Westra, S. and Sharma, A.: Dominant modes of interannual variability in Australian rainfall analyses using wavelets, *J. Geophys. Res.*, 111, D05102, doi:10.1029/2005JD005996, 2006.
- Willems, P.: Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types, *J. Hydrol.*, 233(1–4), 189–205, 2000.