Response to Reviewer

We would like to thank the reviewer of the paper "Regime shifts in Annual Maxima Rainfall across Australia– Implications for Intensity-Frequency-Duration (IFD) relationships". We are particularly pleased with the positive comments provided on the revised version of this manuscript. We have considered the Reviewers' comments and provided descriptions of how each comment will be addressed in the revised manuscript below:

Comment 1:

Line 233: I hope that the authors should clarify the version of 'Australian Rainfall and Runoff (AR&R)' because there was no mention of the AR&R 1999 in Introduction and AR&R 1999 suddenly appeared in line 233, I was really confused with other versions of AR&R. It should be explained the AR&R 1999 and the relation between different versions in Introduction.

Response: The 1987 AR&R edition was republished in book form in 1999. With only the chapter on the estimation of extreme to large floods updated. This has been clarified in the revised manuscript.

Comment 2:

Line 378: The authors considered a simple bootstrap procedure suggested by the referee #1. However, the authors did a repetition number of 100 times. I think that it is too small in the bootstrap resampling method. As the referee #1 said in the comments, I think that more repetition in resampling should be carried out since a small number of repetition causes large sampling error. If more repetition, the range of percent difference in rainfall intensity and the return period with zero crossing may be changed.

Response: While we agree that repeating the bootstrap procedure additional times times may help to further quantify the uncertainty we feel that resampling 100 times is sufficient to represent the effects of sampling and parameter estimation uncertainties under the hypothesis of the existence of two different regimes. Indeed Reviewer 1 who originally suggested the procedure was happy with our method and results.

Comment 3:

In Fig. 7, it would be better to provide the differences between rainfall intensities for two climate phases through resampling method for Melbourne and Brisbane. If the results of these two stations are provided in Fig. 7, the discussion and conclusions would be more reasonable.

Response: The reason that Melbourne and Brisbane are not included in Fig 7 is that the IPO positive and negative rainfall distributions were found not to be significantly different based on the KS test for all durations, whereas the distributions were significantly different for Sydney rainfall. Therefore the Sydney station was chosen to further analyse the impact of regime shifts on the IFD estimates (and associated bootstrap test). We intend to extend the study to look at non-IPO related regime shifts at other stations around Australia but this is beyond the scope of the current paper.

Comment 4:

In future research, it is needed to consider the non-stationary models and the return period definitions (the number of exceedance event and the waiting time in Salas and Obeysekera (2013)) as suggested by referee

Response:

Thank you. We agree with this comment and this will be investigated as part of our ongoing research.

All minor comments have been corrected in the revised paper.

Regime shifts in Annual Maxima Rainfall across Australia– Implications for Intensity-Frequency-Duration (IFD) relationships

Danielle C. Verdon-Kidd[1], Anthony S. Kiem[1]

[1]Environmental and Climate Change Research Group, School of Environmental and Life Sciences, University of Newcastle, AUSTRALIA

Corresponding author address: Danielle C. Verdon-Kidd, Environmental and Climate Change Research Group, School of Environmental and Life Sciences, University of Newcastle, Callaghan, New South Wales, AUSTRALIA.

E-mail: danielle.verdon@newcastle.edu.au

Ph: +61 2 49215749

Abstract

Rainfall Intensity-Frequency-Duration (IFD) relationships are commonly required for the design and planning of water supply and management systems around the world. Currently IFD information is based on the 'stationary climate assumption' - that weather at any point in time will vary randomly and that the underlying climate statistics (including both averages and extremes) will remain constant irrespective of the period of record. However, the validity of this assumption has been questioned over the last 15 years, particularly in Australia, following an improved understanding of the significant impact of climate variability and change occurring on interannual to multidecadal timescales. This paper provides evidence of regime shifts in annual maxima rainfall timeseries using 96 daily rainfall stations and 66 sub-daily rainfall stations across Australia. Further, the effect of these regime shifts on the resulting IFD estimates are explored for three long-term sub-daily rainfall records (Brisbane, Sydney and Melbourne) utilising insights into multidecadal climate variability. It is demonstrated that IFD relationships may under- or over-estimate the design rainfall depending on the length and time period spanned by the rainfall data used to develop the IFD information. It is recommended that regime shifts in annual maxima rainfall be explicitly considered and appropriately treated in the ongoing revisions of Engineers Australia's guide to estimating and utilising IFD information, 'Australian Rainfall and Runoff', and that clear guidance needs to be provided on how to deal with the issue of regime shifts in extreme events (irrespective of whether this is due to natural or anthropogenic climate change). The findings of our study also have important implications for other regions of the world that exhibit considerable hydroclimatic variability and where IFD information is based on relatively short data sets.

1. Introduction

Information on rainfall event intensity, frequency and duration (IFD, or IDF as it is known in some countries) plays a critical role in the design of dams, bridges, stormwater drainage systems and floodplain management. Dependent upon the application, information is required for event-durations ranging from hours to several days. The development of IFD relationships were first proposed by Bernard (1932) and since then different versions of this relationship have been developed and applied worldwide (e.g. Bara et al. 2009, Chen 1983, Hershfield 1961, IHP-VII 2008, Nhat et al. 2006, Raiford et al. 2007).

Historically, in Australia, IFD design rainfall curves were developed by the Australian Bureau of Meteorology (BoM) for durations ranging from 5 minutes to 72 hours and Average Return Intervals (ARI) of 1 year to 100 years (however, recently additional durations and ARIs have also been developed). Up until very recently IFD information available to (and used by) engineers and hydrologists were developed 25 years ago, as part of Engineers Australia publication Australian Rainfall and Runoff (AR&R) in 1987. New IFD information was released early in 2013 after a major revision of IFD information carried out by Engineers Australia. Importantly, the revised IFD information is based on a longer and more extensive rainfall data set (http://www.bom.gov.au/water/designRainfalls/ifd/). However, the BoM and Engineers Australia still recommend to use the AR&R 1987 information for existing flood studies and the probabilistic rational method and to conduct sensitivity testing with the revised 2013 AR&R parameters including the new IFD design rainfalls (http://www.bom.gov.au/water/designRainfalls/ifd/index.shtml).

At the time of writing, the revised IFD information does not take into account the impact of climate change on IFD estimates. This is part of ongoing research commissioned through Engineers Australia. It is also not yet clear how or if the role of natural climate variability is going to be considered. Of concern is the fact that currently, estimates of IFD are based on the assumption that “climatic trend, if it exists in a region, has negligible effect on the design intensities” (Pilgrim 1987). This is known as the ‘stationary climate assumption’ (i.e. the statistical properties of the rainfall do not change over time) and implies that the chance of an extreme event occurring is the same at any point in time (past or future). However, the validity of this assumption has been questioned over the last 15 years following demonstration of the significant impact of climate variability occurring on interannual to multidecadal timescales in Australia. For example, research has shown that annual maximum flood risk estimates in Australia vary depending on climate state (e.g. Ishak et al. 2013, Kiem et al. 2003, Leonard et al. 2008). Importantly these studies demonstrate that founding flood risk estimates on an unsuitable time period has the potential to significantly underestimate (or overestimate) the true risks. This may apply to design rainfall also given that current IFD estimates are based on varying lengths of data spanning different time periods (the latest IFD estimates are based on all daily-read stations with 30 or more years of record and all continuously-recording stations with more than 8 years of record).

Khaliq et al. (2006) explained that the traditional idea of probability of exceedance and return period are no longer valid under non-stationarity. Recently, Jakob et al (2011a) found that rainfall quantile estimates derived for Sydney Observatory Hill for the period 1976 to 2005 show significant decreases across durations from 6 minutes to 72 hours. Jakob et al (2011b) subsequently extended the sub-daily rainfall data analysis to 31 sites located in southeast Australia, assessing variations in frequency and magnitude of intense rainfall events across durations from 6 minutes to 72 hours. This study identified two different trends in the data sets, a decreasing trend in frequency of events at durations of 1-hour and longer for sites in the north of the study region, while sites in the south cluster displayed an increase in frequency of events, particularly for sub-hourly durations. Importantly Jakob (2011a, 2001b) concluded that, for at least some regions of Australia, trends found in historical records has

Danielle Verdon-Kidd 4/11/15 4:40 PM
Deleted: n

the potential to significantly affect design rainfall estimates. Westra and Sisson (2011) also investigated evidence of trends in extreme precipitation at sub-daily and daily timescales (1965-2005) using a spatial extreme value model. They identified a statistically significant increasing trend in precipitation extremes for the sub-daily data set, however at the daily timescale no change in annual maximum rainfall could be detected with the exception of southwest Western Australia (Westra and Sisson 2011). Further, Yilmaz and Perera (2014) conducted change point analysis for extreme rainfall data for storm durations ranging from 6 minutes to 72 hours in Melbourne, and found evidence of regime shifts, concluding the year 1966 as a statistically significant change point. Yilmaz et al (2014) then investigated changes in extreme rainfall through trend analysis, non-stationarity tests and non-stationary GPD models (NSGPD) for Melbourne. They found statistically significant extreme rainfall trends for storm durations of 30 minutes, 3 hours and 48 hours, however for above storm durations there was no evidence of a regime shift (which they termed 'non-stationarity') according to statistical non-stationarity tests and non-stationary GPD (Yilmaz et al (2014).

A limitation of the analysis presented by Westra and Sisson (2011) and Jakob et al (2011a, 2011b) is that they tested for linear trends in the rainfall data series based on the hypothesis that extreme rainfall events would have either decreased, increased or exhibited no trend over the time period being investigated. However these are not the only attributes of trend detection, since annual rainfall maxima may also cycle through interannual to multidecadal periods (note that Westra and Sisson (2011) also investigated possible links between extreme rainfall and annual fluctuations in the El Niño/Southern Oscillation (ENSO)). Therefore, depending on what time period the annual rainfall maxima data are derived from (in reference to any long term cycles or epochs) the observed trends may be misleading or even not apparent (leading to the misconception that regimes shifts are non-significant or not an important consideration). Recently Yilmaz et al (2014) investigated the potential impact of the Interdecadal Pacific Oscillation (IPO) on extreme rainfall and resulting IFD for a case study in Melbourne. They concluded that, the IPO negative phase can be the driver of higher rainfall intensities for long durations and high return periods. However, the trends in extreme rainfall data and differences in rainfall intensities for short storm durations and return periods could not be explained with the IPO influence. Given that Melbourne is located in south-east Australia, where the influence of the IPO is temporally variable due to other climate drivers operating (acting to enhance or suppress impacts, see Kiem and Verdon-Kidd 2010; 2009), the research by Yilmaz et al (2014) provides promise for developing relationships between extreme rainfall and the IPO for regions where the IPO may have a more consistent influence (due to fewer competing climate modes), such as north-eastern Australia.

Therefore this paper aims to establish if there is evidence of regime shifts in the annual maxima rainfall timeseries (1-hour to 7-days) across Australia by testing for shifts (regardless of direction or timing) in the long term sub-daily and daily data. Further, the implications on IFD estimation are explored, along with the potential influence of the IPO on extreme rainfall and resulting IFD. Recommendations are then provided as to how these insights may be incorporated in future revisions of AR&R.

2. Data and methods

2.1 Data

2.1.1 Rainfall data

Sub-daily and daily rainfall data for Australia were obtained from the BoM. Sub-daily data records from continuously recording (i.e. pluviograph) rainfall stations in Australia tend to be relatively short, hindering the ability to conduct trend and attribution studies. In this study pluviograph rainfall stations were chosen with data spanning at least 40 years and at least 90% complete, resulting in 66 stations (see Figure 1a). In order to address the concerns raised in the Introduction about short term data analysis (note that according to Raiford et al. (2007) ARI should not be extrapolated from more than twice the record length), three long-term data sets, highlighted in Figure 1a, were chosen for further analysis that contained data from at least 1913 onwards (Brisbane Aero, Sydney (Observatory Hill) and Melbourne Regional Office).

Daily rainfall stations with data spanning the period 1900 to 2009 were selected in order to capture as much temporal variability as possible (see Figure 1b). These stations were filtered according to the amount of data missing in order to identify the highest quality stations recording rainfall during this period, resulting in 96 being considered suitable for further analysis. Due to variability in the quality and quantity of rainfall data in each State of Australia, the following selection criteria were applied:

- New South Wales, Queensland and Victoria - selected stations are at least 97% complete;
- Tasmania- selected stations are at least 90% complete; and
- South Australia, Northern Territory and Western Australia - selected stations are at least 85% complete.

*****Figure 1 about here****

2.1.2 Climate index data

The climate of Australia has experienced a number of regime shifts in climate during its history, resulting in sustained periods of above average rainfall and storminess and abnormally cool temperatures, followed by the reverse conditions (i.e. droughts and elevated bushfire risk) (e.g. Erskine and Warner 1988, Franks and Kuczera 2002, Kiem et al. 2003, Kiem and Franks 2004, Verdon et al. 2004). These shifts have tended to occur every 20 to 30 years and are associated with changes in the Interdecadal Pacific Oscillation (IPO, Power et al. 1999). The IPO represents variable epochs of warming (i.e. positive phase) and cooling (i.e. negative phase) in both hemispheres of the Pacific Ocean (Folland et al. 2002). Importantly, the IPO has been shown to influence the magnitude and frequency of flood and drought cycles across eastern Australia (Kiem et al. 2003, Kiem and Franks 2004). In New Zealand, the IPO is also associated with similar shifts in flood frequency (McKerchar and Henderson 2003). It has been noted that, following the abrupt shift in the IPO in the mid 1970s, the period, amplitude, spatial structure and temporal evolution of ENSO markedly changed (Wang and An, 2001). Historically, during negative phases of the IPO there tends to be more La Niña (wet) events and fewer El Niño (dry) events (Kiem et al. 2003, Verdon and Franks 2006), resulting in an overall 'wet' epoch for eastern Australia and New Zealand . While during the positive phase of the IPO there tends to be a higher frequency of El Niño events and fewer La Niña events (Kiem et al. 2003, Verdon and Franks 2006), resulting in an overall 'dry' epoch. In this study negative phases of the IPO were defined as 1913-1920 and 1945-1977, while positive phases included 1921-1944 and 1978 to 2010.

2.2 Statistical tests

A 20 year moving window was used to test for low frequency variability in the annual maxima timeseries (1-hour, 1-day and 7-day). A Mann-Whitney U test was then used to determine the statistical significance of possible regime shifts by testing if the first 10 years of data was significantly different from the second 10 years, within the 20 year window (the null-hypothesis in this case was that the data was independently distributed). If the difference in medians was found to be statistically significant (i.e. $p\text{-value} < 0.05$) and there was a change in sign of the median values (e.g. switch from negative to positive), a climate shift was postulated to have occurred during the 10th year of the window. The Mann-Whitney U test is a robust test that does not place implicit assumptions on the underlying distribution of the data (i.e. it is a distribution free test), which is particularly appropriate here due to the small number of years used in each window (Kundzewicz and Robson 2004). Note that a number of different size windows were also tested, however this did not change the results or conclusions.

A second test was also applied to identify step changes in the 1-day and 7-day annual maxima time series known as the distribution free CUSUM with resampling (note that the test was not applied to the shorter sub-daily data as longer data sets are recommended for this method). CUSUM tests whether the means in two parts of a record are different (for an unknown time of change). The second test was applied as it does not require the use of a moving window (which is a limitation of the Mann-Whitney U test described above). However the CUSUM test sequentially splits the timeseries into two portions (which are not necessarily equal), which may be a problem if more than one cycle/shift is present in the timeseries.

The existence of serial correlation (or autocorrelation) in a time series will affect the ability of tests (such as the Mann-Whitney U and CUSUM) to assess the site significance of a trend (e.g. Yu et al. 2003, Serinaldi and Kilsby 2015b). The presence of cross-correlation among sites in a network will also influence the ability of the test to evaluate the field significance of trends over the network (e.g. Yu et al. 2003, Douglas et al. 2000, Guerreiro et al. 2013). Therefore, prior to applying the change point analysis as described above, the Durbin-Watson (DW) statistic was used to test for autocorrelation in the annual maxima timeseries (Durbin and Watson (1950, 1951)). In this case the null hypothesis is that the residuals from an ordinary least-squares regression are not autocorrelated against the alternative that the residuals follow an AR1 process. All DW statistic values were found to be greater than the 1.562 (the upper bound for 1% significance and a sample size of ~100) providing no evidence to reject the null hypothesis. Therefore, any regime shifts detected using the change point methods above are not likely to be artefacts resulting from hidden persistence.

The potential issue of cross- correlation was also investigated. It was found that less than 9% of all possible pairings of rainfall data sets display a significant (yet weak) correlation at the 5% level (r >0.2, significance based on n=100). Only eight pairings (out of 4465) were correlated at 0.5 or higher. It was also found that stations located more than 500km apart were unlikely to be correlated and that the strength of the correlation reduced as distance increased between the pairs. This is not surprising given annual maximum rainfall events are due to synoptic scale processes. Therefore observations relating to spatial consistency of regime shifts are unlikely to be due to spatial correlation between sites.

2.3 IFD Calculation

Danielle Verdon-Kidd 4/11/15 4:41 PM
Deleted: 1

The standard process for obtaining IFD information for a location is to refer to the six master charts of rainfall intensity for various durations and ARIs covering all of Australia in Volume 2 of AR&R 2001. Alternatively, IFD curves can be obtained for any location in Australia via the BoM website (both the AR&R 1987 and revised IFDs are available). This information is based on regionalised estimates of IFDs that are spatially and temporally consistent. However, this approach cannot be adopted when using the instrumental rainfall data required for the analysis presented in this study. As such, the IFD information generated for this project follows the methodology on which the IFDs were based for AR&R 1999 (note the1987 edition was republished in book form in 1999 with only the chapter on the estimation of extreme to large floods updated), which utilises point source data with no regionalisation. It should be noted that it is not the purpose of this paper to compare different methods of generating IFDs, rather, one method has been adopted in order to provide a comparative assessment of the impact of non stationarity on IFD estimation. The AR&R 1999 procedure used to generate IFDs from raw rainfall data (i.e. point based estimates) is summarised as follows:

- A log-Pearson III distribution was fitted to the annual maxima timeseries using the method of moments (for annual maxima series of 30 minutes to 72 hours duration). This is the standard distribution that has historically been adopted for generating IFDs in Australia; however other distributions have recently been tested as part of the revision of AR&R. To test if this distribution is suitable for the region being studied, the goodness of fit for the log-Pearson III was tested using a Kolmogorov Smirnov (KS) test. Here the null hypothesis is that the data fits the Log-Pearson lll distribution (the alternate is that the data does not follow the Log Pearson III distribution). All p-values were greater than 0.05 (average p-value was 0.75), for all series (30min to 72hr durations at Brisbane, Sydney and Melbourne), therefore we accept the null hypothesis at the 5% significance level.;
- The coefficient of skewness was determined for each duration (30 minutes to 72 hours);
- The coefficient of skewness was then used to obtain a frequency factor, K_Y, for use with Log-Pearson III Distribution. K_Y was obtained from Table 2.2 (positive skew coefficients) and Table 2.3 (negative skew coefficients) in AR&R 1999 Book 4;
- Rainfall intensities for a range of ARI were calculated using the following formula:

$$\log RI_Y = M + K_Y S \qquad (1)$$

Where: RI_Y = rainfall intensity having an ARI of 1 in Y

M = mean of the logarithms of the annual maxima rainfalls

S= Standard deviation of the logarithms of the annual maxima rainfalls

K_Y = frequency factor for the required ARI of 1 in Y

- ARIs of 2 years to 10 years were adjusted to partial-duration series estimates. In AR&R 1999, the following correction factors were used (note: for ARI greater than 10 years, no corrected factor is required): 2 year ARI – 1.13, 5 year ARI - 1.04, 10 year ARI – 1.0.

It should be noted that this approach is likely to result in different estimates of IFDs than those obtained from the standard maps provided by AR&R 1999 or the revised IFD estimates released in 2013. Here we are using point based rainfall data, whereas AR&R 1999 have derived regionalised estimates based on multiple rainfall stations with varying lengths of data, varying resolution (daily and pluviograph) and varying quality of records. It is recognised that analysis of rainfall data from single stations is often unreliable, is not temporally or spatially consistent and should generally not be used for design purposes. However, the use of point based rainfall data satisfies the specific aims of this study (which is a comparative analysis) and is therefore considered appropriate.

3. Results

3.1 Test for regime shifts in the annual maxima rainfall timeseries

Significant step changes identified in the extreme rainfall timeseries are shown in Figure 2. Of the 66 sub-daily rainfall stations tested, 40 (61%) displayed at least one step change in the 1-hour annual maxima timeseries (Figure 2a), with some stations exhibiting multiple shifts. Of the 96 daily rainfall stations tested, 86 displayed at least one step change in the 1-day annual maxima timeseries (Figure 2b), while 92 exhibited at least one shift in the 7-day annual maxima timeseries (Figure 2c), and some stations exhibited multiple shifts. Figure 2 collectively shows that observed step changes (or regime shifts) in annual maxima rainfall are not confined to any one particular region of Australia, with most stations analysed exhibiting at least one statistically significant shift.

****Figure 2 about here****

As shown in Figure 3, the CUSUM test yielded fewer stations with statistically significant step change in the annual maxima timeseries (only 18 stations out of 96) and many of these were clustered along the coastal fringe of eastern Australia (note that, although the total number of stations displaying significant change points was the same for both the 1-day and 7-day annual maxima, in some cases the location of the stations differed between the two). However, as stated previously a limitation to this method is that only one significant change can be detected using the CUSUM test (given that the data is sequentially split into two portions during testing). This can be a problem if more than one step change or cycle in the data is present (see example timeseries in Figure 4). Therefore, while the number of stations displaying a step change is reduced using the alternative method, the results do in fact support the theory that regime shift(s) in the annual timeseries are present for some stations at different durations.

****Figure 3 about here****

****Figure 4 about here****

The temporal consistency of step changes in the annual maxima timeseries was further investigated (Figure 5a) and it was found that the timing of observed shifts were not necessarily consistent across Australia. However, for some regions (e.g. the east coast of Australia) periods such as the 1940s (Figure 5b) and to a lesser degree 1970’s (Figure 5c) display a higher degree of spatial consistency.

****Figure 5 about here****

Instability and storminess can result during periods when a number of climate driving mechanisms interact (e.g. El Niño/Southern Oscillation, Indian Ocean Dipole and the

Southern Annular Mode) to influence the occurrence of regional weather systems such as east coast lows and cut off lows (Pook et al. 2006, Verdon-Kidd and Kiem 2009). However, the large-scale climate phenomena impact various regions of Australia at different times of the year and to varying degrees, therefore it is not surprising that the timing of shifts in the annual maxima timeseries varies spatially and temporally. This highlights the limitations of trying to assess and attribute variability in annual maxima rainfall based on a single climate driver (e.g. ENSO) or attempting to address the issue of climate trends for the whole of Australia using one simple approach or model.

3.2 Effect of non-stationarity on IFD estimation

Section 3.1 provided evidence of non-stationarity in the annual maxima timeseries for a range of durations. This non-stationarity may ultimately influence the IFD estimation depending on the length of data, and the time period it comes from, and therefore the underlying climatic state (or combination of states). Current IFD estimates for Australia (both the 1987 and 2013 versions) are based on data as short as 30 years for the daily-read stations and 8 years for the sub-daily data. Therefore IFD estimates based on relatively short-term data sets may under- or over-estimate rainfall intensities, depending on where the data series fits within the long term context (i.e. before or after a shift in annual maxima).

For many east coast stations a shift in 1-day annual maxima (along with the 7-day) occurred around the 1940s - 1950s and again in the 1970s. This timing also corresponds to well-known periods of change in the IPO (see Section 2.1.2 for a description of the IPO and its influences). Therefore, to further explore the issue of regime shifts, breakpoints in the IPO were used to stratify the annual maxima rainfall timeseries into IPO positive and negative epochs for the three long sub-daily data sets described in Section 2.1.1 (i.e. Brisbane, Sydney and Melbourne, see Figure 1a for location). The reason for selection of these stations was twofold. Firstly, for all three stations, a shift in the annual maxima timeseries (for 1-day and 7-day) was observed during the 1940s and again in the 1970s, and secondly the stations contain long records of pluviograph data (the shortest being from 1913 onwards). Figure 6a shows the modulating effect of the IPO on total annual rainfall for the three east coast stations. Annual maxima at the three east coast stations during the two IPO epochs are also shown in Figure 6 (b-d) for event durations of 30 minutes to 72 hours (durations that are critical for flood design applications). A two-sample Kolmogorov-Smirnov (KS) test was applied to determine if the observed differences between the IPO positive and negative rainfall distributions are statistically significant. Here the null hypothesis is that the two samples are drawn from the same distribution.

****Figure 6 about here****

It is evident from Figure 6a that the effect of the IPO on annual rainfall totals (as measured by the largest difference between the two rainfall distributions associated with each climate phase and the results of the KS test) is greatest for Sydney. Although there does appear to be some impact in Brisbane, the result was not statistically significant according to the KS test. Melbourne does not appear to be greatly influenced by the IPO in terms of annual rainfall variability. This is due to the fact that the southern regions of Australia are affected by other climate modes than those arising from the Pacific (i.e. the Southern Annular Mode and the Indian Ocean Dipole (e.g. Kiem and Verdon-Kidd 2010, Gallant et al, 2012)). Regions such as Brisbane and Sydney tend to be dominated by Pacific Ocean influences (e.g. Verdon et al. 2004). Figure 6b shows annual maxima rainfall tends to be higher during IPO negative on average for durations 6 hours and longer at Brisbane (though not statistically significant according to the KS test), while Figure 6c confirms the same to be true for Sydney for durations 2 hours and longer (statistically significant at 95%). However, for Sydney, the outliers (represented by circles) tend to be larger during IPO positive, indicating that the less frequent events might be more intense during this phase.

Irrespective of the fact that the annual rainfall totals for Melbourne do not show any significant difference between the two phases of the IPO, there does appear to be a consistent relationship between IPO and the sub-daily and daily statistics (Figure 6d), whereby the median of the IPO positive distribution is higher across all durations, however IPO negative is associated with less frequent but more extreme events (although results were not statistically significant based on the KS test). For events 24 hours and longer, the IPO negative distribution also shows a much higher degree of variability than smaller durations, with the ‘box and whiskers’ extending beyond the IPO positive counterpart for these longer durations. This suggests that while IPO might not be as dominant in southeastern Australia as it is further to the north it still has some influence that needs to be better understood.

Based on the analysis presented in Figure 6 and the results of the KS test, the Sydney record was chosen to further investigate the effects of regime shifts on IFD estimation. IFD information was generated for the Sydney record using rainfall data from the two IPO phases and the methodology outlined in Section 2.1 for durations 6 minutes through to 72 hours and ARI 2 years to 200 years. In order to test the robustness of the point estimates of rainfall return levels and estimate the uncertainty in their calculation, a simple bootstrap procedure was carried out. Firstly the IPO positive and IPO negative rainfall timeseries were resampled with replacement to obtain two new B-samples. Then for each B-sample the log-Pearson III distribution was fitted and the rainfall intensities calculated for the various return intervals. The difference between the rainfall intensities (of the two B-samples) was then calculated. This procedure was repeated 100 times to build the empirical distribution of the differences (which represents the effects of sampling and parameter estimation uncertainties under the hypothesis of the existence of two different regimes).

Figure 7 shows the difference in rainfall intensity between IPO positive and IPO negative estimates, along with the 95% confidence intervals (CIs) derived using the procedure above.

****Figure 7 about here****

Figure 7 demonstrates clear differences in the resulting rainfall intensities for Sydney estimated for each duration and ARI using the two regimes (i.e. rainfall data from either IPO negative or IPO positive). The difference in rainfall intensity estimated is as great as 65% in some cases. In all cases, the magnitude of the difference in rainfall intensity estimated using the different data regimes is greater for less frequent events (e.g. 50-year, 100-year, 200-year ARIs), highlighting that uncertainty is greatest with less frequent events. The rainfall intensity is greater in IPO positive for the very short duration events (6 minutes) at all return intervals and for 30min duration events for return intervals of 10 years or more. Similarly, for the 24 and 72 hour duration events rainfall intensity in the positive IPO phase is higher for return intervals of 5 years or more. For 2 hour and 6 hour events, the negative phase results in higher intensity events for more frequent return levels (20 years or less) but lower intensities for less frequent events (50 years or more).

4. Discussion and conclusions

An analysis of regime shifts in the annual maxima timeseries (1-hour, 1-day and 7-day) has been carried out using a set of high quality rainfall stations across Australia. It was found that the annual maxima timeseries does indeed exhibit statistically significant step changes/shifts for the majority of stations for various durations. Further it was demonstrated using three long term sub-daily rainfall stations along the east coast that this impacts upon the resulting IFD estimation. The potential for Pacific Influences (i.e. the IPO) to influence the resulting IFD estimation was explored in order to demonstrate this issue. The authors acknowledge that the IPO is unlikely to be the only driver of variability in the annual maxima timeseries across

Australia, and it is recommended that future research should aim to identify other potential drivers of this variability.

These findings highlight the fact that in some instances the IFD estimates currently being used are likely to be either under- or over-estimated at any one time depending on the length of data, and climatic state, from which they were derived. This is a particular concern given that current regionalised IFD information is based on data of varying length (as short as 8 year in the case of sub-daily data) spanning different time periods. An over estimation of rainfall intensity for a given duration could impact on construction costs, while the risks of underestimating rainfall intensities could result in failure of design criteria. That is, the risk is dependent on the application and length of time over which the risk is assessed.

Further revisions of AR&R are currently underway to include an assessment of the potential impacts of climate change on IFD estimates. However, there are many uncertainties associated with climate change projections, particularly when extracting information on timescales shorter than a season and particularly for hydrological extremes (e.g. Blöschl and Montanari 2010, Kiem and Verdon-Kidd 2011, Koutsoyiannis et al. 2008, 2009, Montanari et al. 2010, Randall et al. 2007, Stainforth et al. 2007, Stephens et al. 2012, Verdon-Kidd and Kiem 2010). Therefore, assessing future changes in extreme events that occur over short durations (e.g. minutes to days) is inherently difficult. Furthermore, climate projections are presented in terms of a percent change from a particular baseline. However, the baseline is often inconsistent and ill-defined leading to very different estimates of risk depending on the time over which the baseline is calculated (as has been demonstrated in this paper). Importantly, for regions where large-scale climate drivers operate on a multi-year to multi-decadal timescales and are known to influence extreme rainfall events, we can use this information to determine if the climate statistics on which the IFD are based are likely to be biased or missing crucial information.

It is recommended that regime shifts in annual maxima rainfall be considered and appropriately treated in any further updates of AR&R. One way to do this may be to only utilise data sets of similar length ensuring that they span a sufficient number of years in order to capture data from epochs of both high or low annual maxima (to remove bias towards one climatic phase or another). However, it is acknowledged that this would potentially result in discarding a large amount of data. alternatively, a separate set of IFDs could be developed for use in high risk modelling for engineers who need to account for the 'worst case' (in a similar manner to climate change allowances). This second set of IFD could be developed based on the periods of elevated annual maxima alone (for those stations with clearly defined epochs of annual maxima) such that if we were to enter such an epoch, designs based on these estimates would be robust for the duration of such a period. Salas and Obeysekera (2014) provide similar recommendations to deal with changing exceedence probabilities over time. This would have to be assessed and calculated on a region by region basis given that Australia is a country associated with high spatial and temporal rainfall variability caused by numerous large-scale climate drivers and regional weather phenomena. Finally, any revised estimates of annual maxima should be compared in terms of uncertainty bounds (e.g. following Koutsoyiannis (2006)). Uncertainty analysis, which takes into account both the data availability and variability within the observation period would provide relevant information to practitioners about the reliability of IFD estimates.

This study has highlighted the existence of regime shifts in annual maxima rainfall data in Australia. The driving mechanisms of these regime shifts are likely to vary from location to location and decade to decade. However, these shifts are typical of many natural phenomena and can be described by processes characterized by long range dependence (or regime-switching processes) and captured by hidden Markov models (or similar), resulting in a mixture of distributions that alternate stochastically according to the transition probability from one regime to the next (e.g. Serinaldi and Kilsby, 2015a). While the strategy of defining IFDs for two (or more) different regimes (e.g Serinaldi and Kilsby (2015a)) currently only partially solves the problem, as we often do not know the beginning or the end of a specific regime (be it rainfall or climate driver), recent work has focused on optimizing designs and planning strategies based on the range of what is plausible rather than a reliance on knowing the current and future climate state (e.g. Mortazavi-Naeini et al., 2015). At the same time, work is also underway on seamless prediction at a range of timescales and if/when this eventuates the results discussed here become even more important/useful. Nevertheless, the immediate usefulness of the insights presented here occurs when first establishing the IFD, as an approach similar to that employed here can be used to determine if the underlying data are biased to a mostly wet or mostly dry regime (or a mix of both) which then provides an indication as to whether the IFD is likely to be an over- or underestimate of the true risk. Importantly, this issue needs to be considered and accounted for when attempting to estimate IFD design rainfalls and prior to quantifying how those IFD estimates might change in both the near and long-term future.

While the analysis presented here has been conducted using rainfall data from Australia alone, the recommendations provided are likely to be applicable for other regions of the world where IFD information is based on short term records and particularly for locations with a highly variable climate.

5. Acknowledgements

The authors wish to acknowledge the Australian BoM for supplying the rainfall data used in this study and the UK Meteorological Office for kindly making the IPO data available. We would also like to thank Mr Andrew Magee for assisting with statistical analysis of the rainfall data and the two reviewers of the paper who provided feedback that greatly improved the paper.

6. References

Bara, M., Gaal, L., Kohnova, S., Szolgay, J. and Hlavcova, K. (2009) Estimation of IDF curves of extreme rainfall by simple scaling in Slovakia. Contributions to Geophysics and Geodesy, 39(3), 187–206.

Bernard, M. M. (1932) Formulas for rainfall intensities of long duration. Transactions, ASCE, 96, 592-624.

Blöschl, G. and Montanari, A. (2010) Climate change impacts - throwing the dice?. Hydrological Processes, 24, 374–381,10.1002/hyp.7574.

Chen C. L. (1983) Rainfall intensity-duration-frequency formulas. ASCE J. Hydraulic Eng., 109, 1603–1621.

Douglas EM, Vogel RM, Kroll CN. (2000) Trends in floods and low flows in the United States: impact of spatial correlation. Journal of Hydrology 240: 90–105.

Durbin, J. and Watson, G. S. (1950). Testing for Serial Correlation in Least Squares Regression, I. Biometrika 37 (3–4): 409–428. doi:10.1093/biomet/37.3-4.409. JSTOR 2332391.

Durbin, J.; Watson, G. S. (1951). Testing for Serial Correlation in Least Squares Regression, II. Biometrika 38 (1–2): 159–179. doi:10.1093/biomet/38.1-2.159. JSTOR 2332325

Erskine, W. D. and Warner, R. F. (1988) Geomorphic effects of alternating flood and drought dominated regimes on a coastal NSW river, in: Fluvial Geomorphology of Australia. Academic Press, Sydney, Australia, p 223-244.

Folland, C. K., Parker, D. E., Colman, A. W. and Washington, R. (1999) Large scale modes of ocean surface temperature since the late nineteenth century, in: Beyond El Nino: Decadal and Interdecadal Climate Variability. Springer, Berlin, p 73-102.

Franks, S. W. and Kuczera, G. (2002) Flood frequency analysis: Evidence and implications of secular climate variability, New South Wales, Water Resources Research 38(5), doi:10.1029/2001WR000232.

Gallant, A.J.E., Kiem, A.S., Verdon-Kidd, D.C., Stone, R.C. and Karoly, D.J. (2012) Understanding climate processes in the Murray-Darling Basin: Utility and limitations for natural resources management, Hydrological and Earth System Sciences, 16, 2049-2068.

Guerreiro S.B., Kilsby C.G., Serinaldi F. (2014) Analysis of time variation of rainfall in transnational basins in Iberia: abrupt changes or trends? Int J Climatol 34(1):114–133

Hershfield D.M. (1961) Rainfall Frequency atlas of the United States for durations from 30 minutes to 24 hours and return period s from 1 to 100 years, tech. paper 40, U.S. Department of Comm., Weather Bureau, Washington, D.C.

International Hydrological Programme IHP-VII | Technical Documents in Hydrology | No. 2, (2008) Asian Pacific FRIEND, Rainfall Intensity Duration Frequency (IDF), Analysis for the Asia Pacific Region, Edited by: Trevor M. Daniell and Guillermo Q. Tabios III, Reported by

Regional Steering Committee for Southeast Asia and the Pacific, UNESCO Office, Jakarta 2008.

Ishak, E.H., Rahman, A., Westra, S., Sharma, A., and. Kuczera, G. (2013) Evaluating the non-stationarity of Australian annual maximum flood, Journal of Hydrology, 494, 134-145.

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

Jakob, D., Karoly, D.J. and Seed, A. (2011b) Non-stationarity in daily and sub-daily intense rainfall – Part 2: Regional assessment for sites in south-east Australia, Nat. Hazards Earth Syst. Sci., 11, 2273–2284.

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

Kiem, A.S. and S. W. Franks (2004), Multi-decadal variability of drought risk - Eastern Australia, Hydrological Processes, 18(11), 2039-2050.

Kiem, A.S., Franks, S.W. and Kuczera, G. (2003) Multi-decadal variability of flood risk. Geophysical Research Letters, 30(2), 1035, doi:10.1029/2002GL015992.

Kiem AS, Verdon-Kidd DC (2009), 'Climatic drivers of Victorian streamflow: Is ENSO the dominant influence?', Australian Journal of Water Resources, 13 17-29.

Kiem AS, Verdon-Kidd DC (2010), 'Towards understanding hydroclimatic change in Victoria, Australia - preliminary insights into the 'Big Dry'', Hydrology and Earth System Sciences, 14 433-445.

Kiem, A.S. and Verdon-Kidd, D.C. (2011) Steps towards ‘useful’ hydroclimatic scenarios for water resource management in the Murray-Darling Basin. Water Resources Research 47, W00G06, doi:10.1029/2010WR009803.

Koutsoyiannis, D. (2006) Nonstationarity versus scaling hydrology. Journal of Hydrology, 324, 1-4, 239-254, doi: 10.1016/j.jhydrol.2005.09.022.

Koutsoyiannis, D., Efstratiadis, A., Mamassis, N. and Christofides, A. (2008) On the credibility of climate predictions. Hydrological Sciences Journal, 53 (4), 671–684.

Koutsoyiannis, D., Montanari, A., Lins, H.F. and Cohn, T.A. (2009) Climate, hydrology and freshwater: towards an interactive incorporation of hydrological experience into climate research—DISCUSSION of “The implications of projected climate change for freshwater resources and their management”. Hydrological Sciences Journal 54 (2), 394–405.

Koutsoyiannis D., Montanari A. (2014) Negligent killing of scientific concepts: the stationarity case Hydrol Sci J http://dx.doi.org/10.1080/02626667.2014.959959

Kundzewicz, Z. W. and Robson, A. J. (2004) Change detection in hydrological records – a review of the methodology. Hydrological Sciences Journal, 49(1), 7-19

Leonard, M., Metcalfe A. and Lambert, M. (2008) Frequency analysis of rainfall and streamflow extremes accounting for seasonal and climatic partitions, Journal of Hydrology, 348(1-2), 135-147

McKerchar, A. I. and Henderson, R. D. (2003) Shifts in flood and low-flow regimes in New Zealand due to inter-decadal climate variations. Hydrological Sciences Journal 48(4), 637-654.

Montanari, A., Blöschl, G., Sivapalan, M. and Savenije, H. (2010) Getting on target. Public Service Review: Science and Technology, 7, 167-169.

Nhat L., Y. Tachikawa and K. Takara (2006) Establishment of Intensity-Duration-Frequency Curves for Precipitation in the Monsoon Area of Vietnam, Annuals of Disas. Prev.Res.Inst., Kyoto Univ., No. 49

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

Pook, M. J., McIntosh, P. C. and Meyers, G. A. (2006) The Synoptic Decomposition of Cool-Season Rainfall in the Southeastern Australian Cropping Region. Journal of Applied Meteorology & Climatology, 45, 1156-1170.

Power, S., Casey, T., Folland, C., Colman, A. and Mehta, V. (1999) Inter-decadal modulation of the impact of ENSO on Australia. Climate Dynamics 15(5), 319-324.

Pui, A, Sharma, A, Santoso, A and Westra, S. (2012) Impact of ENSO, Indian Ocean Dipole, and Southern Annular Mode on Daily to Subdaily Rainfall Characteristics in East Australia. Mon. Wea. Rev., 140, 1665–1682.

Raiford J.P., N.M. Aziz, A.A. Khan and D.N. Powell, (2007) Rainfall Depth-Duration-Frequency Relationships for South Carolina, North Carolina, and Georgia. American Journal of Environmental Science 3(2), 78-84.

Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R.J., Sumi, A. and Taylor, K.E. (2007) Climate Models and Their Evaluation. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. and Miller, H.L. (Eds.) Climate Models and Their Evaluation, Cambridge University Press, Cambridge, United Kingdom and New York, USA.

Salas J.D., Obeysekera J. (2014) Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. J Hydrol Eng, 19(3):554–68.

Serinaldi F., Kilsby C.G. (2015a),Stationarity is undead: Uncertainty dominates the distribution of extremes. Advances in Water Resources, 77, 17-36

Serinaldi F., Kilsby C.G. (2015b) The importance of prewhitening in change point analysis under persistence. Stochastic Environmental Research and Risk Assessment

Stainforth, D.A., Allen, M.R., Tredger, E.R. and Smith, L.A. (2007) Confidence, uncertainty and decision-support relevance in climate predictions. Philosophical Transactions of the Royal Society, 365, 2145-2161; doi:10.1098/rsta.2007.2074.

Stephens, E.M., Edwards, T.L. and Demeritt, D. (2012) Communicating probabilistic information from climate model ensembles—lessons from numerical weather prediction. WIREs Climate Change 2012 (3), 409–426, doi:10.1002/wcc.187.

Verdon-Kidd, D.C. and Kiem, A.S. (2009) On the relationship between large-scale climate modes and regional synoptic patterns that drive Victorian rainfall, Hydrological and Earth System Sciences, 13, 467-479.

Verdon-Kidd, D.C. and Kiem, A.S. (2010) Quantifying drought risk in a non-stationary climate. Journal of Hydrometeorology 11(4), 1019-1031.

Verdon, D. C., A. M. Wyatt, A. S. Kiem and S. W. Franks (2004), Multi-decadal variability of rainfall and streamflow - Eastern Australia. Water Resources Research, 40, W10201, doi:10210.11029/12004WR003234.

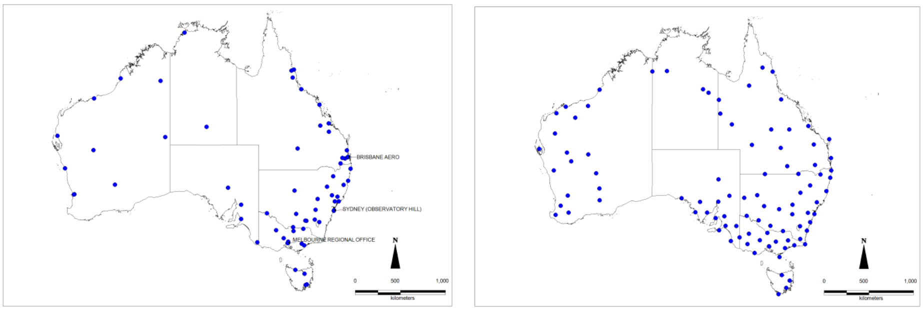
Westra, S. and Sisson, S.A (2011) Detection of non-stationarity in precipitation extremes using a max-stable process model, Journal of Hydrology, 406, 119-128.

Yilmaz, A. G., Hossain, I., and Perera, B. J. C. (2014) Effect of climate change and variability on extreme rainfall intensity–frequency–duration relationships: a case study of Melbourne, Hydrol. Earth Syst. Sci., 18, 4065-4076, doi:10.5194/hess-18-4065-2014.

Yilmaz, A. and Perera, B. (2014) Extreme Rainfall Nonstationarity Investigation and Intensity-Frequency-Duration Relationship, J. Hydrol. Eng., 19, 1160–1172, doi:10.1061/(ASCE)HE.1943- 5584.0000878.

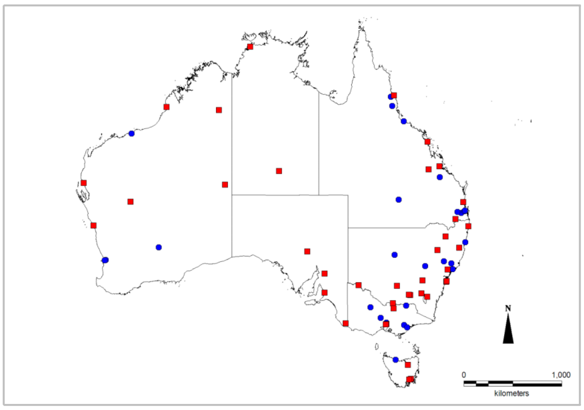
Yue S, Pilon P, Phinney B. (2003) Canadian streamflow trend detection: impacts of serial and cross-correlation. Hydrological Sciences Journal 48(1): 51–64.

FIGURES

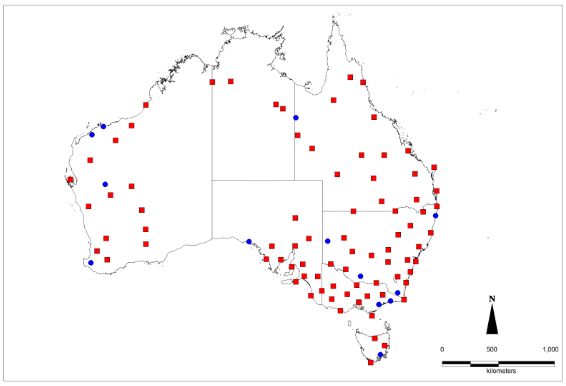


a) b)

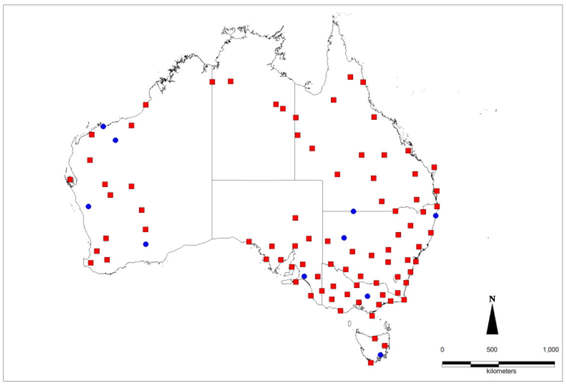
Figure 1 a) Reference stations for sub-daily stations, b) Reference stations for daily rainfall. Note the three long term sub-daily stations used in the IFD analysis are also labelled.



a)

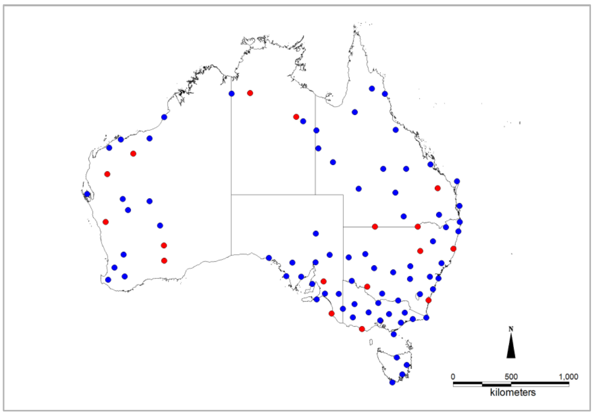


b)

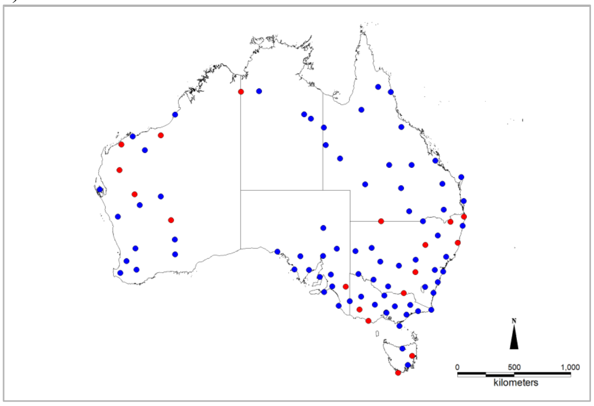


c)

Figure 2 Stations (in red) with at least one statistically significant step change in the a) 1-hour, b) 1-day, c) 7-day annual maximum rainfall (using the Mann-Whitney U test)



a)



b)

Figure 3 Stations (in red) with at least one statistically significant step change in a) the 1-day and b) 7-day annual maximum rainfall (using the CUSUM test with resampling)

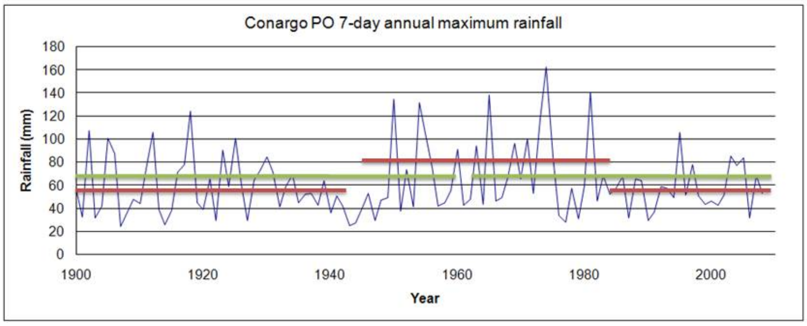
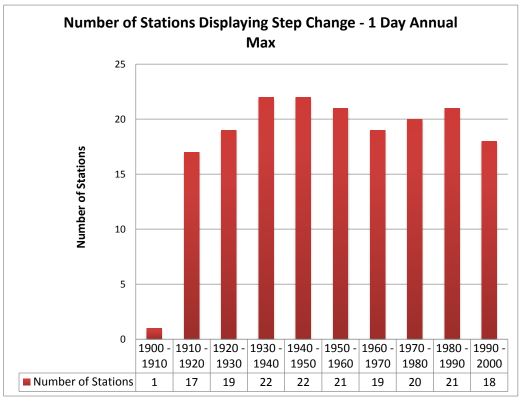
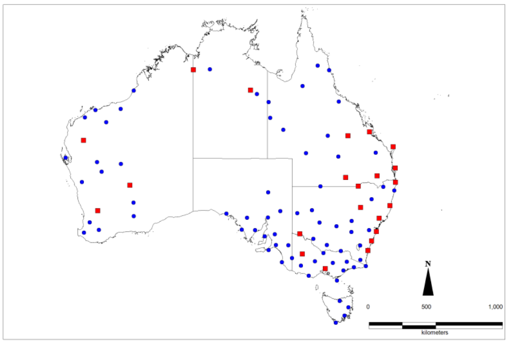


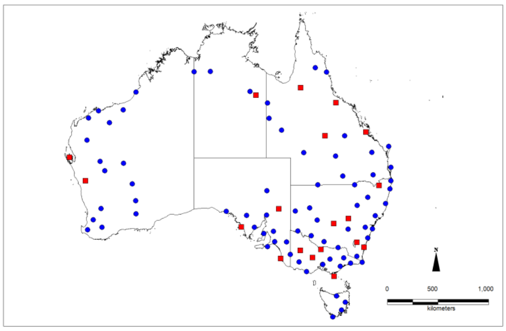
Figure 4 Example of inadequate identification of non-stationarity using CUSUM test (red line highlights three distinct epochs of high/low rainfall, while green line demonstrates effect of splitting the data into two sections for CUSUM test)



a)



b)



c)

Figure 5 a) number of stations each decade displaying evidence of a step change in 1-day annual max, b) Stations (in red) with at least one statistically significant step change in the 1-day annual max during 1940-1950 (using the Mann-Whitney U test), c) Stations (in red) with at least one statistically significant step change in the 1-day annual max during 1970-1980 (using the Mann-Whitney U test)

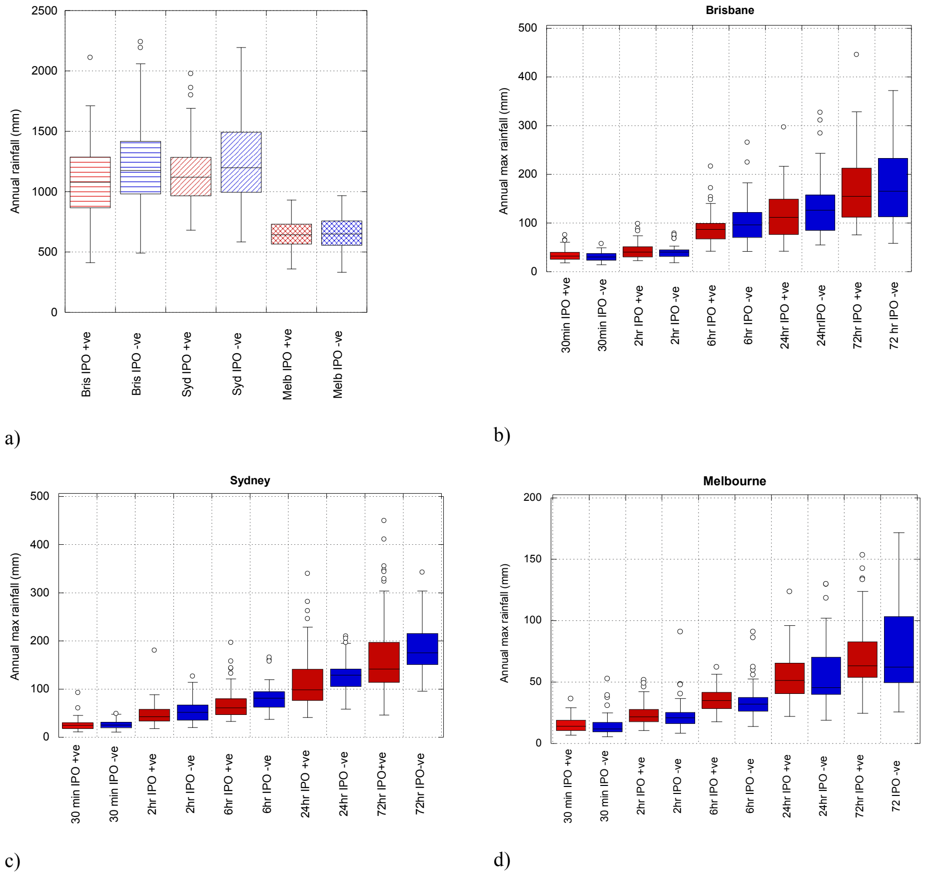


Figure 6 Relationship between IPO and a) total annual rainfall, and annual maximum rainfall at various durations for b) Brisbane, c) Sydney and d) Melbourne.

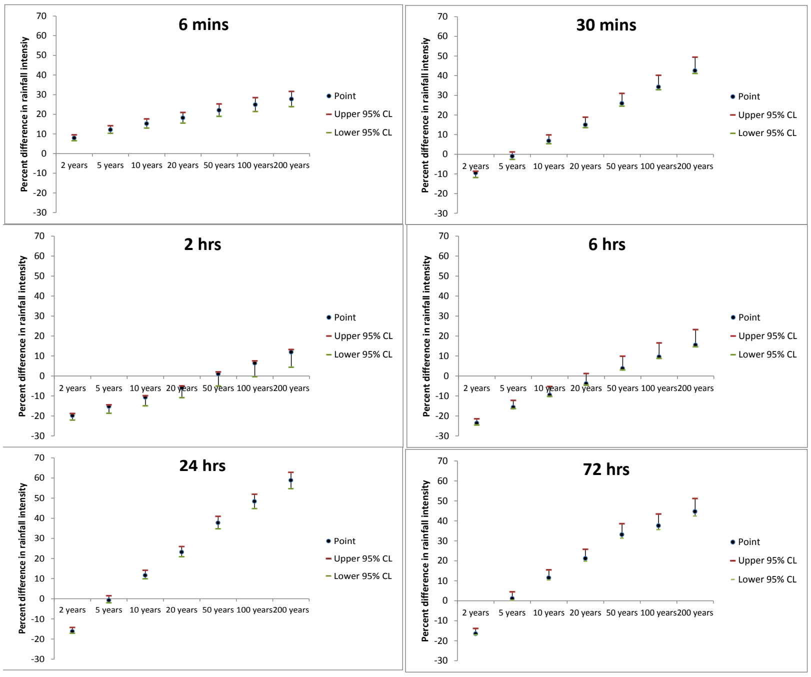


Figure 7 Difference in rainfall intensity for each duration and ARI. Positive (negative) values represent an increase (decrease) in rainfall intensity during IPO positive compared to IPO negative