INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2012)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.3588

The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia

Andrew D. King,[a,b]* Lisa V. Alexander[a,b] and Markus G. Donat[a]
[a] *Climate Change Research Centre, University of New South Wales, Sydney, NSW, Australia*
[b] *ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, NSW, Australia*

ABSTRACT: A 0.05° × 0.05° gridded dataset of daily observed rainfall is compared with high-quality station data at 119 sites across Australia for performance in capturing extreme rainfall characteristics. A range of statistics was calculated and analysed for a selection of extreme indices representing the frequency and intensity of heavy rainfall events, and their contribution to total rainfall. As is often found for interpolated data, we show that the gridded dataset tends to underestimate the intensity of extreme heavy rainfall events and the contribution of these events to total annual rainfall as well as overestimating the frequency and intensity of very low rainfall events. The interpolated dataset captures the interannual variability in extreme indices. The spatial extent of significant trends in the frequency of extreme rainfall events is also reproduced to some degree. An investigation into the performance of this gridded dataset in remote areas reveals issues, such as the appearance of spurious trends, when stations come in and out of use. We recommend masking over areas of low station density for this particular gridded data. It is likely that in areas of low station density, gridded datasets will, in general, not perform as well. Therefore, caution should be exercised when examining trends and variability in these regions. We conclude that this gridded product is suitable for use in studies on trends and variability in rainfall extremes across much of Australia. The methodology employed in this study, to examine extreme rainfall over Australia in a gridded dataset, may be applied to other areas of the world. While our study indicates that, in general, gridded datasets can be used to investigate extreme rainfall trends and variability, the data should first be subjected to tests similar to those employed here. Copyright © 2012 Royal Meteorological Society

KEY WORDS extreme rainfall; gridded data; observations; Australia

Received 11 May 2012; Revised 8 August 2012; Accepted 11 August 2012

1. Introduction

Studies of observed climate mostly use station data to examine trends in extreme rainfall indices both globally (Groisman *et al*., 2005) and regionally (Aguilar *et al*., 2005; Haylock *et al*., 2006; Sen Roy, 2009) including studies focusing on Australia (Haylock and Nicholls, 2000; Alexander *et al*., 2007; Gallant *et al*., 2007). However, this makes comparison with gridded output from climate models difficult (Haylock *et al*., 2008), so modelling evaluation studies frequently rely on gridded datasets based on *in situ* measurements (Meehl *et al*., 2007; Sillmann and Roeckner, 2008; Alexander and Arblaster, 2009), satellite retrievals (Wilcox and Donner, 2007) or reanalysis data (Kharin *et al*., 2007). Satellite data and reanalyses have disadvantages. Rainfall derived from satellite measurements only extends back to the 1970s and significant biases exist in the data (Gerstner and Heinemann, 2008), whilst Hanson *et al*. (2007) found, using the National Centers for Environmental Protection/National Center for Atmospheric Research (NCEP/NCAR) 40 year reanalysis, that there is significant underestimation of precipitation extremes. Gridded data based on *in situ* measurements offer some advantages over these datasets, especially in temporal extent. However, there are issues that need to be considered when using gridded data (Klein Tank *et al*., 2009; Zhang *et al*., 2011), including the representativeness of the gridded values and the sensitivity of the chosen interpolation technique to changing network density. For example, Hofstra *et al*. (2010) used the European E-OBS high-resolution gridded dataset to illustrate the importance of having a large enough station network in order to gain an accurate dataset of daily climate. Hofstra *et al*. (2010) also noted the danger of over-smoothing precipitation data when there are few stations used in the analysis and that this was particularly a problem when examining extreme rainfall data.

Previous research into the uses of gridded data for studying precipitation extremes has been focused on the United States (Chen and Knutson, 2008) and Europe (Hofstra *et al*., 2010) primarily. Australia is a suitable area for further investigation as it encompasses many climate zones (12 in the Koppen climate classification) with very different precipitation regimes whilst being of a similar area to both the United States and Europe. There is also a high-resolution gridded daily rainfall dataset

* Correspondence to: A. D. King, Climate Change Research Centre, Level 4, Mathews Building, University of New South Wales, Sydney, NSW 2052, Australia. E-mail: andrew.king@student.unsw.edu.au

over a century in length available for use in investigating precipitation extremes (Section 2). The gridded dataset in use in this study has a substantially finer resolution than the four used in the E-OBS dataset (Haylock *et al.*, 2008) discussed in Hofstra *et al.* (2010) and the Climate Prediction Center (CPC) Daily US Unified Precipitation dataset used by Chen and Knutson (2008). Chen and Knutson (2008) note that even on high-resolution grids, rainfall extremes are likely to be underestimated if compared directly to point measurements. So, the question still remains whether gridded datasets are sufficient to assess changes in extremes.

This study therefore looks at a broader range of statistics than other studies that have compared the performance of gridded data with station data in capturing precipitation extremes. This comprehensive examination improves the general understanding of the suitability of gridded data to examine rainfall extremes.

The data used are described in Section 2 and the methodology in Section 3. The results and discussion are presented in Section 4 and the conclusions in Section 5.

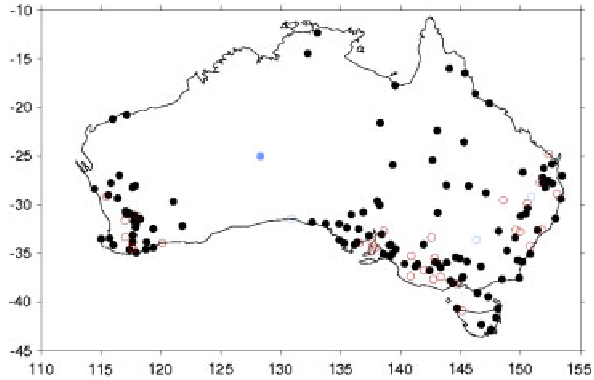

Figure 1. Map of the 152 station locations considered for this study. Filled black circles represent the 119 high-quality stations selected following testing. The filled grey circle shows the location of Giles, chosen to study the performance of AWAP in remote regions. Unfilled circles represent the 3 stations that lacked enough available data to be included and the 30 stations that had unreasonably low and high frequencies of heavy rain events on certain days of the week. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

2. Data

The Australian Water Availability Project (AWAP) gridded daily rainfall dataset was compiled by the Bureau of Meteorology and CSIRO and provides the basis for this study (Jones *et al.*, 2009). The AWAP daily rainfall dataset extends back to 1900 and has a $0.05^\circ \times 0.05^\circ$ resolution. AWAP draws on rainfall measurements from a varying number of stations across Australia peaking at over 7000 in the early 1970s. The non-stationary network used in the development of AWAP is, to some degree, accounted for through the use of rolling climatologies. An anomaly-based approach (Hunter and Meentemeyer, 2005; Xie *et al.*, 2007) was used to generate the spatial analyses with the Barnes successive-correction method (Koch *et al.*, 1983) and three-dimensional smoothing splines applied (Hutchinson, 1995) to three different climatological periods. The methods used in the creation of AWAP do not remove temporal inhomogeneities in rainfall and this is a known caveat. Jones *et al.* (2009) tested the accuracy of the rainfall dataset using several cross-validation methods and acknowledged issues with the accuracy of daily rainfall analyses. Prior to 1907, AWAP has poor spatial coverage over much of Western Australia. Many stations in this region have digitized daily records beginning on 1 January 1907, including ten stations in the high-quality dataset (Lavery *et al.*, 1992), so for the purposes of this study only the period from 1 January 1907 to 31 December 2009 is considered.

To test the ability of AWAP to capture rainfall characteristics, we compared it to the high-quality dataset of daily rainfall measurements discussed in Lavery *et al.* (1992). Originally, data from 191 stations were compiled but following further data quality tests and some station closures this was reduced to 152 sites. These sites were chosen from an original set of around 2100 after passing a series of rigorous statistical tests that searched for inhomogeneities due to non-climatic factors. It is worth noting that the high-quality dataset has not been homogenized and there are likely to be inconsistencies remaining in the data. Of the remaining 152 sites in the dataset (Figure 1), our study required that at least 70% of data at a station from a given year had to be available for extreme statistics for that year to be calculated (see Section 3 for a discussion of the statistics used) and that there had to be extreme statistics for at least 70% of years for that station to be included. Also, to ensure that no bias existed due to seasonality in missing observations, at least 70% of data had to be available for each calendar date through the time series. As a result of these additional tests, a further three stations were removed.

A further issue which needed to be considered was the possibility of accumulations being recorded erroneously against individual days. Viney and Bates (2004) found evidence of untagged accumulations significantly affecting some stations in the high-quality dataset. To examine whether this would affect analysis of extreme rainfall statistics, the days of the week when the highest four rain events in each year fell were compiled for each station and corresponding AWAP gridbox. The frequency of these extreme events would be expected to be equal for each day in the week; however, some sites exhibited high variability in the frequency of extreme events, often with Sundays having very few extreme events and Mondays having significantly more than would be expected. This has been caused by observers not taking measurements on Sundays and allowing accumulations to gather, giving larger rainfall totals on Mondays. This is a particular problem when investigating extreme rainfall totals.

In order to determine quantitatively whether stations should be removed from the analysis, a bootstrapping method was used to estimate the distribution of frequencies that could be expected to occur, within reason, on a given weekday. The expected frequency was generated by a random sample of the days of the week. The sample size is the same as the number of days of extreme intense rainfall being considered (a maximum of 412 for the period 1907–2009, but often less if there is some missing data or the station record is shorter). This random sample was then bootstrapped 1000 times to generate an expected distribution of frequencies. If a station had a day of the week where the observed frequency was more than two standard deviations below the mean of the expected distribution and the following weekday's observed frequency was more than two standard deviations above the mean, then the station was removed from the analysis. Figure 2 shows examples of a histogram of the relative frequency of extreme rainfall days on each day of the week for a station that fails this test [Ardrossan (34.42 °S, 137.92 °E)] and a station that passes this test [Yass (34.74 °S, 148.89 °E)]. This test resulted in another 30 stations being removed from the dataset used for this study leaving a total of 119 stations, 29 of which had too few extreme rainfall events on Sundays and too many on Mondays.

If a station started recording daily rainfall after 1 January 1907 and/or finished prior to 31 December 2009, the corresponding AWAP grid's dates were restricted to the dates the station has been in operation. Also, if a station had missing data for a set of dates, then

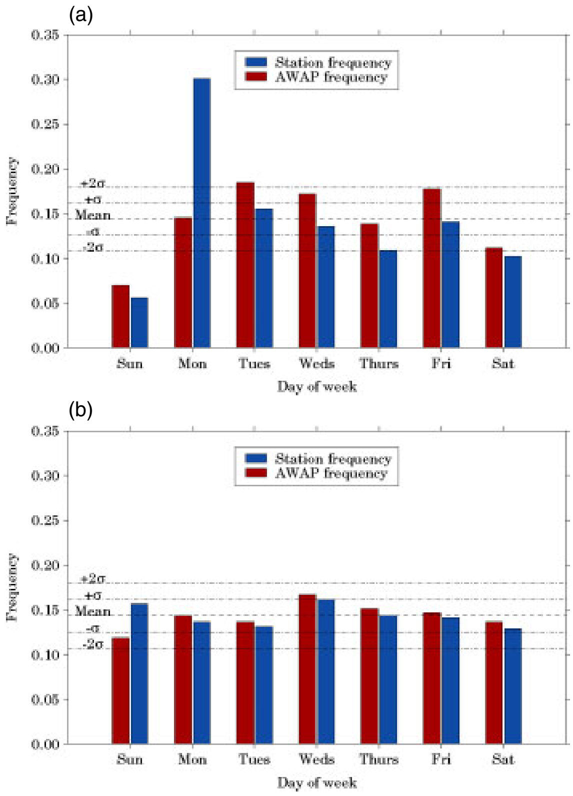

Figure 2. Histograms of the relative frequencies of days of the week when the four heaviest daily rainfall events occurred in each year at (a) Ardrossan (34.42 °S, 137.92 °E) and (b) Yass (34.74 °S, 148.89 °E). The grey bars represent the station relative frequencies and the black bars represent the AWAP relative frequencies at the site of the station. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

missing data were inserted into the AWAP grid for the corresponding dates.

The 119 stations used here are spread across Australia (Figure 1), but are most densely concentrated in the south-eastern part of the continent and south-west Western Australia. Fortunately, the high-quality stations that failed to pass the additional tests imposed in this study are not clustered in a specific region of Australia and the remaining 119 sites still allow for a nationwide comparison of AWAP and station data. A greater number of stations could have been used in this study, but the quality of the results may have been compromised. Quality control is especially important when extreme rainfall is being studied as errors are likely to show up as 'extreme'. The final dataset chosen is not truly independent of AWAP as these stations have contributed to the formation of AWAP and will have a strong influence on the gridboxes being considered. However, these stations provide the greatest knowledge of the climate at each site and are less likely to be subjected to biases and have a longer temporal record than other datasets that could have been used instead. The stations are compared with the gridboxes in which they are located. Testing of a sub-sample of six sites suggested that if neighbouring gridboxes or an average of surrounding gridboxes had been used instead, changes to results would have been negligible. Therefore, the results detailed in Section 4 are robust to the methods applied to the data.

There is a large sparse region of data across much of the centre of the continent where few stations are sited and none of the stations in the high-quality dataset are located [although work to define a high-quality network in Western Australia is being carried out (Marinelli *et al.*, in press)]. Rainfall measurements from Giles (25.03 °S, 128.30 °E), in Western Australia, were used to provide supplementary verification of AWAP data as detailed in Section 4.5 of this study.

3. Methods

The ability of AWAP to capture station rainfall characteristics at all intensities was examined through a comparison of the entire rainfall distributions at all 120 sites in both datasets. The intense rainfall portion of the distribution was then investigated by calculating statistics of extreme indices for the station and AWAP data.

To compare the rainfall distribution between a station and its corresponding grid, plots of every days' station and AWAP grid rainfall were plotted against each other for each of the 120 sites considered (including Giles). The station and AWAP gridbox rainfall were rank-correlated and lines of best fit were plotted. Spearman's rank correlation method was chosen to calculate the correlation between station and gridbox data as the data are unevenly spread with fewer points for heavier events. A rank correlation technique is robust to outliers, so it is more suitable for this application.

A wide range of statistics was required to examine the ability of AWAP to capture characteristics of extreme rainfall seen in station data. Three indices of extreme rainfall were calculated for each year (and season) at each station and AWAP gridbox: extreme frequency, extreme intensity and extreme contribution. Haylock and Nicholls (2000) used three similar extreme indices in their study. The *extreme frequency* is calculated as the frequency of exceedance of the climatological 95th percentile of rainfall. The *extreme intensity* is a measure of the average intensity of the four heaviest daily rainfall events in a year or season. The *extreme contribution* is the proportion of the total rainfall in a season or year that is due to the four heaviest rainfall events. Other indices for extremes could have been used; however, these indices are relatively simple to understand and are appropriate for examining the suitability of AWAP in capturing extreme rainfall characteristics.

The calculation of extreme frequency requires the 95th percentile precipitation. This has been computed for each calendar date in the year using the 103 year time series for each station and AWAP grid (differing from most calculations of similar indices which are based on 30 year climatological periods) and includes days with no rainfall. Excluding days with no rainfall from the calculation of the 95th percentile precipitation would retain seasonality in the number of rain-days and this would have a large effect in some parts of Australia. The rainfall values for each calendar date were placed in ascending order and the nearest value to the point 95% of the way along this string of data (accounting for missing data) was assigned as the 95th percentile rainfall value for that date. The 95th percentile precipitation is then smoothed across calendar dates through the application of a 21 day binomial smoothing filter. Note that using a date-dependent 95th percentile rainfall spreads the dates on which rainfall events used in the calculation of extreme frequency occur through the year (Figure 3). This means that some light rainfall events during a normally dry spell of the year are incorporated in the calculation and heavy rain events in a wetter period of the year are not included [e.g. for Katherine (Figure 3(b)) light rainfall events during winter (JJA) are included whilst comparatively wet ones during summer (DJF) are not].

When examining extreme frequency, it was decided that it is important to incorporate the seasonal cycle of precipitation so that the index provided a good representation of extreme events across the entire year and did not only capture intense rainfall events in one season. A date-dependent percentile method could also have been applied in calculating the extreme intensity and contribution indices instead of using an absolute number of events for each season/year. The authors decided, however, that the extreme intensity and contribution indices should represent only the heaviest rainfall events (and the ones that are most likely to have the largest impact on society) even if they were more likely to occur during a specific period of the year or season being considered. Consecutive days of intense precipitation from the same weather system are counted separately in the calculations of our indices.

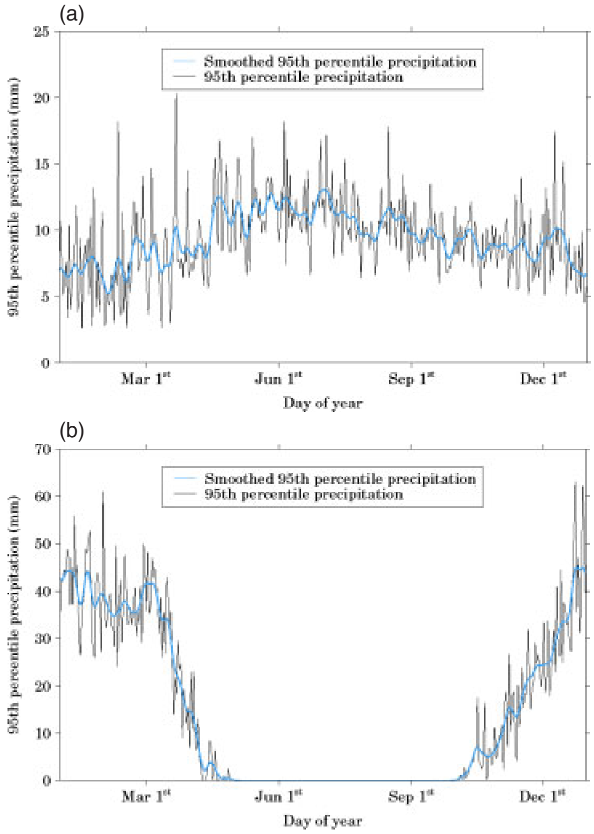

Figure 3. Line graphs (black lines) of the date-dependent 95th percentile precipitation at (a) Deal Island (39.48 °S, 147.32 °E) and (b) Katherine Council (14.46 °S, 132.26 °E) with a 21 day binomial smoothing applied (grey lines). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

For each of the three indices, a suite of statistics was calculated annually and for each season across the 1907–2009 time series. For each index, X, and each year/season, k, up to n years/seasons, these statistics are the:

- Mean average of the index for the station and its corresponding AWAP grid.

$$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k \quad (1)$$

- Bias between the time averages of the station and grid calculated as a percentage.

$$\text{Bias} = 100\left(\frac{\overline{X_{\text{AWAP}}} - \overline{X_{\text{Station}}}}{\overline{X_{\text{Station}}}}\right) \quad (2)$$

- Root mean square deviation between the station and grid.

$$\text{RMSD} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(X_{\text{AWAP}} - X_{\text{Station}})_k^2} \quad (3)$$

- Mean absolute deviation between the station and grid.

$$\mathrm{MAD} = \frac{1}{n}\sum_{k=1}^{n} |X_{\mathrm{AWAP}} - X_{\mathrm{Station}}|_k \qquad (4)$$

- Standard deviations of the index distribution for both the station and grid.

$$\text{Standard deviation} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2} \qquad (5)$$

- Correlation (Pearson's) of the station and gridbox time series. The time series of the three extreme indices are assumed to be near-Gaussian as each point in the time series is calculated based on several events.
- Significant trends (at the 5% significance level using the Mann–Kendall statistic) in station and grid time series.

The mean rainfall (not shown) and mean 95th percentile rainfall were also calculated for each station and its AWAP grid annually and for each season.

4. Results and discussion

Annual and seasonal time series of each of the three indices were produced for each of the 119 stations and their corresponding grids for 1907–2009. The statistics listed in Section 3 were calculated and plotted onto maps. We first discuss the entire rainfall distributions in station and AWAP data (Section 4.1), before going on to discuss annual (Sections 4.2 and 4.3) and seasonal results (Section 4.4). Finally, we discuss results for a remote location (Section 4.5).

4.1. The distribution of rainfall at stations and AWAP gridboxes

For each station and corresponding AWAP grid, all the daily rainfall measurements were plotted against each other to compare characteristics of the rainfall distributions between the point and gridded data. Lines of best fit (all linear) were plotted and the rank correlations calculated. Examples of these plots for two sites are shown in Figure 4. There is very little difference in some of the characteristics of these plots between the 119 different sites. The gradient of every line of best fit is between 0.7 and 1 and the intercept is >0 in all cases. AWAP, consistently, has a tendency towards larger rainfall values for very low rainfall events (and sometimes non-zero rainfall values when the station measurement is zero), but also lower values than station measurements for extreme rainfall events. AWAP has fewer non-zero rainfall days than are observed in the station data at all sites examined. Possible reasons for this difference include AWAP accounting for isolated non-zero rainfall values at station sites and spreading them over an area or spurious zero values in station data where rainfall is sometimes under-recorded for small accumulations (Trewin, 2001). In coastal regions, zero values of rainfall are known to be extrapolated to non-zero values in nearby areas where topography exists. Further investigation into the performance of AWAP in regions of differing topography is required. Also, the use of rolling climatologies creates surfaces that are non-stationary in time. Therefore, the construction of AWAP may also provide some explanation for there being fewer zero rain-days at AWAP gridboxes than observed at stations.

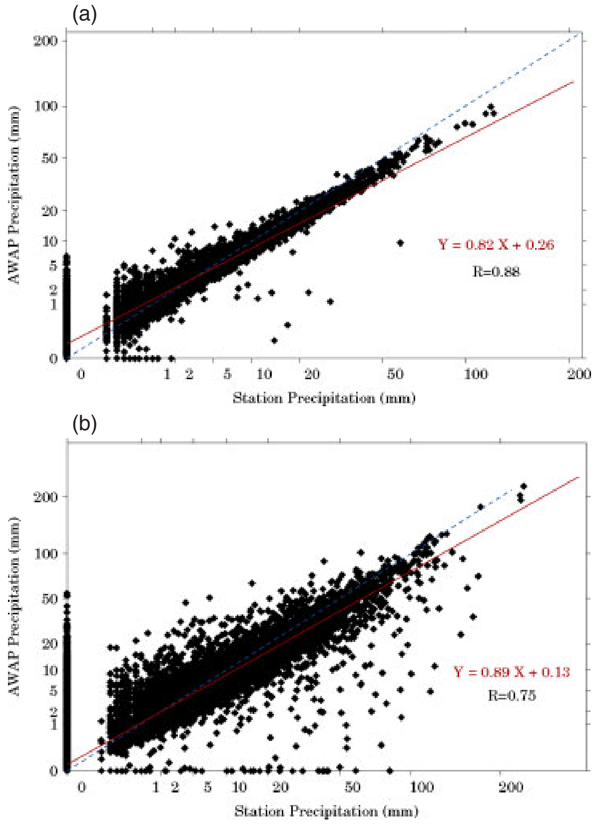

Figure 4. Plots of daily station precipitation against daily AWAP precipitation at (a) Deal Island (39.48 °S, 147.32 °E) and (b) Katherine Council (14.46 °S, 132.26 °E). The line of best fit is shown in solid and the dashed line represents a 1 : 1 relationship. The equation of the line of best fit is shown for both sites as well as rank correlation (R) values. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

The rank correlation shows the strength of the relationship between the station and AWAP rainfall and is plotted for each location (Figure 5). Rank correlation values are generally larger in coastal areas (most noticeably in the south-west and south-east of Australia) and lower in inland areas. The reason for this may, in part, be related to stations in coastal locations having greater influences on their AWAP gridboxes as there is no additional information affecting the gridbox value from the area off the coast. It may also be related to the local station density or, indeed, the nature of rainfall in different parts of Australia. In coastal regions of the south-west and south-east, a larger proportion of total rainfall is due to synoptic-scale systems than mesoscale systems, compared with

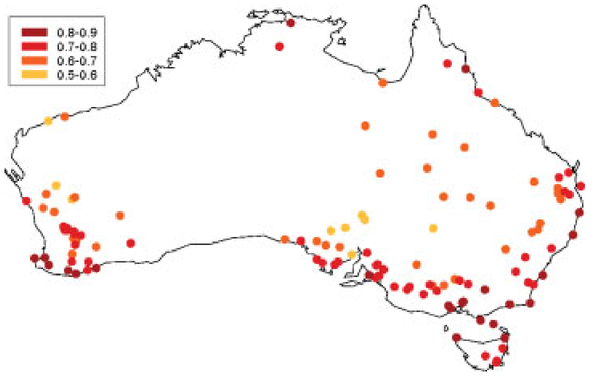

Figure 5. Map of the rank correlation coefficients between daily station and corresponding AWAP rainfall values.

other parts of Australia. Therefore, in these coastal areas of the south, neighbouring stations are more likely to record similar rainfall values when larger-scale frontal systems pass through. This will increase the consistency of AWAP values through the region and each AWAP value is then more likely to be close to its corresponding station value, thus, increasing consistency between the station and AWAP values and increasing the rank correlation statistic.

4.2. Annual indices

The annual mean 95th percentile precipitation values (Figure 6(a)) are generally higher at stations in coastal regions, particularly in the east and north, while those in central and western regions are considerably lower. This pattern is also seen at the AWAP grids (Figure 6(b)). AWAP estimates of the 95th percentile precipitation are lower than at the stations at the majority of locations (86 of 119 sites). In coastal areas, a greater proportion of sites have higher 95th percentile rainfall values at stations than gridboxes, as opposed to in areas, further inland, where AWAP gridboxes often have greater 95th percentile rainfall than station sites. Sites with lower 95th percentile rainfall (i.e. those in inland and western areas) are more likely to have greater 95th percentile rainfall estimates for AWAP grids than at the station sites. At many of these sites, there are times in the year when the station 95th percentile rainfall is 0 mm and AWAP often has small non-zero rainfall totals instead. This is likely to be the cause of the positive difference in these regions. Overall, it is worth noting that the average difference between station and AWAP 95th percentile rainfall estimates is $<20\%$ at all 119 locations and is $<10\%$ at around three-fourth of sites. Similarly to E-OBS (Hofstra *et al.*, 2010), AWAP tends to over-smooth precipitation extremes.

Annual time series of extreme frequency, extreme intensity and extreme contribution were analysed for each of the 119 sites. Figure 7 shows examples for Dowerin-Ejanding (31.01 °S, 117.13 °E) and Curlewis-Pine Cliff (31.18 °S, 150.03 °E). The AWAP time series of the extreme indices are generally well correlated ($r > 0.7$) with the corresponding station time series and

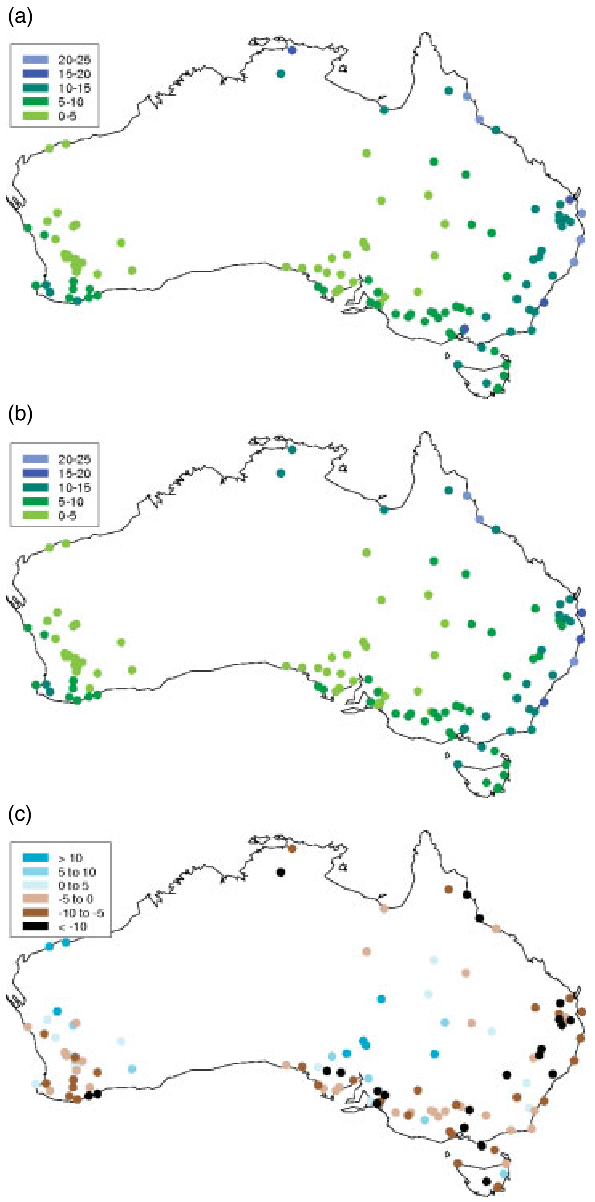

Figure 6. Maps of the mean (smoothed) 95th percentile daily rainfall in (a) station data (mm), (b) AWAP gridbox data (mm) and (c) the bias (%).

the standard deviation is similar (i.e. any interannual variability in the extreme indices at a given station is usually detected in AWAP as well). There is little geographical variability in the correlation coefficients (not shown); however, extreme contribution correlation coefficients are on average slightly lower (0.85) than for extreme frequency (0.88) and extreme intensity (0.9). The high correlation coefficients between the time series suggest that it is likely that the same rainfall events are contributing to the calculations of the extreme indices for the station sites and AWAP grids.

Biases in the mean of each of the annual extreme indices between AWAP and the stations were calculated. Maps of the mean biases in each of the extreme indices were plotted in order to examine geographical patterns and consistency across regions.

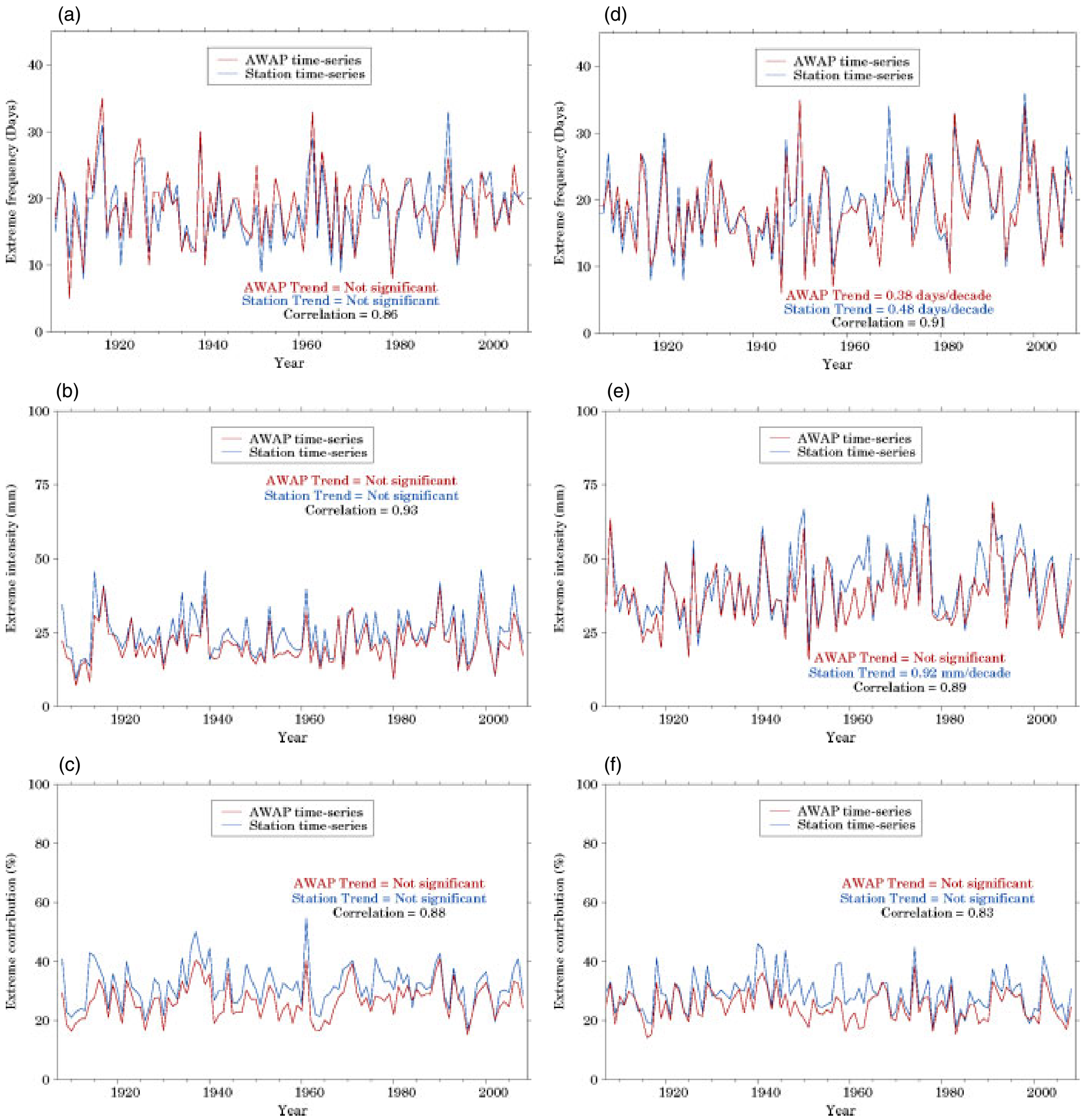

Figure 7. Time series of (a and d) extreme frequency above the 95th percentile (days), (b and e) extreme intensity (mean of highest four daily rainfall totals each year (mm)) and (c and f) extreme contribution [contribution of highest four daily rainfall totals to annual rainfall each year (%)] at Dowerin-Ejanding (31.01 °S, 117.13 °E) and Curlewis-Pine Cliff (31.18 °S, 150.03 °E). Time series of each index are shown for the station and corresponding AWAP gridbox. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

The extreme frequency is calculated using the 95th percentile precipitation (where days with no rain are included). Owing to the definition, a mean extreme frequency of roughly 18 would be expected for each year, and this is close to what is generally found at stations and gridboxes. Biases in the mean extreme frequency (not shown) are typically low (between −3 and 3%) at the majority of sites, and are much greater at sites in the north and north-west of Australia where the mean extreme frequency at the AWAP grid was often more than 25% greater than at the station. At the sites with large biases, the station mean extreme frequencies are well below those that would be expected. This is, in part, due to there being fewer than 5% of days with recorded rainfall in some seasons (discussed further in Section 4.4). There are more rain-days in total in AWAP (as discussed in Section 4.1 and shown for two sites in Figure 4), so biases are generated.

Biases in mean extreme intensity and contribution are consistently negative across all but 3 (for extreme intensity only) of the 119 sites with AWAP grids having lower extreme intensity and contribution values than stations. Extreme intensity (the average intensity of the heaviest four daily rainfall events in each year) at the stations is larger in coastal areas in the east of Australia (Figure 8(a)). This pattern can also be seen when looking at the AWAP grids, albeit with lower values (Figure 8(b)). The bias (Figure 8(c)) appears to be greater in inland and lower on the coast, particularly on the coast of New South Wales and south-east Queensland. This is likely due to lower absolute values of extreme intensity at inland stations (Figure 8(b)), but may also be related to the

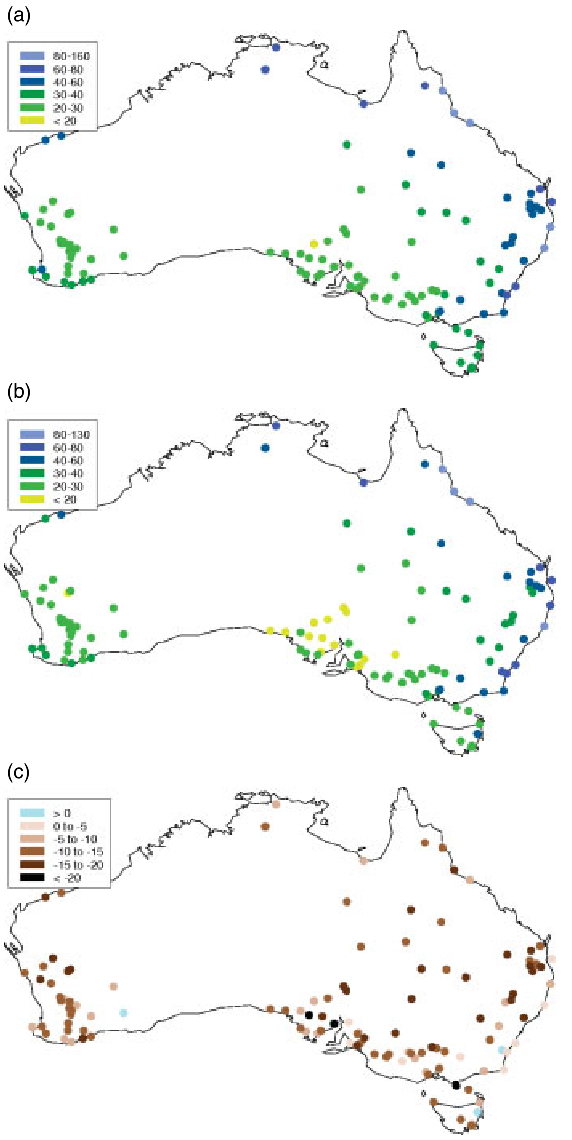

Figure 8. Maps of the mean annual extreme intensity in (a) station data (%), (b) AWAP gridbox data (%) and (c) the bias (%).

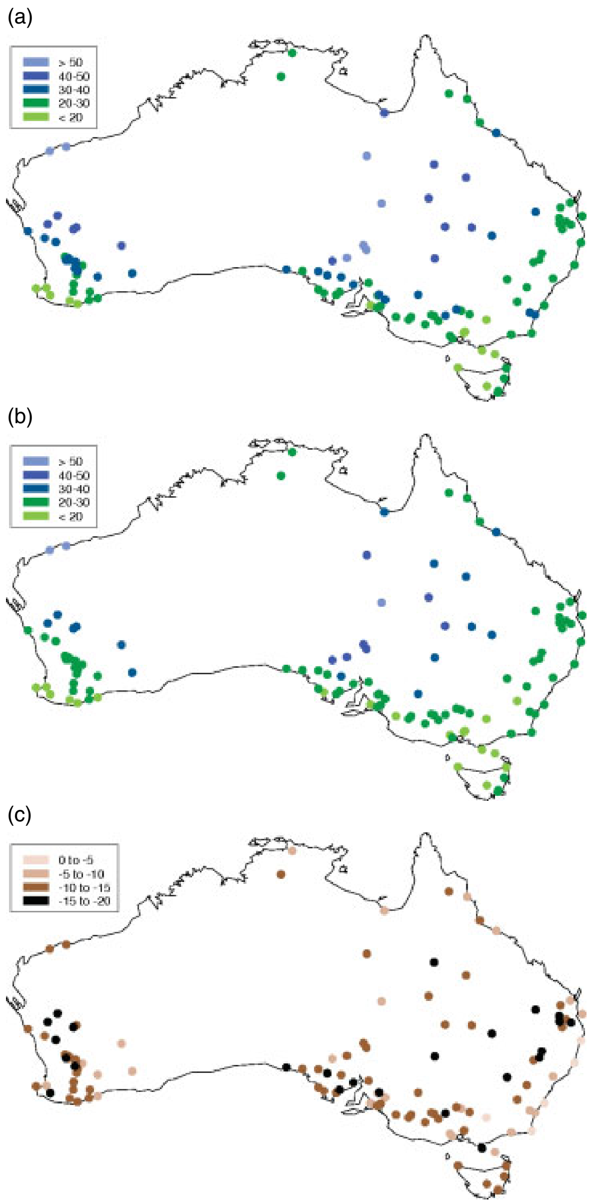

Figure 9. Maps of the mean annual extreme contribution (%) in (a) station data, (b) AWAP gridbox data and (c) the bias.

interpolation technique of AWAP in coastal areas. The average bias (AWAP-Station) is −11.4%.

There is a very different spatial pattern in the extreme contribution index (the contribution to total annual rainfall from the heaviest four events in each year) across Australia (Figure 9). The largest extreme contribution values are at stations in central and western areas with considerably lower values at coastal sites, particularly in south-west Western Australia and Tasmania (Figure 9(a)). This geographical distribution is also seen at the corresponding AWAP grids (Figure 9(b)). The sites with the largest extreme contributions are also those with the fewest rain events, so it is unsurprising that the heaviest four rain events make up a greater proportion of the total rainfall. Negative biases exist at all sites, although there is no obvious spatial pattern in the magnitude of these biases (Figure 9(c)). The average bias (AWAP-Station) is −11.5%.

4.3. Trends of annual indices

The trends in each of these extreme indices (over the period 1907–2009) were calculated using the Mann–Kendall trend statistic and tested at the 5% significance level. Maps of the locations of significant trends are shown for each index (Figure 10) where trends are significant at the station, the AWAP grid or both. At no locations, for any of the extreme indices, are there significant trends of opposing signs between the stations and corresponding AWAP grids. There are 26 locations with significant trends in the extreme frequency index at stations, AWAP grids or both (Figure 10(a)). In south-west Western Australia, there are four stations and six AWAP gridboxes showing significant decreasing trends in extreme frequency. Haylock and Nicholls (2000) also found a significant decreasing trend in extreme frequency in this region when looking at a grouping of high-quality

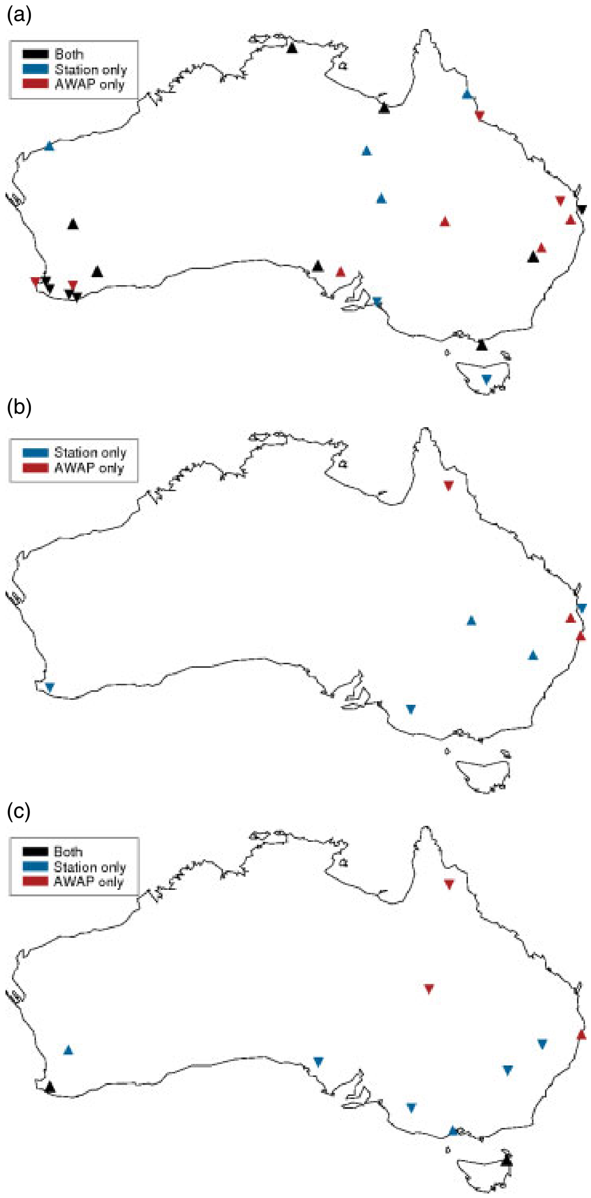

Figure 10. Maps of sites showing significant trends in annual (a) extreme frequency, (b) extreme intensity and (c) extreme contribution. Blue triangles represent sites of significant trends in station data only, red triangles represent sites of significant trends in AWAP data only and black triangles represent sites of significant trends in both. The orientation of the triangle represents the sign of the trend.

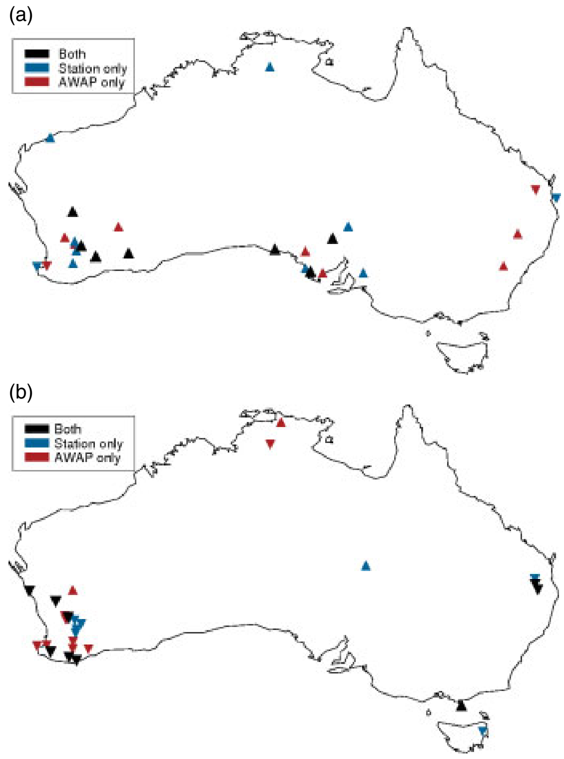

Figure 11. Maps of sites showing significant trends in seasonal extreme frequency for (a) summer (DJF) and (b) winter (JJA). Blue triangles represent sites of significant trends in station data only, red triangles represent sites of significant trends in AWAP data only and black triangles represent sites of significant trends in both. The orientation of the triangle represents the sign of the trend.

stations in the area together over a similar period of time. Many studies have noted a decrease in total rainfall and extreme rainfall in the south-west of Western Australia (Hennessy *et al.*, 1999; Li *et al.*, 2005; Gallant *et al.*, 2007). There are two locations in this area where the AWAP grid shows a significant decreasing trend and the station does not. Outside of this region, most of the significant trends are positive suggesting that the majority of the country has experienced an increase in the number of extreme rainfall events. There are far fewer locations that have experienced significant trends in extreme intensity (Figure 10(b)) and extreme contribution (Figure 10(c)) and there are no obvious spatial patterns.

The other statistics listed in Section 3 were calculated for each of the extreme indices. The root mean square difference, mean absolute difference and standard deviations of each index are generally higher at locations where the value of the index is higher (not shown). This means that larger extreme index values are associated with greater interannual variability and larger differences between the station and AWAP time series.

4.4. Seasonal statistics

The same statistics that were calculated annually were also calculated for each index seasonally. The mean 95th percentile precipitation varies greatly between seasons in some locations. This seasonal variability is most apparent at sites in the north of Australia where during the wet summer (DJF) and autumn (MAM) seasons some stations and grids have mean 95th percentile rainfall in excess of 30 mm, whereas during the drier winter (JJA) and spring (SON) seasons locations in this region have zero or near-zero mean 95th percentile rainfall values. In the south-west, the largest 95th percentile rainfall values have occurred during the winter months, although there is less variability in this region than in the north. In the south-east, there is considerably less variability.

Examining the extreme frequency index in each season explains some of the issues with the large biases between mean annual extreme frequency at stations and AWAP grids. During the winter (JJA), stations in some parts of northern Australia experience very few rain-days and in some years no rainfall is recorded at all. This results in very low station mean extreme frequencies in winter at some stations, e.g. mean extreme frequency is 0.58 at Katherine (14.46 °S, 132.26 °E) and 0.84 at Oenpelli (12.33 °S, 133.06 °E) both located in the Northern Territory. The expected mean extreme frequency in each season is close to 4.5. The corresponding AWAP grids also have lower values for mean extreme frequency than expected, but still greater than 2.5 as there are more rain-days at the grids. This results in large positive biases at these locations.

The correlations between the station and grid extreme frequency time series remain high for each season, except where large biases exist. There are few obvious patterns in mean absolute difference or root mean square difference. The locations and signs of significant extreme frequency trends for summer and winter (Figure 11) show high variability between seasons and AWAP captures the pattern seen in the station data. No location has statistically significant trends of opposing signs between the station and gridbox in any season. Whilst there are many locations where only the station or the AWAP grid has a significant trend, the areas represented are generally the same. For example, in winter, in south-west Western Australia, both AWAP and the stations show a clear decreasing trend in the extreme frequency index, despite there being some differences in the locations where these significant trends are observed between the two datasets.

The seasonal extreme intensity and contribution indices also show some seasonal variability. The heaviest four rain events in each season were used to construct these indices, thus making these seasonal indices less ‘extreme’ than the annual indices (where the heaviest four rain events in each year were used). This was done in order to make the sample size large enough. One difficulty is that when investigating the extreme contribution if there are four or fewer rainfall events in a given season, the extreme contribution for that season will be 100%. This index is, therefore, less useful when investigating seasonal rainfall extremes, unless a smaller sample is taken. The biases are less consistent when looking at an individual season; however, they are still largely negative at sites where there are enough rainfall events in each season (i.e. AWAP gridboxes tend to have lower mean index values through the time series). A bias correction would still be reasonable even when examining extreme rainfall in individual seasons, although the error on the bias correction is likely to be increased. There are few significant trends in the seasonal extreme intensity or extreme contribution (not shown).

4.5. Investigating AWAP performance in remote regions

The analysis of AWAP performance in capturing rainfall characteristics at locations where station density is high is useful, however, much of Australia has low station density and an examination of AWAP performance in these regions is also needed. Giles (25.03 °S, 128.30 °E) is located in the east of inland Western Australia around 60 km west of the border with Northern Territory (Figure 1). It is not in the high-quality dataset; however, it is a Bureau of Meteorology staffed site and has a complete and relatively long record of daily rainfall measurements for stations in this region (extending back to August 1956), making it the most suitable station for comparison with AWAP. Figure 12 shows the annual extreme intensity time series and station precipitation plotted against the AWAP grid precipitation at Giles. The station and AWAP time series of extreme intensity have a correlation in excess of 0.99, considerably higher than the correlations between stations and their corresponding AWAP grids at all other locations for all indices. The root mean square differences and mean absolute differences are also much lower than for other locations. The AWAP precipitation is well correlated with the station precipitation with considerably less spread in points around the line of best fit (Figure 12(b)). The gradient of the line of best fit (0.99) is closer to 1 than at any of the other 119 locations considered. The AWAP grid much more closely reflects the rainfall recorded at Giles due to the lack of nearby stations. Prior to rainfall measurements being first taken at Giles, there is a period when

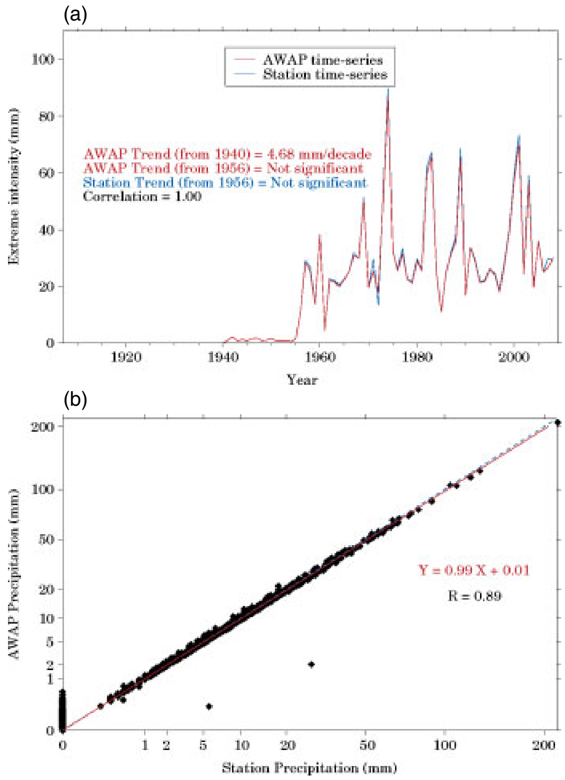

Figure 12. Plots of (a) time series of extreme intensity (mean of highest four daily rainfall totals each year) and (b) daily station precipitation against daily AWAP precipitation at Giles (25.03 °S, 128.30 °E). The trends in extreme intensity (shown in (a)) are not significant from September 1956 when Giles starts recording daily rainfall. In (b) lines of best fit are shown in solid red and the blue dashed line represents a 1 : 1 relationship. Equations of lines of best fit are shown as well as rank correlation (R) values.

AWAP interpolates data from other stations to the gridbox within which Giles is located. Rainfall values are considerably lower and there is a dramatic increase once measurements at Giles start being taken. This creates an artificial wetting trend in this area simply due to a station being set up in Giles.

Such dubious results suggest that AWAP performance in more remote areas of Australia may not be as good as in areas where station density is higher and that an appropriate mask, therefore, should perhaps be applied if using AWAP to investigate rainfall characteristics in Australia. Isolated stations with records shorter than the period of the AWAP dataset are likely to impact the AWAP dataset and, therefore, should be investigated further.

5. Conclusions

For 119 sites across Australia, high-quality station measurements of daily rainfall and corresponding grid data from AWAP have been compared with a focus on more extreme rainfall events. The study aimed to test whether the grid data shares the same characteristics as the station data and, therefore, whether it could justifiably be used to examine extreme rainfall.

It was found that AWAP tends to have lower extreme rainfall estimates than those observed at stations in the high-quality dataset, a result common to gridded analyses. It is important to note that this does not mean the AWAP value is 'wrong'. The AWAP estimate is for a $0.05^{\circ} \times 0.05^{\circ}$ gridbox, whereas the station measurement is for a single point. However, the extreme AWAP rainfall values tend to be lower than the extreme station rainfall measurements at all locations and this can be seen where all the station and AWAP values are plotted against each other. The mean climatological 95th percentile rainfall is lower at AWAP gridboxes than at stations at 86 of the 119 sites considered. The extreme intensity annual index is biased, being lower in AWAP grids at all but 3 of the 119 sites.

There are few obvious geographical patterns in the biases of extreme indices and the biases are generally consistent allowing for the possibility of applying a correction. The station and AWAP time series of extreme indices are strongly correlated and have similar standard deviations. This would suggest that AWAP could be used in a study on interannual or interdecadal variability of extreme rainfall. Also, the areas where significant trends in extreme rainfall have been observed are, largely, similar between the stations and AWAP. The strength of the correlation between all rainfall values at the station and corresponding gridbox varies between coastal and inland areas and is strongest in southern parts of Australia. This is likely to be, in part, due to the interpolation method used to produce AWAP, so that coastal stations play a larger role in influencing the AWAP gridbox value. It may also be related to the more homogeneous rainfall patterns in southern Australia due to the greater role of synoptic-scale systems in driving intense rainfall in this region.

AWAP gridded values more closely track station values if there are few other stations nearby, such as at Giles, Western Australia. However, stations dropping in and out of the network in remote areas appear to have a large influence on AWAP and so extreme caution is required when investigating rainfall characteristics in such regions. Application of a mask in such areas is suggested to avoid spurious results.

Overall, AWAP appears to be reasonably consistent with station values and, therefore, we believe that it could be used in further studies of extreme rainfall characteristics in Australia. However, caution still needs to be exercised. It is likely that some similar problems to those found in AWAP would appear in other gridded datasets, such as the over-smoothing of extreme precipitation and the appearance of spurious trends in data sparse areas. A thorough examination of the data should therefore be conducted before it is applied to an investigation into extreme rainfall variability or trends.

Acknowledgements

This study was supported by Australian Research Council grants CE110001028 and LP100200690. We thank the Bureau of Meteorology, the Bureau of Rural Sciences and CSIRO for providing the Australian Water Availability Project data. The station data used in this study was also provided by the Bureau of Meteorology. We are grateful to Blair Trewin and an anonymous reviewer whose suggestions helped to improve the quality of this article.

References

Aguilar E, Peterson TC, Ramírez Obando P, Frutos R, Retana JA, Solera M, González Santos I, Araujo RM, Rosa García A, Valle VE, Brunet India M, Aguilar L, Álvarez L, Bautista M, Castañón C, Herrera L, Ruano R, Siani JJ, Hernández Oviedo GI, Obed F, Salgado JE, Vázquez JL, Baca M, Gutíerrez M, Centella C, Espinosa J, Martínez D, Olmedo B, Ojeda Espinoza CE, Haylock M, Núñez R, Benavides H, Mayorga R. 2005. Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *Journal of Geophysical Research* **110**: D233107, DOI: 10.1029/2005JD006119.

Alexander LV, Arblaster JM. 2009. Assessing trends in observed and modelled climate extremes over Australia in relation to future projections. *International Journal of Climatology* **29**: 417–435, DOI: 10.1002/joc.1730.

Alexander LV, Hope P, Collins D, Trewin B, Lynch A, Nicholls N. 2007. Trends in Australia's climate means and extremes: a global context. *Australian Meteorological Magazine* **56**: 1–18.

Chen C-T, Knutson T. 2008. On the verification and comparison of extreme rainfall indices from climate models. *Journal of Climate* **21**: 1605–1621.

Gallant AJE, Hennessy KJ, Risbey J. 2007. Trends in rainfall indices for six Australian regions: 1910: 2005. *Australian Meteorological Magazine* **56**: 223–239.

Gerstner E-M, Heinemann G. 2008. Real-time areal precipitation determination from radar by means of statistical objective analysis. *Journal of Hydrology* **352**: 296–308, DOI: 10.1016/j.jhydrol.2008.01.016.

Groisman PY, Knight RW, Easterling DR, Karl TR, Hegerl GC, Razuvaev VAN. 2005. Trends in intense precipitation in the climate record. *Journal of Climate* **18**: 1326–1350.

Hanson CE, Palutikof JP, Livermore MTJ, Barring L, Bindi M, Corte-Real J, Durao R, Giannakopoulos C, Good P, Holt T, Kundzewicz Z, Leckebusch GC, Moriondo M, Radziejewski M, Santos J, Schlyter P, Schwarb MC, Stjernquist I, Ulbrich U. 2007. Modelling the impact of climate extremes: an overview of the MICE project. *Climatic Change* **81**: 163–177.

Haylock M, Nicholls N. 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998. *International Journal of Climatology* **20**: 1533–1541.

Haylock MR, Peterson TC, Alves LM, Ambrizzi T, Anunciação YMT, Baez J, Barros VR, Berlato MA, Bidegain M, Coronel G, Corradi V, Garcia VJ, Grimm AM, Karoly D, Marengo JA, Marino MB, Moncunill DF, Nechet D, Quintana J, Rebello E, Rusticucci M, Santos JL, Trebejo I, Vincent LA. 2006. Trends in total and extreme South American rainfall in 1960–2000 and links with sea surface temperature. *Journal of Climate* **19**: 1490–1512.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research* **113**: D20119, DOI: 10.1029/2008JD010201.

Hennessy KJ, Suppiah R, Page CM. 1999. Australian rainfall changes 1910–1995. *Australian Meteorological Magazine* **48**: 1–13.

Hofstra N, New M, McSweeney C. 2010. The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. *Climate Dynamics* **35**: 841–858.

Hunter RD, Meentemeyer RK. 2005. Climatologically aided mapping of daily precipitation and temperature. *Journal of Applied Meteorology* **44**: 1501–1510.

Hutchinson MF. 1995. Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems* **9**: 385–403.

Jones DA, Wang W, Fawcett R. 2009. High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal* **58**: 233–248.

Kharin VV, Zwiers F, Zhang X, Hegerl GC. 2007. Changes in temperature and precipitation extremes in the IPCC Ensemble of Global Coupled Model Simulations. *Journal of Climate* **20**: 1419–1444.

Klein Tank AMG, Zwiers FW, Zhang X. 2009. Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. WCDMP-72, WMO-TD/No.1500. 56 pp.

Koch SE, DesJardins M, Kochin PJ. 1983. An interactive Barnes objective map analysis scheme for use with satellite and conventional data. *Journal of Climatology and Applied Meteorology* **22**: 1487–1503.

Lavery B, Kariko A, Nicholls N. 1992. A historical rainfall data set for Australia. *Australian Meteorological Magazine* **40**: 33–39.

Li Y, Cai W, Campbell EP. 2005. Statistical modelling of extreme rainfall in southwest Western Australia. *Journal of Climate* **18**: 852–863.

Marinelli M, Braganza K, Collins D, Jones D, Maguire S, Cook G. In press. Defining a high-quality daily rainfall candidate network for Western Australia. *Australian Meteorological and Oceanographic Journal*.

Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C. 2007. Global climate projections. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Avery KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge and New York.

Sen Roy S. 2009. A spatial analysis of extreme hourly precipitation patterns in India. *International Journal of Climatology* **29**: 345–355.

Sillmann J, Roeckner E. 2008. Indices for extreme events in projections of anthropogenic climate change. *Climatic Change* **86**: 83–104.

Trewin B. 2001. *The Under-Reporting of Small Daily Rainfall Amounts*. The 14th Australia-New Zealand Climate Forum, Darwin.

Viney NR, Bates BC. 2004. It never rains on Sunday: the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high-quality data set. *International Journal of Climatology* **24**: 1171–1192.

Wilcox EM, Donner LJ. 2007. The frequency of extreme rain events in satellite rain-rate estimates and an atmospheric general circulation model. *Journal of Climate* **20**: 53–69.

Xie P, Yatagai A, Chen M, Hayasaka T, Fukushima Y, Liu C, Yang S. 2007. A gauge-based analysis of daily precipitation over east Asia. *Journal of Hydrometeorology* **8**: 607–626.

Zhang X, Alexander L, Hegerl GC, Jones P, Klein Tank A, Peterson TC, Trewin, B, Zwiers FW. 2011. Indices for monitoring changes in extremes based on daily temperature and precipitation data. *WIREs Climate Change* **2**(6): 851–870, DOI: 10.1002/wcc.147.

