INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **29**: 417–435 (2009)
Published online 21 July 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1730

RMetS
Royal Meteorological Society

Assessing trends in observed and modelled climate extremes over Australia in relation to future projections

Lisa V. Alexander[a,b]* and Julie M. Arblaster[c,d,e]

[a] *School of Geography and Environmental Science, Monash University, Clayton, VIC, Australia*
[b] *Met Office, Hadley Centre, Exeter, UK*
[c] *School of Earth Sciences, University of Melbourne, Melbourne, VIC, Australia*
[d] *National Center for Atmospheric Research, Boulder, CO, USA*
[e] *Bureau of Meteorology Research Centre, Melbourne, VIC, Australia*

ABSTRACT: Multiple simulations from nine globally coupled climate models were assessed for their ability to reproduce observed trends in a set of indices representing temperature and precipitation extremes over Australia. Observed trends over the period 1957–1999 were compared with individual and multi-modelled trends calculated over the same period. When averaged across Australia, the magnitude of trends and interannual variability of temperature extremes were well simulated by most models, particularly for the index for warm nights. The majority of models also reproduced the correct sign of trend for precipitation extremes although there was much more variation between the individual model runs. A bootstrapping technique was used to calculate uncertainty estimates and also to verify that most model runs produce plausible trends when averaged over Australia. Although very few showed significant skill at reproducing the observed spatial pattern of trends, a pattern correlation measure showed that spatial noise could not be ruled out as dominating these patterns. Two of the models with output from different forcings showed that the observed trends over Australia for one of the temperature indices was consistent with an anthropogenic response, but was inconsistent with natural-only forcings. Future projected changes in extremes using three emissions scenarios were also analysed. Australia shows a shift towards warming of temperature extremes, particularly a significant increase in the number of warm nights and heat waves with much longer dry spells interspersed with periods of increased extreme precipitation, irrespective of the scenario used. Copyright © 2008 Royal Meteorological Society

KEY WORDS Australian climate; extremes; observations; climate models; projections

Received 7 September 2007; Revised 10 February 2008; Accepted 7 May 2008

1. Introduction

Extremes research is particularly important for Australia, given the vulnerability of its unique flora, fauna and ecosystems to even slight variations in climate (Fitzharris *et al*., 2007). A previous body of work concluded that significant changes in temperature and precipitation extremes have already occurred across the country during the 20th century (e.g. Hennessy *et al*., 1999; Plummer *et al*., 1999; Collins *et al*., 2000; Haylock and Nicholls, 2000; Trewin, 2001; Alexander *et al*., 2007; Gallant *et al*., 2007). Regional studies across the Asia–Pacific area (e.g. Manton *et al*., 2001; Griffiths *et al*., 2005) have shown statistically significant increases in occurrences of hot days and warm nights and decreases in occurrences of cool days and cold nights over the past few decades. Over the past century, there has been a significant decrease in the frequency and intensity of extreme precipitation events in the southwest region of Western Australia and a significant increase in the proportion of total precipitation from extreme events in eastern Australia (Haylock and Nicholls, 2000; Li *et al*., 2005). While these studies have been thorough, they have focussed on the analysis of extremes at station locations. This makes it difficult to compare observations objectively with simulations from climate models that output data on spatial grids. Some work has suggested increases in hot days and hot spells and decreases in cold days and cold spells in the future (CSIRO, 2001) and an increase in extreme precipitation (Groisman *et al*., 2005), but relatively little has been published about how extremes might change in the future over Australia or, indeed, if climate models are able to adequately reproduce the observed trends in extremes (thus increasing our confidence in future projections).

Global studies comparing observed and modelled trends in climate extremes have shown reasonably good agreement with temperature trends but poor agreement (or multi-model disagreement) with observed precipitation patterns or trends (e.g. Kharin *et al*., 2007; Kiktev *et al*., 2007). Kiktev *et al*. (2007) also comment that a 'super ensemble' from multiple climate models appears to perform better than any individual ensemble member

* Correspondence to: Lisa V. Alexander, School of Geography and Environmental Science, Monash University, Clayton, Vic 3800, Australia. E-mail: Lisa.Alexander@arts.monash.edu.au

or model, particularly when there is some skill in the contributing ensemble members. Kiktev *et al*. (2003) found that only the inclusion of human-induced forcings in a climate model could account for observed changes in global temperature extremes. Robust anthropogenic changes have been detected globally in indices of extremely warm nights, although with some indications that the model overestimates the observed warming (Christidis *et al*., 2005). However, other recent studies show that the *regional responses* of observed trends in temperature and precipitation extremes can also largely be driven by large-scale processes, which might not be adequately simulated in global climate models (GCMs)(Meehl *et al*., 2004, 2005; Scaife *et al*., 2008). On regional scales, (e.g. Sillmann and Roekner, 2008 for Europe; Meehl and Tebaldi, 2004 and Meehl *et al*., 2007a for USA) while the modelled trends have been shown to capture observed trends reasonably accurately, the results are somewhat dependent on the extreme under consideration. To date, no such analysis has been carried out for Australia.

Recent initiatives by the World Climate Research Programme (WCRP) in preparation for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) in 2007 have now made it possible to compare multiple model simulations with high-quality observations of extremes. On the request of the Joint Scientific Committee (JSC)/CLIVAR Working Group on Coupled Models, coupled modelling groups worldwide submitted a standard set of 'extremes indices' to the WCRP's Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset at the Program for Climate Model Diagnosis and Intercomparison (PCMDI) in California (hereafter named the CMIP3 archive). Ten extremes indices, calculated from daily data and based on the definitions of Frich *et al*. (2002), were submitted with five temperature-based indices (e.g. heat wave duration, occurrence of frosts) and five precipitation-based indices (e.g. heavy precipitation events, consecutive dry days). The study by Tebaldi *et al*. (2006) was the first one to analyse these indices for both historical and future simulations on global and hemispheric scales. Global averages of the temperature-based indices were found to be dominated by trends in the Northern Hemisphere, with Southern Hemisphere trends much weaker. This is intuitively explained by the presence of a relatively stable ocean in close proximity to all Southern Hemisphere landmasses, sheltering them from extreme cold and warm air masses. However, in that study, observational datasets with which to compare the multi-model output were not yet adequate or available, so an assessment of the ability of the models to reproduce observed trends in extremes was not possible. The Commission for Climatology (CCl)/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) [previously known as the Expert Team on Climate Change Detection, Monitoring and Indices (ETCCDMI), see http://www.clivar.org/organization/etccdi/etccdi.php for details], initiated a project aimed at addressing gaps in observed data availability and analysis in previous global studies (e.g. Frich *et al*., 2002). Following on from this, Alexander *et al*. (2006) updated and extended the analysis of Frich *et al*. (2002) using the best global observations available, gridding a total of 27 indices onto a regular latitude–longitude grid from 1951 to 2003.

The CMIP3 multi-model dataset and recently available high-quality gridded datasets provide us with an unprecedented opportunity to directly compare observed trends in extremes over Australia with multiple model-simulated trends and to compare projections in these extremes, both across models and across scenarios. In this study, we first briefly discuss how the extremes indices are calculated from the observed and modelled datasets followed by a comparison of the models with observations and the future projections for the selected temperature and precipitation extremes across Australia.

2. Extremes indices data

The indices used in this study, based on the definitions of Frich *et al*. (2002), are given in Table I. Nine annual indices are analysed (four derived from daily maximum and/or minimum temperature and five from daily precipitation) providing one value per grid box per year per index. The indices chosen contain more robust statistical properties than could be expected from the analysis of more infrequent events and allow GCMs the possibility to adequately simulate these events. Therefore, some of the indices may be viewed as not particularly 'extreme', but given their statistical properties and their availability in the CMIP3 archive, we chose to use these definitions as the basis for our analysis. Note that the indices given by Frich *et al*. (2002) also contained a definition for *growing season length*, which is not analysed here since it has little meaning for the Australian climate (Collins *et al*., 2000). *Frost days* are included in the analysis but note that this index is only meaningful for parts of southern Australia. There are some differences between the observed and model definitions for three of the indices (Table I) and the potential effects these could have on the results are discussed in Section 3.3. Model indices for which climatologies were required were calculated relative to each model's own climatology, thus partially removing inherent model bias.

2.1. Observations

High-quality daily maximum and minimum temperature (Trewin, 2001) and daily precipitation (Haylock and Nicholls, 2000) data composed the Australian contribution to the Alexander *et al*. (2006) study, which created 2.5° of latitude by 3.75° of longitude gridded datasets (HadEX) of observed extremes indices for the globe (data available from www.hadobs.org). Extremes indices were first calculated for each station and then were transformed to the grid. For this study, we extract from this dataset those grid boxes that cover the Australian continent for each extremes index from Table I. Alexander *et al*. (2006) used a distance weighting method, which

DOI: 10.1002/joc

Table I. Extremes indices used in this study.

Index name	Index definitions		Units
	Model (Frich *et al*., 2002)	**Obs** (Alexander *et al*., 2006)	
Warm nights (**TN90**)	Percent of time Tmin > 1961–1990 90th percentile of daily minimum temperature	Percentile calculation differs from model definition in that the bootstrapping technique of Zhang *et al*., 2005 is used	%
Frost days (**FD**)	Total number of days with absolute minimum temperature <0 °C	As model	days
Extreme temperature range (**ETR**)	Difference between the highest and lowest temperature observation in a calendar year	As model	°C
Heat wave duration (**HWDI**)	Maximum period >5 consecutive days with Tmax >5 °C above the 1961–1990 daily Tmax normal	Known as *warm spell duration index* (WSDI) – maximum period >5 consecutive days with Tmax > 1961–1990 90th percentile of daily maximum temperature. Percentiles are calculated using the bootstrapping technique of Zhang *et al*., 2005 and spells can continue across calendar years	days
Heavy precipitation days (**R10**)	Number of days with precipitation ≥10 mm	As model	days
Maximum 5-day precipitation (**R5D**)	Maximum precipitation total over a 5-day period	As model	mm
Simple daily intensity (**SDII**)	Ratio of annual total precipitation to number of days ≥1 mm	As model	mm d^{-1}
Consecutive dry days (**CDD**)	Maximum number of consecutive days <1 mm	Basic definition is the same as model except a spell can continue across calendar years	days
Very heavy precipitation contribution (**R95T**)	Fraction of annual total precipitation due to events exceeding the 1961–1990 95th percentile	As model	%

required that at least three stations be within a pre-defined search radius from the centre of a grid box, in order for an extreme to be calculated for that grid box. Since Australia is large but sparsely populated, high-quality observations tend to be lacking in more remote areas. This means that for some of the indices (especially the precipitation indices, which have small decorrelation length scales) there is little or no coverage in inland or northern areas. Other indices such as *warm nights*, however, provide almost complete observational coverage over the country.

2.2. Model data

Extremes indices from nine models were available for inclusion in the IPCC AR4, as analysed by Tebaldi *et al*. (2006) and presented in the AR4 by Meehl *et al*. (2007). As noted above, each modelling group calculated the indices based on the definitions of Frich *et al*. (2002) and submitted them to the CMIP3 archive at PCMDI (http://www-pcmdi.llnl.gov). There were four models from the USA (CCSM3, PCM, GFDL-CM2.0, and GFDL-CM2.1), three from Japan (MIROC3.2 (medres), MIROC3.2 (hires), and MRI-CGCM2.3.2), one from France (CNRM-CM3) and one from Russia (INM-CM3.0). Simulations of the climate of the 20th century (20C3M) and three special report on emissions scenarios (SRES) experiments, B1 (low-range emissions), A1B (mid-range emissions) and A2 (high-range emissions) were available for most models. Each model varies in resolution, but the indices from each of the nine models were interpolated here onto the observational grid, i.e. 2.5° of latitude by 3.75° of longitude so that a direct comparison between the observations and the model simulations could be made. Multiple ensemble members were submitted for five out of the nine models (PCM, GFDL-CM2.0, GFDL-CM2.1, MIROC3.2 (medres), and MRI-CGCM2.3.2), with single runs available for the remaining four. In total, there were 22 20C3M simulations from the nine models. The multi-model mean values shown here are the average across all ensemble members and then across all models.

3. Comparison between observed and modelled extremes over Australia

3.1. Spatial and temporal comparison

To compare the modelled and observed indices, trends were calculated between 1957 and 1999 for each grid box with available data. The start date was chosen as the date from when high-quality temperature station data are available for Australia (Trewin, 2001) and the end

DOI: 10.1002/joc

date chosen based on when some of the model groups end their climate of the 20th century simulations. Trends in precipitation indices were also calculated over this period for consistency even though high-quality station data exist prior to this (Haylock and Nicholls, 2000). In all cases, for computational efficiency, trends are calculated using ordinary least squares (OLS) regression and trend significance is calculated at the 5% level using a non-parametric Mann–Kendall test (Mann, 1945; Kendall, 1975). However, because OLS is sensitive to outliers in the series, which may be present in the extremes indices analysed here, an additional non-parametric iterative technique to estimate trends and significance (Wang and Swail, 2001) was used in some cases to test the robustness of the OLS results. This method makes no assumptions about the distribution of the time series residuals and is robust to the effect of outliers in the series. Our general conclusions, however, remain unchanged irrespective of the trend calculation method used. Trends were only calculated in grid boxes if at least 40 out of the 43 years of observed indices data were available. In order that comparable analyses could be performed, the model output was masked by the regions where observed trend data exist. In addition to trend calculation, time series were produced for each of the nine indices using areally averaged data from the masked grid boxes to compare the magnitude and interannual variability of the observed and simulated extremes and these are plotted in Figure 1. Average trends for Australia and associated significance for the observations and multi-model ensemble are given in Table II, while ensemble mean trends for each individual model are given in Table III. To assess the uncertainty in the multi-modelled trends, we provide confidence intervals using a bootstrapping technique described in Section 3.2 with all 22 model runs. The associated uncertainties in the trend calculation method for the observations are given as two standard errors using restricted maximum likelihood (Trenberth *et al*., 2007). Spatial trend patterns of the observations and multi-model simulations for each of the nine extremes indices were compared for temperature (Figure 2) and precipitation (Figure 3).

3.1.1. *Temperature extremes*

Trends for the observed and multi-model simulation are given in Table II while the mean trends associated with each of the nine models used in this study are given in Table III. Table II shows that all of the observed trends for the four temperature indices are statistically significant and commensurate with warming. The majority of models (Table III) and individual model runs (not shown)

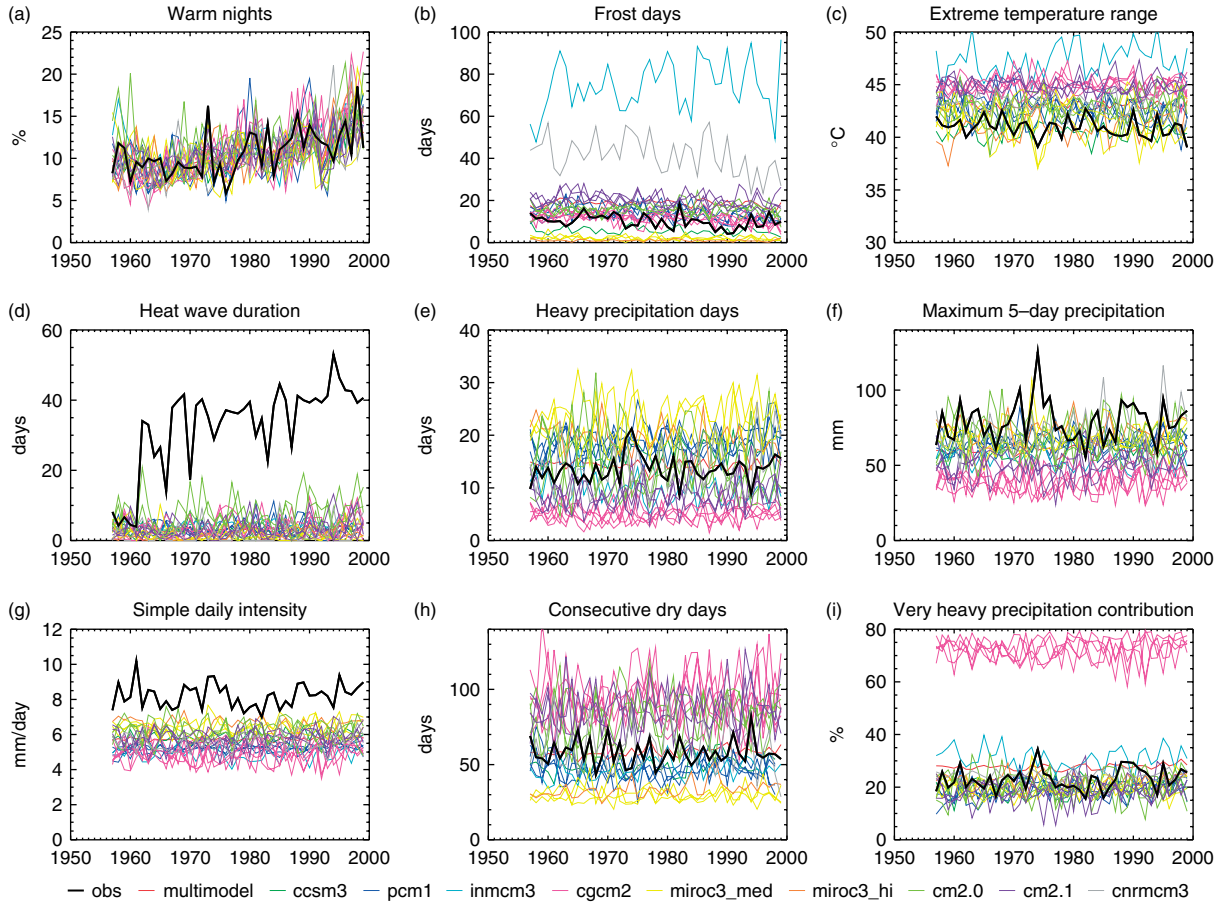

Figure 1. Observed (black line) and modelled (grey lines) time series of areally averaged extremes indices (Frich *et al*., 2002) from 1957 to 1999 using grid boxes in Australia with observed data. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

Table II. Observed and simulated decadal OLS trends calculated over 1957–1999 for each index (Table I) averaged across Australia using grid boxes containing observations from Figures 2 and 3. Boldface signifies trends that are significant at 5% level. Observations are shown with two standard errors in the trend calculation estimated using restricted maximum likelihood (Trenberth *et al*., 2007) while 10–90% confidence intervals are shown in brackets for the model data by randomly resampling the bootstrapped trends (see text) across all model runs to give an estimate of the uncertainty from using multiple model simulations. Units as Table I (per decade).

Index	Obs	Multi-model
Warm nights	**1.11** ± 0.06	**1.15** (0.48/1.87)
Frost days	**−0.89** ± 0.07	−0.19 (−1.46/0.22)
Extreme temperature range	**−0.19** ± 0.02	0.04 (−0.29/0.31)
Heat wave duration	**7.05** ± 0.33	**0.26** (−0.31/0.91)
Heavy precipitation days	0.28 ± 0.06	−0.06 (−0.79/0.89)
Maximum 5-day precipitation	0.42 ± 0.33	0.32 (−1.37/2.32)
Simple daily intensity	0.04 ± 0.02	0.02 (−0.06/0.13)
Consecutive dry days	−0.14 ± 0.15	**1.04** (−1.68/3.36)
Very heavy precipitation contribution	0.60 ± 0.12	0.26 (−0.58/1.23)

obtain the correct sign of trend for each temperature index when averaged across Australia. In fact, 21 of the 22 model runs exhibit a statistically significant increasing trend in *warm nights*; the multi-model trend of 1.15% per decade is comparable to the observed trend of 1.11% per decade. Eight of the nine models and the multi-model ensemble produce trends in *frost days* of the same sign as the observed trends. Five of the nine models agree with the observations that *extreme temperature range* is decreasing on average and seven of the nine models show increases in *heat wave duration* in agreement with the observed trends. However, the confidence intervals for these two indices compared to *warm nights* and *frost days* shows that there is much greater uncertainty both within and between models, as to the sign of the trend over the latter part of the 20th century. Moreover, while the sign of the temperature trends was correct, the magnitude was generally very different from the observed values. *Warm nights* is the only index where the confidence intervals in all models and multi-model ensemble do not overlap with zero, indicating that there is very good consensus between the models that the recent trend in *warm nights* over Australia is positive. The lowest resolution model, INM-CM3.0, is the only model that shows an increase in *frost days* while the highest resolution model, MIROC3 (hires), is the only model that

Table III. As Table II, but showing OLS trends and 10 and 90% confidence intervals for each model used in this study. For models where the ensemble mean is calculated from multiple simulations, the confidence intervals are calculated using all ensemble members.

	Temperature extremes			
	Warm nights	Frost days	Extreme temperature range	Heat wave duration
CCSM3	**1.41** (1.11/1.71)	−0.47 (−0.73/−0.18)	−0.02 (−0.16/0.12)	0.02 (−0.24/0.30)
PCM	**1.22** (0.60/1.91)	**−0.92** (−1.58/−0.01)	**−0.19** (−0.40/0.03)	−0.09 (−0.50/0.30)
INM-CM3.0	**0.81** (0.34/1.24)	2.44 (0.46/4.44)	0.31 (0.06/0.54)	0.10 (−0.09/0.27)
MRI-CGCM2.3.2	**1.71** (1.13/2.60)	**−0.91** (−1.33/−0.45)	0.02 (−0.18/0.23)	0.30 (−0.17/0.77)
MIROC3.2(med)	**0.78** (0.30/1.26)	**−0.19** (−0.38/−0.04)	**−0.22** (−0.55/0.08)	−0.18 (−0.52/0.17)
MIROC3.2 (hi)	**1.37** (0.96/1.78)	−0.04 (−0.12/0.03)	**0.48** (0.30/0.67)	**0.62** (0.41/0.81)
GFDL-CM2.1	**0.85** (0.33/1.31)	−0.07 (−0.52/0.42)	0.14 (−0.06/0.35)	**0.75** (0.11/1.51)
GFDL-CM2.0	**0.78** (0.50/1.08)	**−0.78** (−1.48/−0.10)	−0.02 (−0.17/0.13)	**0.39** (−0.06/0.83)
CNRM-CM3	**1.59** (1.16/2.01)	**−2.91** (−4.27/−1.62)	−0.17 (−0.32/−0.01)	0.37 (−0.07/0.83)

	Precipitation extremes				
	Heavy precipitation days	Maximum 5-day precipitation	Simple daily intensity	Consecutive dry days	Very heavy precipitation contribution
CCSM3	0.01 (0.33/0.36)	0.25 (−0.89/1.23)	0.06 (0.02/0.09)	1.20 (0.05/2.18)	0.15 (−0.41/0.63)
PCM	**0.51** (−0.03/1.43)	**1.31** (0.19/2.41)	**0.05** (−0.01/0.12)	−0.27 (−1.82/1.59)	**0.83** (0.30/1.45)
INM-CM3.0	−0.41 (−0.78/−0.03)	**−1.54** (−2.48/−0.63)	−0.04 (−0.11/0.02)	**1.65** (0.67/2.63)	−0.25 (−0.92/0.39)
MRI-CGCM2.3.2	0.08 (−0.32/0.48)	0.58 (−0.79/2.00)	**0.05** (−0.07/0.17)	0.64 (−2.79/5.02)	0.26 (−0.56/1.13)
MIROC3.2(med)	0.33 (−0.36/0.97)	0.31 (−1.41/2.70)	0.03 (−0.04/0.09)	−0.07 (−0.79/0.62)	0.15 (−0.53/0.86)
MIROC3.2 (hi)	**−1.07** (−1.52/−0.64)	−0.04 (−1.25/1.19)	−0.03 (−0.09/0.04)	**1.95** (1.12/2.80)	0.08 (−0.47/0.60)
GFDL-CM2.1	**−0.70** (−1.64/0.14)	−0.42 (−3.11/2.03)	−0.02 (−0.13/0.09)	**2.58** (−0.15/5.16)	0.04 (−0.98/1.07)
GFDL-CM2.0	0.25 (−0.16/0.68)	1.01 (−0.29/2.42)	0.05 (−0.03/0.13)	0.80 (−1.09/2.76)	0.51 (−0.25/1.34)
CNRM-CM3	0.89 (0.21/1.56)	2.95 (0.99/5.03)	0.07 (−0.01/0.16)	−1.18 (−3.53/1.11)	0.15 (−0.10/0.38)

DOI: 10.1002/joc

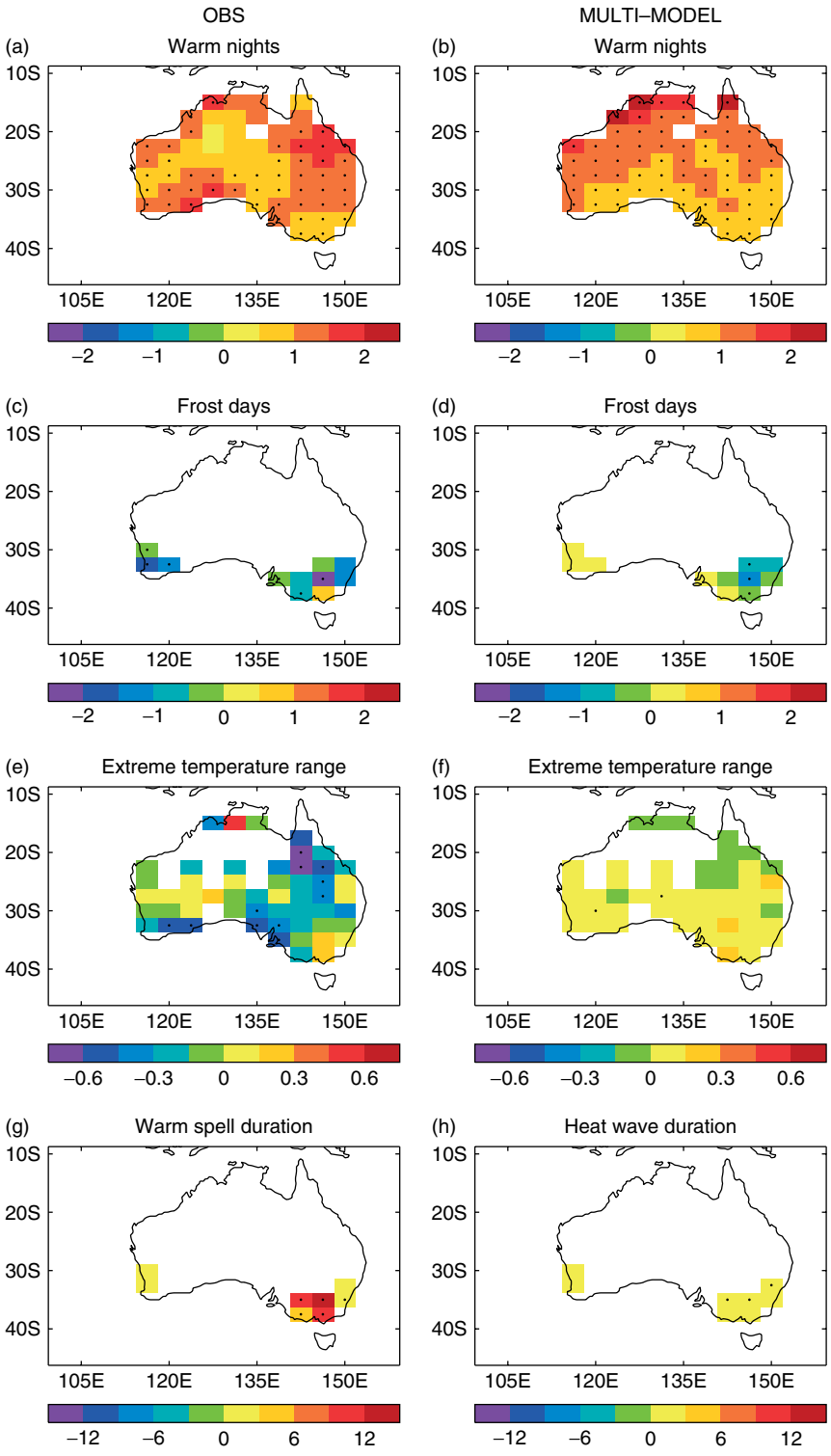

Figure 2. Observed (left column) and modelled (right column) decadal trends calculated between 1957 and 1999 for extreme temperature indices (Table I) for Australia. Model data are masked with grid boxes which have observed data. Stippling indicates trend significance at the 5% level. Units as in Table I (per decade).

shows a significant increase in *extreme temperature range* contrary to the observed trends. Figure 1 shows that the models also do reasonably well at simulating the amount and variability of the temperature extremes. This is particularly true of *warm nights*. However, no one model is consistently 'best' across all indices. Figure 1(a) and Table III show that all models are particularly good at simulating the amount, interannual variability and trend of this index. It is obvious, however, that some models are overestimating the actual value of some of the temperature extremes indices. In addition to showing increased trends in *frost days*, INM-CM3.0 (Table III) also vastly overestimates the amount of frosts that actually occur in Australia (Figure 1(b)) and this is likely

DOI: 10.1002/joc

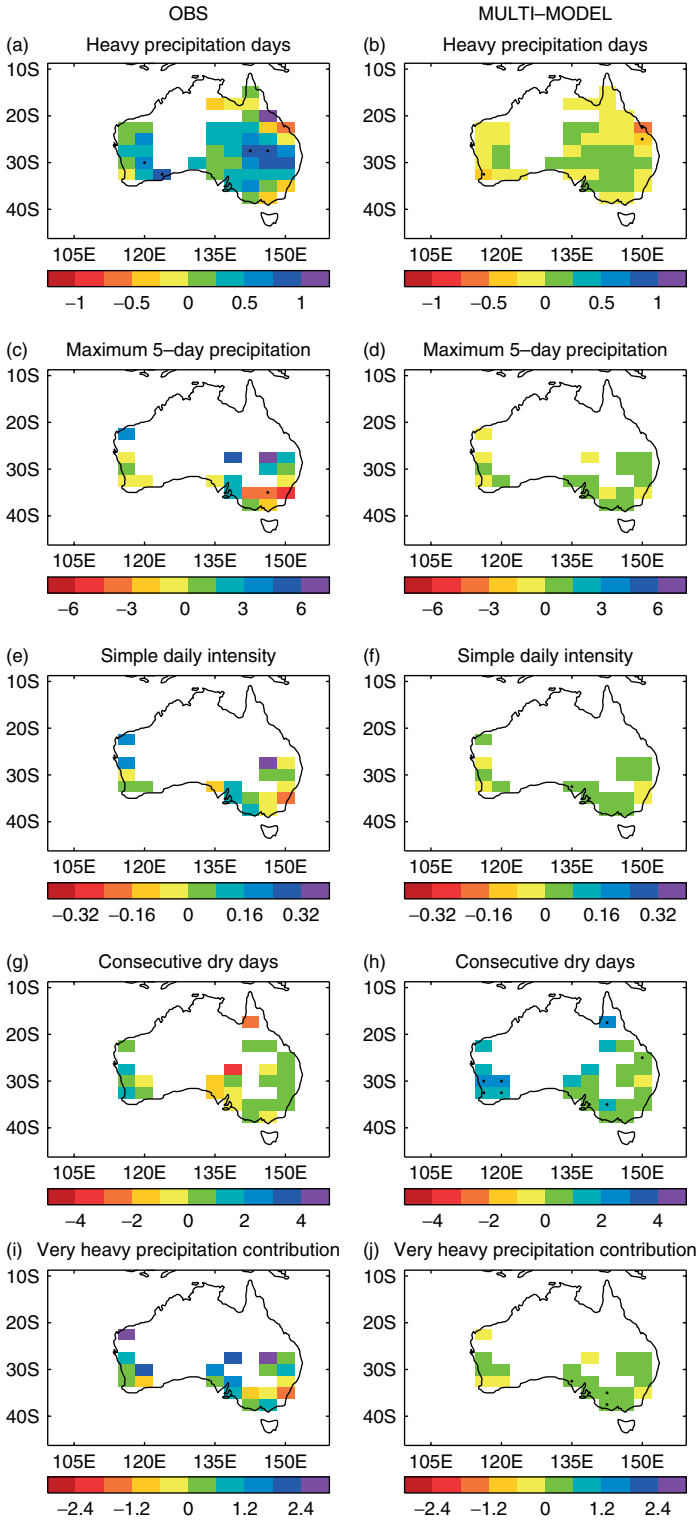

Figure 3. As in Figure 2, but for extreme precipitation indices (Table I).

to contribute to its overestimate of *extreme temperature range* (Figure 1(c)). Similarly, the CNRM-CM3 model also overestimates the amount of *frost days* and *extreme temperature range* although it gets the right sign of trend for both indices. It is likely that the vastly different amounts and magnitude of trends in *heat wave duration* (Figure 1(d); Tables II and III) are related to the different definitions of this index between the model and observations (Table I) and this is discussed in Section 3.3. In addition, this index definition is statistically 'volatile' (e.g. it contains a lot of zeros and no values between one and five) and is particularly sensitive to missing data. Given that we allow a maximum of 3 years of missing data before a trend can be calculated and most of the missing data occurs early on in the record, this creates an apparent inhomogeneity in the observations near the

DOI: 10.1002/joc

beginning of the time series (Figure 1(d)). It is likely, therefore, that the observed trend in *heat wave duration* is smaller than indicated in Table II. However, we chose not to remove this index from the study since the models may be doing a reasonable job using this definition and future changes in this index may have pronounced societal impacts.

The spatial trend patterns of temperature extremes across Australia for the observations and multi-model are shown in Figure 2. As noted, the majority of models are able to simulate the observed sign of change in the temperature indices when averaged across Australia; however, it is clear from Figure 2 that the regional trend pattern is less well captured by the models. While there is an observed increase in *warm nights* across northwest Australia over the period studied, it is small and mostly non-significant (Figure 2(a)). Indeed, if the analysis is extended beyond 1999, we see a non-significant decrease in *warm nights* over the northwest region (Alexander *et al*., 2006) and this is consistent with a cooling in mean minimum temperatures annually, but particularly associated with a decrease in minimum temperatures between December and August (Alexander *et al*., 2007). Figure 2(a) and Figure 2(b) show the multi-model ensemble is overestimating the observed trends in *warm nights* in the northwest and underestimating the observed trends in this index in southern and eastern Australia. In fact, there is very good consensus between the individual models (not shown) that the number of *warm nights* have increased significantly in the northwest region. However, while there is some discrepancy in the magnitude and significance between the observed and multi-modelled trends, it is noteworthy that the sign of the simulated trend in *warm nights* is consistent with the sign of the observed trend in every grid box across Australia. While there are few regions of Australia where *frost days* can be measured, (Figure 2(c)) almost all grid boxes show a consistent and mostly significant decline. Figure 2(d) indicates that the multi-model gets the opposite sign of trend to the observations in the southwest region of Western Australia and southwestern Victoria and there can be large differences between the observed and modelled response. Overall, the multi-model ensemble underestimates the observed trend in *frost days*, but the ensemble average from several individual models such as PCM and MRI-CGCM2.3.2 get very good approximations to the observed trend (-0.92 and -0.91 days/decade, respectively). The observed pattern of trends for *extreme temperature range* (Figure 2(e)) is generally not well simulated (Figure 2(f)). The trend pattern for *heat wave duration* is well simulated by the multi-model although trend magnitudes are underestimated in each grid box (Figure 2(g) and (h)), which, as noted previously, is likely to be related to definitional differences between the observations and models.

3.1.2. Precipitation extremes

There are no significant observed trends in the precipitation indices (Table II), which is perhaps not surprising, given that precipitation extremes are less well spatially correlated and have larger interannual variability over Australia than temperature extremes (Alexander *et al*., 2007). Also, it is clear from Tables II and III that there are generally wider confidence intervals on the simulated trends of precipitation extremes than temperature extremes. Precipitation is also not expected to respond as consistently or strongly to greenhouse gas forcing as temperature (e.g. Lambert *et al*., 2005). Given this, we might expect that it would be more difficult for climate models to capture observed trends in precipitation extremes. So, it is encouraging to find that the majority of models match the sign of the observed trend for four out of the five precipitation extremes. The exception is *consecutive dry days*, where six out of the nine models and the multi-model average have trends of opposite sign to the observations. Figure 1 shows that the models also do reasonably well at simulating the amount and variability of the precipitation extremes (Figure 1(e)–(i)). In addition, all models underestimate the actual amount of observed *simple daily intensity* (Figure 1(g)) while most models also underestimate *maximum 5-day precipitation* amount (Figure 1(f)), although the trend averaged over Australia from the multiple model simulations is close to the observed trend (Table II). The observations of *heavy precipitation days* (Figure 1(e)), *consecutive dry days* (Figure 1(h)) and *very heavy precipitation contribution* (Figure 1(i)) lie within the range of values for those indices simulated by the full suite of models. However, the model ranges are very large.

The spatial trend maps for the precipitation indices (Figure 3) show that it is mostly southern Australia that is covered by observational data. Even so, this corresponds to the region with the highest population density so that it provides useful information for future studies, which relate climate extremes to impacts. Unfortunately, no observed precipitation extremes data exist for the northwest region although current work at the Bureau of Meteorology is aimed at addressing this (Dörte Jakob, personal communication). This is unfortunate, since there is a well-established increase in the mean precipitation in this region since the 1950s, the reason for which is still being debated in the current literature (e.g. Wardle and Smith, 2004; Rotstayn *et al*., 2007; Shi *et al*., 2008). However, most models (not shown) simulate a decrease in *heavy precipitation days* over the northwest region but a mixed response regarding the trends in *simple daily intensity*. As noted, the multi-modelled trend for *maximum 5-day precipitation* is close to the observed trend but, Figure 3(c) and (d) shows that there are some differences in regional response. The multi-modelled trend in *simple daily intensity* also fails to capture some of the strong spatial gradients shown in the observations, e.g. Victoria in southeast Australia exhibits increasing trends in the western part of the state and decreasing trends in the east (Figure 3(e)), whereas the multi-model shows uniform increases in the intensity of precipitation across the region (Figure 3(f)). Observed trends in *consecutive dry*

DOI: 10.1002/joc

days (Figure 3(g)) have not been as uniformly increasing as simulated trends might suggest (Figure 3(h)); the overall trend across those parts of Australia with observed data shows that the simulated multi-model trend is significantly increasing (1.04% per decade) contrasting with the observed decreasing trend of −0.14% per decade (Table II). Also, as with *heavy precipitation days*, it is the models INM-CM3.0, MIROC3.2 (hires) and GFDL-CM2.1, with statistically significant increases in *consecutive dry days*, which largely influence this temporal trend (Table III). The southwestern region of Western Australia has seen a significant and well-documented decline in precipitation since the mid-1970s (IOCI, 2002) and this agrees with an increase in *consecutive dry days* in the region (Figure 3(g)). However, the observed increase is not as large or statistically significant as the models suggest (Figure 3(h)). In general, the simulated and observed spatial distribution of trends of *very heavy precipitation contribution* (Figure 3(i) and (j)) are not in good agreement. However, given that precipitation is a much less spatially coherent variable than temperature and that precipitation extremes could depend on specific convection or storm events, models would be expected to have a more difficult time simulating patterns of precipitation extremes than temperature extremes. Given this fact, it is quite impressive that the models capture some of the observed trends and trend patterns.

In the next section, measures of trend uncertainty are estimated for observed and modelled temperature and precipitation extremes to provide objective comparison of the temporal and spatial similarity between observed and modelled trends.

3.2. Measuring trend uncertainty

For each index, objective measures were calculated to assess the ability of the models to reproduce (1) observed area-averaged trends (temporal similarity) and (2) spatial patterns of observed trends (spatial similarity) over Australia. In each case, a bootstrapping technique was employed to assess the uncertainty associated with the modelled trend estimates over Australia during the latter part of the 20th century. To assess the uncertainty associated with temporal similarity, the modelled time series from Figure 1 are used to calculate the lines of best fit and associated residuals for each index. Next, a moving block bootstrap resampling (following the method of Wilks, 1997) is used to randomly resample the residuals in blocks of 2 years to maintain some of the temporal correlation in the time series. This procedure is performed 1000 times, adding the line of best fit back each time to the resampled residuals and recalculating the trend. Essentially, this produces a distribution of probable or 'plausible' climate trends for Australia. The same bootstrapping method is followed independently for each of the model simulation index time series from Figure 1. Probability distribution functions (PDFs) are then created using the 1000 bootstrapped trends for each index so that the observations and models can be compared. PDFs are centred on the original model trend and models with multiple simulations are combined into one PDF and centred on the ensemble mean trend. Figure 4 shows the temporal similarity PDFs for each index.

In most cases, the spread of plausible model trends overlaps with the observations. *Heat wave duration* (Figure 4(d)) is the only index where none of the PDFs of modelled trends overlap with the observed trends and this is probably associated with the different definitions used (Section 3.3). For *warm nights*, all models support a warming trend and indeed there is very little overlap of any of the model PDFs with zero (Figure 4(a)). This is also supported by the positive confidence intervals on the multi-model trend estimates in Table II. Trends in the other temperature indices, *frost days* (Figure 4(b)) and *extreme temperature range* (Figure 4(c)) are generally less well simulated than *warm nights* although, for instance, the median value of the PCM model PDF for both *frost days* and *extreme temperature range* is centred around the observed trend. Some other models do a relatively poor job of simulating these indices. For instance, the highest resolution model, MIROC3.2 (hires), and the lowest resolution model, INM-CM3.0, exhibit little or no overlap with the observations for both *frost days* and *extreme temperature range*. Note also that two of the models, CNRM-CM3 and INM-CM3.0, have a much larger spread than the rest of the models for trends in frost days (Figure 4(b)). Both these models have only one ensemble member, but the greater variance of the PDFs is most likely due to the larger interannual variability of simulated frost days by these models as shown in Figure 1(b).

To assess the uncertainty associated with spatial trend patterns, i.e. to determine the spatial similarity of trends between the observations and models, the bootstrapping technique is again employed, but time series are randomly resampled this time at each grid point to calculate trends. The bootstrapping is done synchronously at all locations to maintain the spatial coherence of the trends. This produces 1000 gridded fields of plausible spatial trend patterns for the observations and models to reflect the uncertainty associated with natural climate variability. From these 1000 fields, a spatial correlation statistic is calculated as follows. We randomly select an observed trend pattern and independently a model trend pattern. The area-weighted uncentred spatial correlation between these two patterns is calculated (this measure is similar to the congruence statistic described by Kiktev *et al*., 2007). The procedure is repeated 2500 times, and the resulting distribution of spatial correlation values is used to create PDFs for each model run and index. To ensure that the bootstrapped PDFs are measuring the uncertainty around the 'best estimate' of the models, the medians were centred on the original values of pattern similarity between the observed and modelled trend fields. The resulting PDFs are shown in Figure 5. In general, the higher the median value of the PDF the better a model is at simulating the observed pattern of trends for that index. The null hypothesis that the

DOI: 10.1002/joc

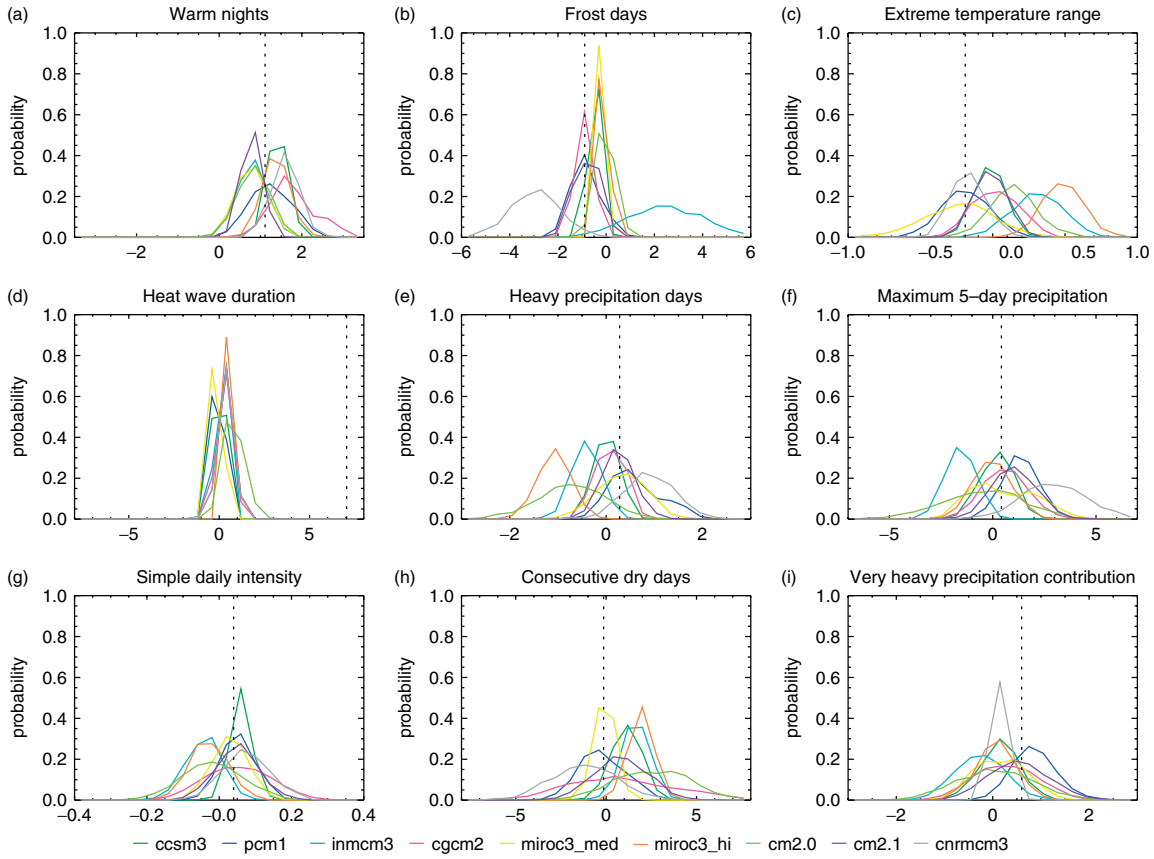

Figure 4. PDFs of plausible areally averaged OLS trends (1957–1999) over Australia using each of the nine climate models in the CMIP3 archive. PDFs are calculated using the 'temporal similarity' bootstrapping technique described in the text. Where there are multiple ensemble members, PDFs are centred on the ensemble mean trend. The dashed lines represent where the observed trends lie over the same period. Units on the x axis as in Table I (per decade). This figure is available in colour online at www.interscience.wiley.com/ijoc

models have no significant skill at reproducing observed spatial trend patterns is tested and rejected if a zero correlation falls within the lower 5% tail of the PDF. Very few of the individual model runs (Figure 5) and none of the ensemble means (not shown) from the nine models showed significant skill at reproducing the observed pattern of trends over Australia for any of the indices. Indeed, only the pattern of *maximum 5-day precipitation* (Figure 5(f)) is significantly well simulated by one run from each of the PCM and MRI-CGCM2 models. However, interestingly, the multi-model ensemble did show significant skill at simulating the trend pattern for *heavy precipitation days* even though there was no significant skill in any of the contributing models (not shown).

3.2.1. *Natural versus anthropogenic forcings*

Two out of the nine models (CCSM3 and PCM) are also available for analysis using natural-only and anthropogenic-only as well as all-forcings runs. The natural-only runs include only forcing from volcanic aerosols and solar variability, while anthropogenic-only runs include only forcing from greenhouse gases, sulphate aerosols, black carbon aerosols (CCSM3 only) and stratospheric ozone depletion. The natural and anthropogenic forcings for PCM and CCSM3 are described in more detail in Meehl *et al.* (2004) and Meehl *et al.* (2006), respectively. Again, the variability in the trends for each of the model runs is calculated using the bootstrapping procedure described above to measure both the temporal and spatial similarity between the observed and modelled trends. The resulting PDFs of temporal similarity are again analysed to assess how well the different forcings runs simulate the observed trends during the latter part of the 20th century for each index. In most cases, the PDFs for the different forcings runs overlapped with the observed trends indicating that, in these cases at least there was no discernible difference in the performance of the models between the natural and all-forcings runs. Figure 6 shows the PDFs using the different forcings from both the PCM and CCSM3 models for *warm nights* and *very heavy precipitation contribution*. For *warm nights,* the PDFs for the natural-only forcings do not overlap (or overlap less than 5% in the case of the PCM model) with the observed trends. Figure 6(a) and (b) show that it is only when anthropogenic forcings are included that the models are able to adequately simulate the observed trend. The precipitation extremes do not show any significant differences between the natural, anthropogenic and all-forcings runs although Figure 6(c) and (d) shows that there does appear to be some separation between the natural-only and all-forcings runs

DOI: 10.1002/joc

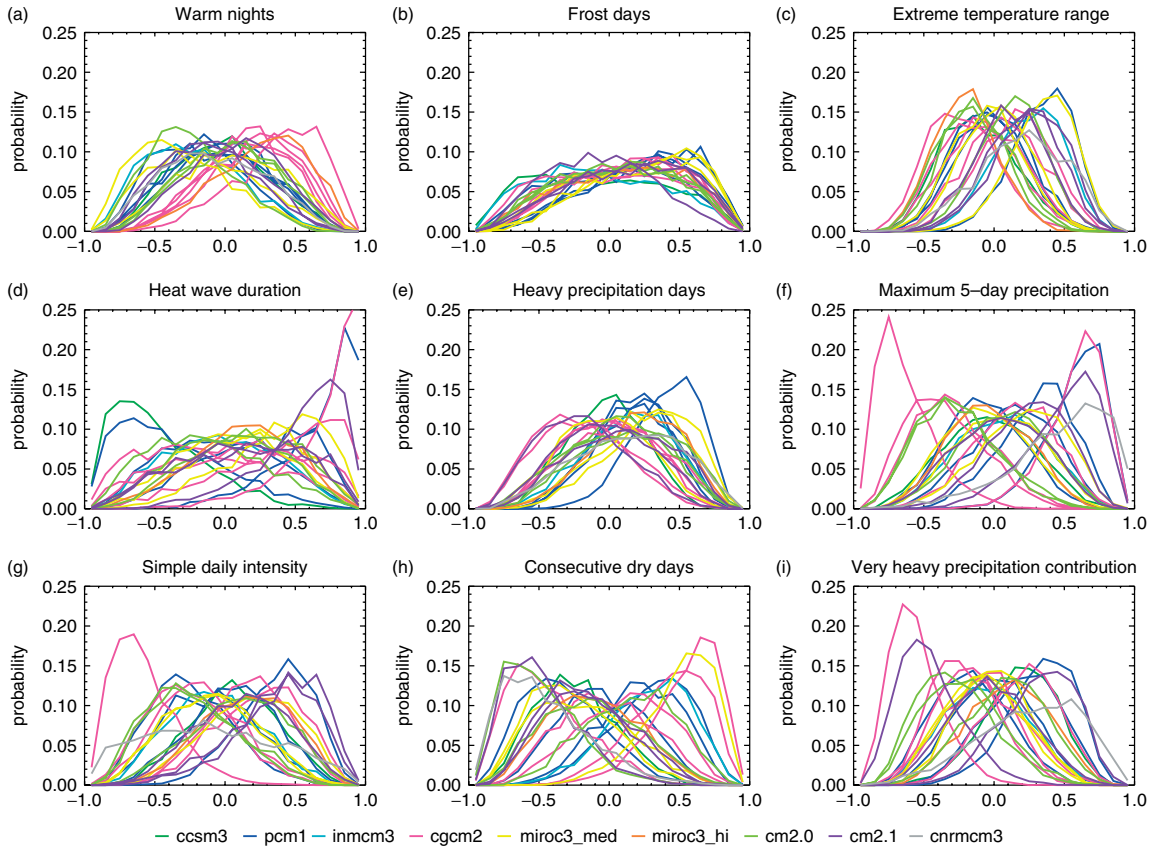

Figure 5. PDFs of the spatial trend correlations (calculated over 1957–1999) between observations and 22 runs from the nine CMIP3 models available for this study over Australia. PDFs are calculated using the 'spatial similarity' bootstrapping technique described in the text. PDFs are not shown for CNRM-CM3 *frost days* and *heat wave duration* due to the masking applied to this model at source, which reduces the number of grid boxes available for the spatial correlation calculation. This figure is available in colour online at www.interscience.wiley.com/ijoc

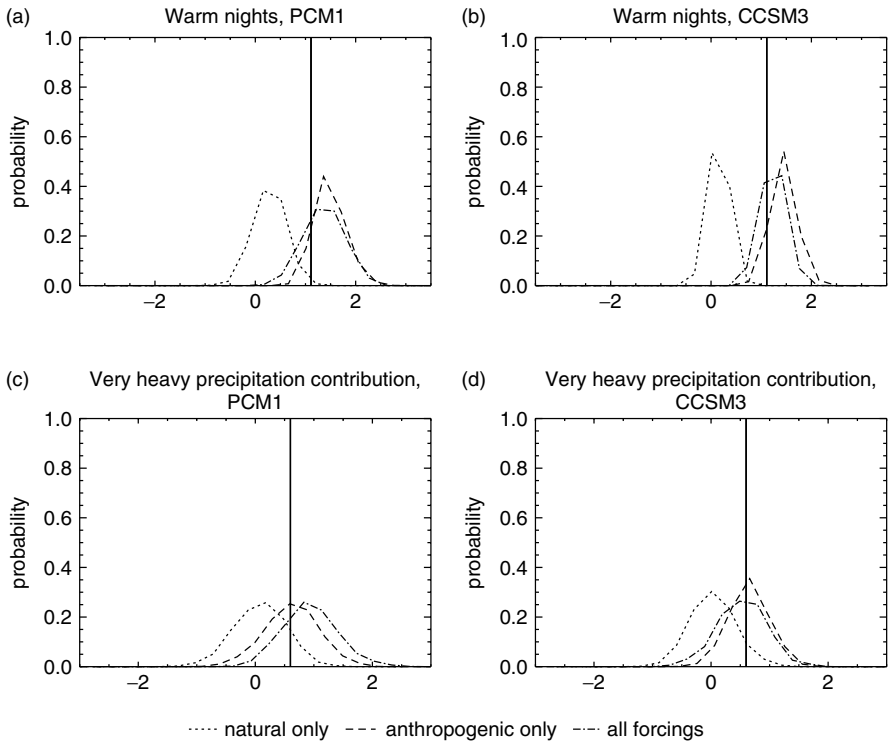

Figure 6. PDFs of annual OLS trends (1957–1999) in *warm nights* for (a) PCM and (b) CCSM3 and *very heavy precipitation contribution* for (c) PCM and (d) CCSM3 over Australia for natural only (dotted), anthropogenic only (dashed) and all forcings (dotted dashed). PDFs are centred on the ensemble mean trend. The solid lines represent where the observed trends lie over the same period. PDFs are calculated using the 'temporal similarity' bootstrapping technique described in the text and trends are calculated as in Figure 4.

DOI: 10.1002/joc

for *very heavy precipitation contribution*. However, when the PDFs of spatial similarity are compared with the observations, there is no significant skill in either the PCM or CCSM3 model in reproducing the spatial trend pattern of any index irrespective of which forcings are used (not shown). Figure 6, however, does highlight that the observed trends in at least one of the temperature extremes during the latter half of the 20th century when averaged over Australia is unlikely to have been produced from natural forcings alone.

3.3. Differences in index definitions

Potential errors or differences among the model results could be due to potentially different computational techniques used by each group. Another complication in comparing the modelled and observed indices is that when the Frich *et al*. (2002) analysis was updated by Alexander *et al*. (2006), some of the definitions using the observed data had to be redefined since statistical inconsistencies were discovered when the original definition was used. Three of the indices in this study have been affected by this update (Table I). In particular, it was shown that the original simple threshold calculation for the 90th percentile of minimum temperatures (*warm nights*) contained an inhomogeneity at the start and end of the 1961–1990 base period (Zhang *et al*., 2005). Unfortunately, these inconsistencies were discovered after the extremes indices had been submitted to the CMIP3 archive, so they could not be recalculated without access to the original daily model data. To try and assess how these definitional differences would affect our comparison, we plotted trends in the original station data for Australia that were used in Frich *et al*. (2002) with trends in the same stations using the Alexander *et al*. (2006) definitions. Figure 7 shows the comparisons for *heat wave duration, warm nights* and *consecutive dry days*, which are the indices where differences in definition occur. *Heat wave duration* has the lowest correlation between the two methods (0.17) and the biggest difference in the trends (the slope of 0.28 of the line of best fit using total least squares regression indicates that, in general, the trends in *heat wave duration* from Frich *et al*., 2002 are much larger than the *warm spell duration* defined by Alexander *et al*., 2006). *Warm nights* are reasonably well spatially correlated between the two methods (0.45), but the slope of the line of best fit (0.5) indicates that using the definition by Frich *et al*. (2002) produces trends about twice as large as those using the definition by Alexander *et al*. (2006) when averaged across Australia. *Consecutive dry days* are highly spatially correlated between the two methods (0.77) and the slope of the line of best fit is close to 1.0 indicating that the trends are reasonably comparable between the definitions by Frich *et al*. (2002) and Alexander *et al*. (2006).

We judge that to adjust for these changes would not be feasible, particularly since bias corrections would probably have to be calculated regionally and our station sample size is simply not large enough to do this rigorously. The best option would obviously be to recalculate these indices using the daily output from the nine climate models. However, at present the model data are not

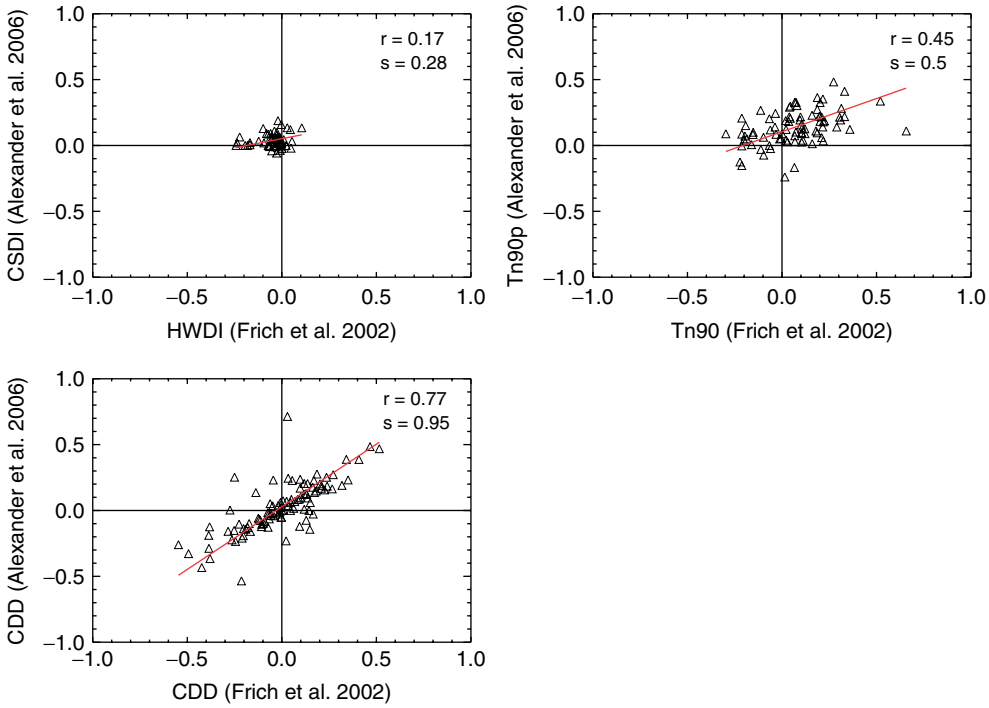

Figure 7. The relationship between two different definitions of the *heat wave duration, warm nights* and *consecutive dry days* indices (Table I) across Australia. Each triangle represents the annual trend calculated between 1957 and 1996 for each index at Australian stations. The solid line represents the line of best fit using total least squares regression, s is the slope of the line and r is the spatial correlation between all points. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

available for us to do this, so the differences discussed above should be considered when assessing the projected future changes in extremes presented in the next section.

4. Future projections of extremes over Australia

To put the future changes in extremes in context, changes in the mean temperature and precipitation for 2080–2099 are shown in Figure 8, along with observed and modelled trends for 1957–1999. Similar to previous studies (e.g. Smith, 2004; Karoly and Braganza, 2005; Gallant *et al*., 2007), temperatures have increased almost Australia-wide, while precipitation trends are of mixed sign. Observed precipitation (Figure 8(b)) has decreased in most of eastern Australia and the southwest (IOCI, 2002) and increased in the northwest (Smith, 2004). The multi-model trends in temperature (Figure 8(c)) and precipitation (Figure 8(d)) capture most of these

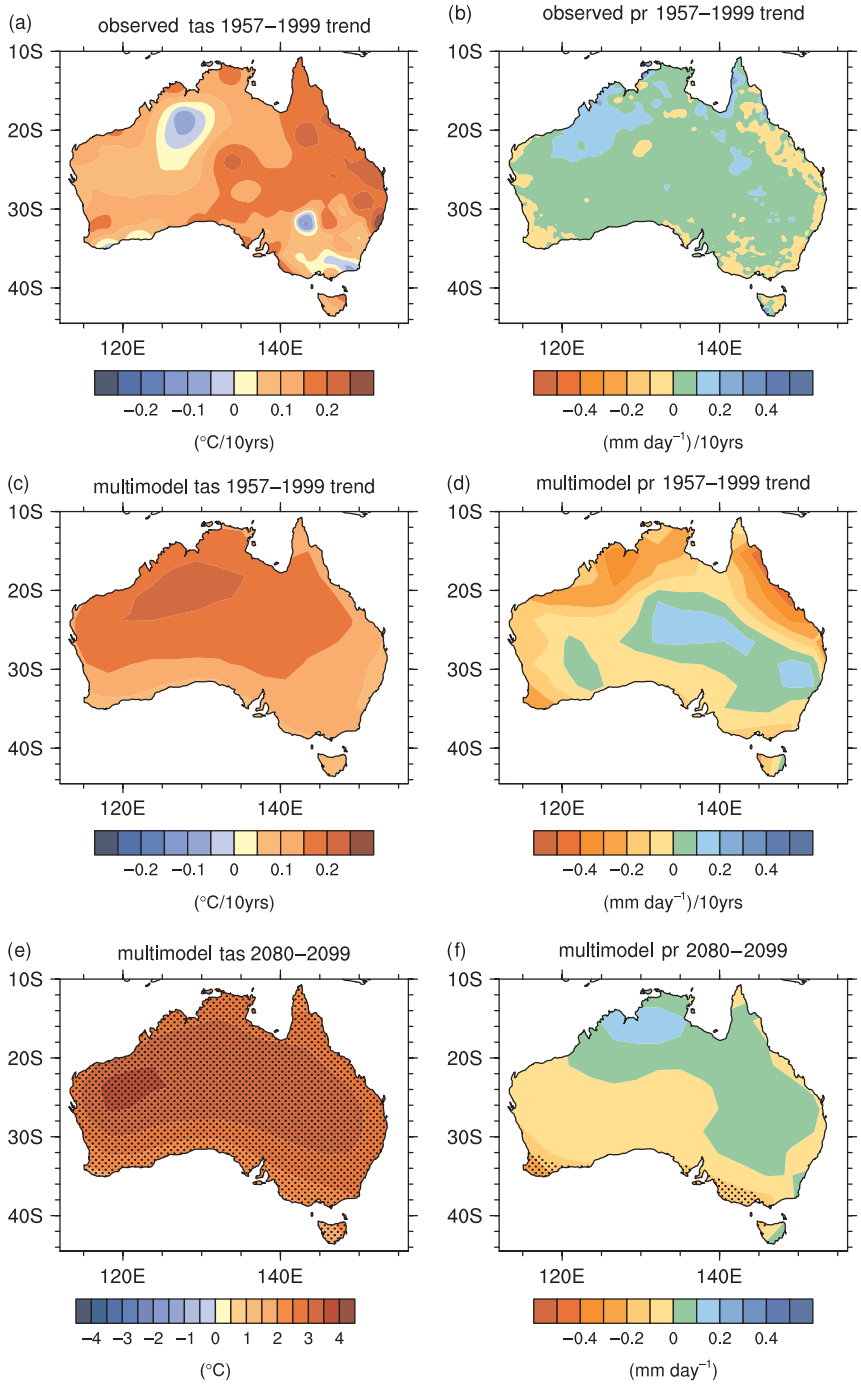

Figure 8. Changes in mean temperature (left column) and precipitation (right column) for observations (a, b), 20th century simulations (c, d) and 21st century SRES A1B simulations (e, f). Twentieth century changes are represented as trends from 1957 to 1999, while future changes are differences of 2080–2099 minus 1980–1999. Stippling in (e and f) indicates regions where the multi-model mean change divided by the intermodel standard deviation of the change is greater than one, a measure of the consistency of the multi-model response. The same nine models for which extremes indices were analysed are used to form the multi-model means here.

DOI: 10.1002/joc

overall changes, with the exception of the northwest. As mentioned earlier, the attribution of observed changes in northwest Australia to anthropogenic sources is still a matter of debate; hence, it is not necessarily clear that we should expect the models to reproduce the observed changes there. Future changes in the annual mean temperature (Figure 8(e)) and precipitation (Figure 8(f)) are represented by the multi-model mean of 2080–2099 minus 1980–1999 for the A1B (mid range) scenario. Warming occurs across the whole continent with largest changes in the interior, while precipitation increases in the north and decreases in the southern regions. Here stippling indicates regions where the multi-model mean divided by the intermodel standard deviation is greater than 1, a measure of the consistency of the response between models. Note that while the temperature change is consistent between the models (as depicted by the stippling of nearly all grid points), precipitation projections are only consistent in the lowest southern latitudes. Thus, there is much intermodel variability in the projected precipitation changes over Australia.

Turning to future projections of the extremes indices, we first show time series of area averages over all Australian grid points (Figure 9), using only grid points for which valid observations are available, as in Figure 1. Multi-model means (solid lines) for the 20C3M and three SRES scenarios, B1 (blue), A1B (green) and A2 (red) are shown for 1870–2099. For the ensemble mean of each of the nine models, anomalies from the entire time series (1870–2099; green line) are first formed for each scenario, followed by the average across models. The multi-model mean is centred around 1980–1999 with a ten-year running mean applied to smooth out the interannual variability, which can be large, particularly for the precipitation extremes. Table IV lists the ratio of changes averaged over Australia for 2080–2099 to changes found globally in Tebaldi *et al.* (2006).

For all temperature-based indices (excepting *extreme temperature range)*, the significant trends observed in the latter half of the 20th century (Table II) are projected to continue into the 21st century. A noticeable increase in *warm nights* is found Australia-wide (Figure 9(a)), with percentage increases between 15 and 40% by the end of the 21st century. This is consistent with a projected rise in *warm nights* globally (Tebaldi *et al.*, 2006), with the Australian-average increase at the end of the century slightly smaller than the global average (Table IV, ratio of 0.86). A consistent increase (decrease) in *heat wave duration* (*frost days)* is found over the 21st century for every scenario, with the changes for these indices markedly smaller than that of the global average. The *heat wave duration* increase over Australia is consistent

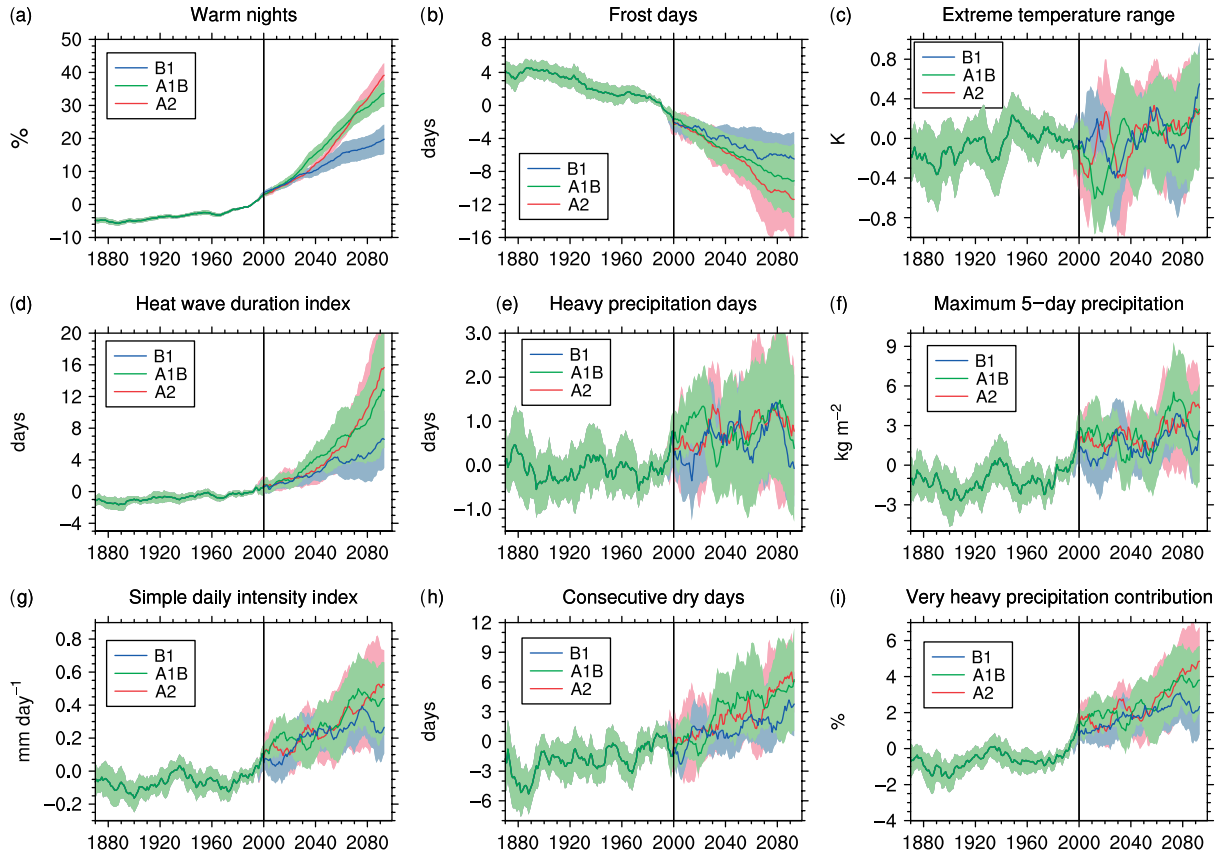

Figure 9. Time series of areally averaged extremes indices (Frich *et al.*, 2002) between 1870 and 2099 using grid boxes in Australia with the observed data from Figure 2 (temperature) and Figure 3 (precipitation). The multi-model ensemble mean (solid lines) of nine models from the CMIP3 dataset is shown for the SRES B1, A1B and A2 scenarios, with shading representing two times intermodel standard deviation. All model time series are smoothed with a ten-year running mean. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

Table IV. Ratios of the 2080–2099 changes in extremes indices in Australia for the SRES A1B scenario relative to (1) changes found globally, (2) the SRES B1 (low-range emissions) and (3) the A2 (high-range emissions). Note that the Australia/Global ratios use only grid points with valid observations over Australia, consistent with the time series in Figure 9, while the B1/A1B and A2/A1B ratios are the weighted averages over all Australian grid points of the patterns of change shown in Figure 10.

Index	Australia/ Global (A1B)	B1/A1B	A2/A1B
Warm nights	0.86	0.65	1.11
Frost days	0.45	0.86	1.15
Extreme temperature range	−0.53	0.58	2.14
Heat wave duration	0.3	0.50	1.40
Heavy precipitation days	0.17	0.79	−0.54
Maximum 5-day precipitation	0.3	0.61	1.49
Simple daily intensity	1.01	0.76	1.09
Consecutive dry days	4.11	0.58	1.19
Very heavy precipitation contribution	0.80	0.52	1.42

with the findings of Tryhorn and Risbey (2006). A notable difference is seen between the observations (Figure 2(e)) and projections in the *extreme temperature range* (Figure 9(c)), with a significant decrease being observed (Table II) compared to an increasing trend in the multi-model projections throughout the 20th and 21st centuries.

Projected changes in the precipitation-based indices are much noisier, with little separation between the scenarios even at the end of the 21st century. Nonetheless, strong increases in *simple daily intensity* (Figure 9(g)), *consecutive dry days* (Figure 9(h)) and *very heavy precipitation contribution* (Figure 9(i)) are projected for Australia over the 21st century, suggesting that our future precipitation regime will have longer dry spells interrupted with heavier precipitation events. Note, however, that other measures of precipitation extremes, *heavy precipitation days* (Figure 9(e)) and *maximum 5-day precipitation* (Figure 9(f)) show no significant trend over Australia. Similar time series over northern and southern Australia (not shown) show no appreciable differences to the Australia-wide averages, except in *extreme temperature range* where a steady increase over the 21st century is found in southern Australia but little change occurs in the north.

Multi-model mean changes in the extremes indices across Australia at the end of the 21st century are seen in Figure 10. This figure is adapted from Tebaldi *et al.* (2006) who used normalized averages for each model to compute the multi-model mean changes. Here, we use non-normalized averages across all model runs, with the expectation that non-normalized units are more meaningful for the user community. The patterns of change are shown for the A1B scenario, the mid range of emissions scenarios, but similar patterns are found for the B1 and A2 scenarios (not shown). Stippling here indicates that at least five out of nine models agree that the change is significant (Tebaldi *et al.*, 2006). In general, warmer and wetter conditions are seen, with significant changes at most grid points for *warm nights* (Figure 10(a)) suggesting a very robust result. Similarly, a large increase in *heat wave duration* (Figure 10(d)) is found across all of Australia, particularly the dry arid regions. Increases in *consecutive dry days* are also seen Australia-wide, although consensus among the models is only found in the interior. Precipitation extremes increase for most indices and locations, with only *heavy precipitation days* having large regions of both positive and negative (albeit small) changes.

Previous studies (e.g. Harvey, 2004) have shown that mean climate change patterns tend to scale with the emissions scenario, i.e. the larger the greenhouse gas forcing, the stronger the response. Tebaldi *et al.* (2006) found this to be true for all temperature indices and a similar result is found here for Australia. A simple quantitative measure of this scaling was computed by dividing the A2 and B1 patterns of change by the A1B pattern shown in Figure 10, thus forming a ratio based on emissions at each grid point. The ratio patterns are then smoothed with a nine-point filter and weighted averages are computed over all Australian grid points. The area average ratios are listed in Table IV. If the patterns were to scale with emissions, we would expect positive numbers at all grid points and smaller values for the B1/A1B ratio than for the A2/A1B ratio. For each index in Table IV, except *heavy precipitation days*, these ratios are indeed positive with values less than one for B1 and greater than one for A2. Note that *extreme temperature range* undergoes more than a doubling from A1B to A2 when averaged across Australia. This is due to the zero line between negative and positive changes (Figure 10(c)) steadily moving southward as the emissions increase. As this pattern is not seen in the observations, it is difficult to draw any conclusions about this result.

5. Discussion

It is encouraging to note that the majority of GCMs analysed in this study were able to generally simulate the sign of the observed trend and, to some extent, the associated variability of temperature and precipitation extremes when averaged over Australia. This gives us some confidence in the projected changes presented here. However, the results also showed that GCMs may not be adequately simulating the spatial trend patterns of extremes across the continent during the late 20th century. Previous studies (e.g. Kiktev *et al.*, 2003; Christidis *et al.*, 2005; Kharin *et al.*, 2007) show that climate models are generally skilful at reproducing global trend patterns of temperature extremes, but have little skill in reproducing trend patterns of precipitation extremes. Hegerl *et al.* (2004) also find that patterns of change in precipitation

DOI: 10.1002/joc

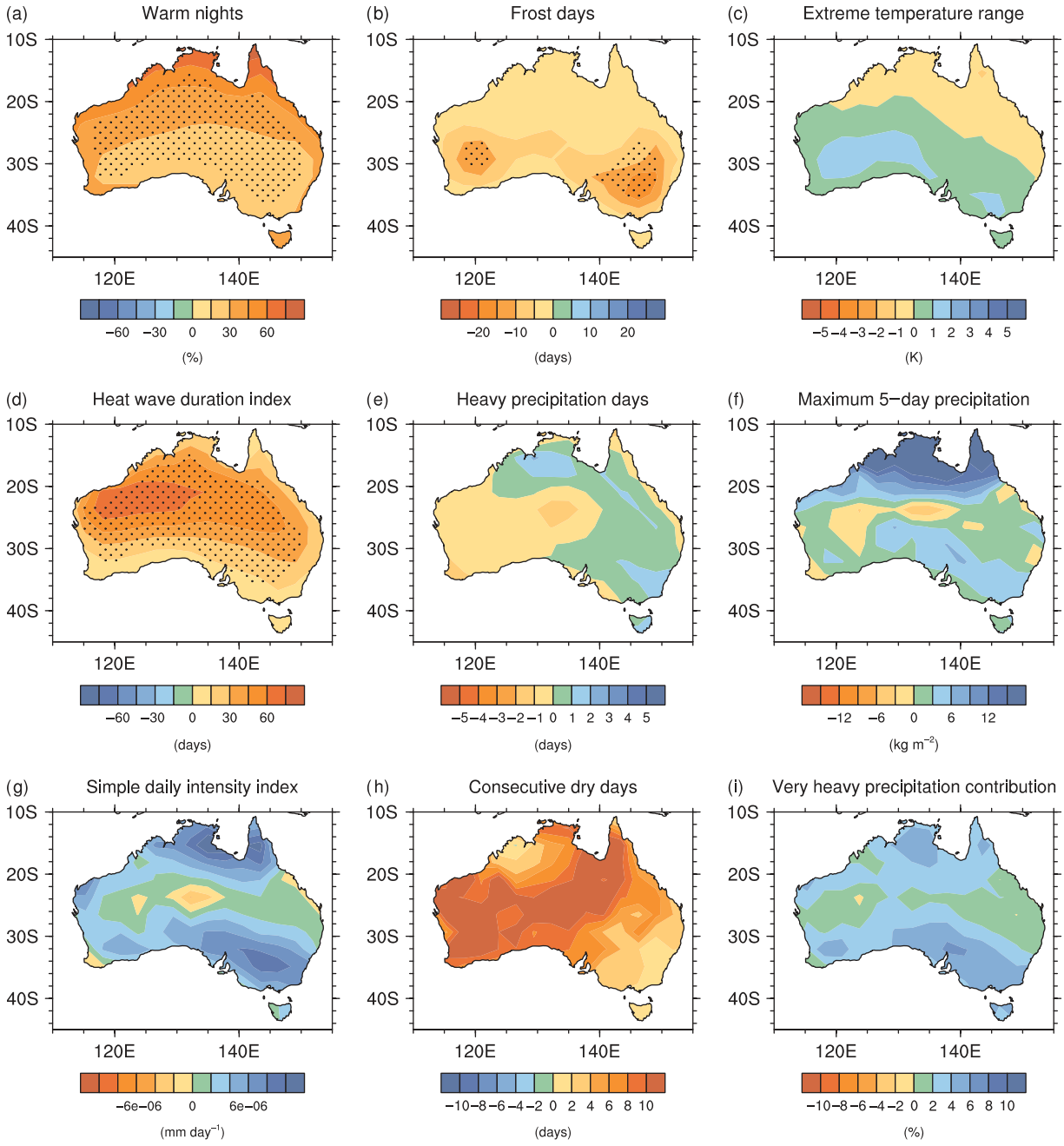

Figure 10. Ensemble mean projected changes (2080–2099 minus 1980–1999) in the extremes indices used in this study (Table I) from the CMIP3 multi-model dataset SRES A1B scenario. Stippling indicates that at least five out of nine models agree that the change is significant (Tebaldi *et al*., 2006).

extremes are more heavily influenced by internal variability of the climate system when compared to temperature extremes. Since the detection of trends in climate variables is a signal-to-noise problem, the noise associated with temperatures at regional scales is greater than at larger continental or global scales (Karoly and Wu, 2005). So, perhaps we cannot expect regional trend patterns to be well simulated, particularly if changes in the extremes in Australia are driven primarily by local influences. Furthermore, the strong influence of ENSO on the Australian climate, which improves predictability on seasonal timescales, also enhances the variability, making attribution on climate scales more difficult.

Some of the driving mechanisms of regional observed trends in Australia are still under debate. While a number of studies have attributed portions of the drying in the southwest to anthropogenic forcing (Cai and Cowan, 2006; Hope, 2006; Timbal *et al*., 2006), the impact of natural variability (Cai *et al*., 2005) and land-cover change (Pitman and Narisma, 2005; Timbal and Arblaster, 2006) also appear to be reasonably large. The increase in precipitation and associated cooling in northwest Australia is well known to researchers (e.g. Nicholls *et al*., 1997; Power *et al*., 1998). Rotstayn *et al*. (2007) suggests that it may be the poor simulation of aerosols in GCMs, which is failing to capture these trends although Shi *et al*. (2008)

DOI: 10.1002/joc

suggest that this may be a model artefact. Wardle and Smith (2004) suggest that the continental warming further south is driving an enhancement of the Australian monsoon. Other possibilities include the known biases of climate models in simulating tropical mean climate and variability including the response of certain cloud feedback to CO_2 that might be causing the sea surface temperature (SST) to warm unevenly (Meehl *et al*., 2000; Barsugli *et al*., 2006). Recent research at the Bureau of Meteorology finds Australian precipitation trends to be consistent with the decadal variability in tropical Pacific SSTs (Harry Hendon, personal communication). Thus, if GCMs could capture the zonal gradient of the SST changes, with a minimum warming in the Central Pacific, they would likely capture the increase in northwest Australian mean precipitation and by extension extremes. Although Santer *et al*. (2006) attribute changes in tropical Atlantic and Pacific SSTs to anthropogenic forcing, the extent to which the pattern of observed trends in tropical SSTs is anthropogenic, is unknown. Further study is required to untangle the contributions of unforced and forced variability to recent changes in the Australian climate.

In this study, model resolution appears not to be critical. While the lowest resolution model, INM-CM3.0, gets the wrong sign of trend of all the precipitation indices and two temperature indices, the highest resolution model, MIROC3 (hires), also does not fare well for four of the five precipitation indices and one of the temperature indices. Perkins *et al*. (2007) find that it is possible to discriminate between models in their ability to simulate daily temperature and precipitation distributions over Australia, but the findings here suggest that no one model is particularly good or bad at reproducing the observed trends or spatial patterns in the extremes of the two variables. In addition, Chen and Knutson (2008) suggest that the way in which the observed precipitation indices are gridded, prohibits a fair comparison between models and observations.

Given that changes in climate extremes will have much larger societal and ecological impacts than mean change (Easterling *et al*., 2000), we need to be confident in our ability to simulate future changes in extremes, if we are to adequately assess their impacts. This study shows that while we can have some confidence, in a general sense, when assessing simulated extremes from coupled climate models over Australia, uncertainties in the patterns of future projections need to be considered when assessing changes on regional scales.

6. Conclusions

In this study, objective measures have been used to analyse the ability of an ensemble of multiple GCMs to simulate observed trends in the climate extremes over Australia and to assess projected changes in these extremes at the end of the 21st century. In general, the models capture the sign of observed trends in both temperature and precipitation extremes, but no one model is consistently good at reproducing all indices. In spite of some differences in definition, the amount, interannual variability and trend of the *warm nights* index is particularly well represented by all the models analysed. A pattern correlation technique, however, has shown that none of the models are skilful at simulating important trend patterns on regional scales, only the multi-model ensemble showing any significant skill at modelling trend patterns of *heavy precipitation days*. This may imply that some regional and/or large-scale process or processes over Australia are not well modelled or resolved, or that unforced variability is contributing largely to these changes.

Future projections show that the significant trends in temperature extremes that have been observed during the latter half of the 20th century are set to continue into the 21st century. A substantial increase in w*arm nights* and *heat wave duration* and decrease in *frost days* are projected by the end of this century under the SRES scenarios. The precipitation indices *simple daily intensity, consecutive dry days* and *very heavy precipitation contribution* are also set to more than double within the next 100 years. In general, the magnitude of changes in both temperature and precipitation indices were found to scale with the strength of emissions.

More work would be required to determine to what degree recent changes in the climate extremes over Australia are due to human causes and more international cooperation is essential to ensure that the modelling and observational groups derive consistent extremes indices. However, it has been important to document the ability of the models to reproduce 20th century changes in context with projections. As the driest inhabited continent with marginal agricultural climate and unique and vulnerable societies and ecosystems, stakeholders and policymakers in Australia urgently need information regarding climate extremes. Here, we highlight both the agreement and uncertainty around the model projections.

Acknowledgements

We acknowledge the modelling groups for making their simulations available for analysis, the PCMDI for collecting and archiving the CMIP3 model output, and the WCRP's Working Group on Coupled Modelling (WGCM) for organizing the model data analysis activity. The WCRP CMIP3 multi-model dataset is supported by the Office of Science, US Department of Energy. Portions of this study were supported by the Office of Science (BER), US Department of Energy, Cooperative Agreement No. DE-FC02-97ER62402, and the National Science Foundation. The National Center for Atmospheric Science is sponsored by the National Science Foundation. Support for this study was also obtained from the Australian Climate Change Science Program (ACCSP) Many thanks also to David Karoly, Amanda Lynch, Jerry Meehl, Neville Nicholls and Claudia Tebaldi for their

DOI: 10.1002/joc

useful comments and help on improving an earlier version of this manuscript and Rob Smalley for his help with the trend calculation software.

References

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.

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahim F, Tagipour A, Kumar Kolli R, Revadekar JV, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez Aguirre JL. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research-Atmospheres* **111**: D05109, DOI:10.1029/2005JD006290.

Barsugli JJ, Shin SI, Sardeshmukh PD. 2006. Sensitivity of global warming to the pattern of tropical ocean warming. *Climate Dynamics* **27**: 483–492.

Cai W, Cowan T. 2006. SAM and regional rainfall in IPCC AR4 models: can anthropogenic forcing account for southwest Western Australian winter rainfall reduction? *Geophysical Research Letters* **33**: L24708, DOI:10.1029/2006GL028037.

Cai W, Shi G, Li Y. 2005. Multidecadal fluctuations of winter rainfall over southwest Western Australia simulated in the CSIRO Mark 3 coupled model. *Geophysical Research Letters* **32**: L12701, DOI:10.1029/2005GL022712.

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

Christidis N, Stott PA, Brown S, Hegerl GC, Caesar J. 2005. Detection of changes in temperature extremes during the second half of the 20th century. *Geophysical Research Letters* **32**: L20716, DOI:10.1029/2005GL023885.

Collins DA, Della-Marta PM, Plummer N, Trewin BC. 2000. Trends in annual frequencies of extreme temperature events in Australia. *Australian Meteorological Magazine* **49**: 277–292.

CSIRO (Commonwealth Scientific and Industrial Research Organisation). 2001. *Climate Projections for Australia*. CSIRO Atmospheric Research: Melbourne; 8, http://www.dar.csiro.au/publications/projections2001.pdf.

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO. 2000. Climate extremes: Observations, modeling, and impacts. *Science* **289**: 2068–2074.

Fitzharris B, Hennessy K, Bates B, Hughes L, Howden M, Salinger J, Harvey N, Warrick R. 2007. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Working Group II contribution to the Intergovernmental Panel on Climate Change Fourth Assessment Report*. Cambridge University Press: Cambridge, New York; 1032.

Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research* **19**: 193–212.

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

Griffiths GM, Chambers LE, Haylock MR, Manton MJ, Nicholls N, Baek H-J, Choi Y, Della-Marta PM, Gosai A, Iga N, Lata R, Laurent V, Maitrepierre L, Nakamigawa H, Ouprasitwong N, Solofa D, Tahani L, Thuy DT, Tibig L, Trewin B, Vediapan K, Zhai P. 2005. Change in mean temperature as a predictor of extreme temperature change in the Asia-Pacific region. *International Journal of Climatology* **25**: 1301–1330.

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.

Harvey LDD. 2004. Characterizing the annual-mean climatic effect of anthropogenic CO2 and aerosol emissions in eight coupled atmosphere ocean GCMs. *Climate Dynamics* **23**: 569–599.

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.

Hegerl GC, Zwiers FW, Stott PA, Kharin VV. 2004. Detectability of anthropogenic changes in annual temperature and precipitation extremes. *Journal of Climate* **17**: 3683–3700.

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

Hope PK. 2006. Projected future changes in synoptic systems influencing southwest Western Australia. *Climate Dynamics* **26**: 751–764.

IOCI. 2002. Climate variability and change in south west Western Australia. *Indian Ocean Climate Initiative Panel*, Perth; 34.

Karoly DJ, Braganza K. 2005. Attribution of recent temperature changes in the Australian region. *Journal of Climate* **18**: 457–464.

Karoly DJ, Wu QG. 2005. Detection of regional surface temperature trends. *Journal of Climate* **18**: 4337–4343.

Kendall MG. 1975. *Rank Correlation Methods*. Griffin: London.

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.

Kiktev D, Caesar J, Alexander LV, Shiogama H, Collier M. 2007. Comparison of observed and multimodeled trends in annual extremes of temperature and precipitation. *Geophysical Research Letters* **34**: L10702, DOI:10.1029/2007GL029539.

Kiktev D, Sexton DMH, Alexander L, Folland CK. 2003. Comparison of modelled and observed trends in indices of daily climate extremes. *Journal of Climate* **16**: 3560–3571.

Lambert FH, Gillett NP, Stone DA, Huntingford C. 2005. Attribution studies of observed land precipitation changes with nine coupled models. *Geophysical Research Letters* **32**: L18704, DOI:10.1029/2005GL023654.

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

Mann HB. 1945. Nonparametric trends against test. *Econometrica* **13**: 245–259.

Manton MJ, Della-Marta PM, Haylock MR, Hennessy KJ, Nicholls N, Chambers LE, Collins DA, Daw G, Finet A, Gunawan D, Inape K, Isobe H, Kestin TS, Lefale P, Leyu CH, Lwin T, Maitrepierre L, Ouprasitwong N, Page CM, Pahalad J, Plummer N, Salinger MJ, Suppiah R, Tran VL, Trewin B, Tibig I, Yee D. 2001. Trends in extreme daily rainfall and temperature in southeast Asia and the South Pacific: 1916–1998. *International Journal of Climatology* **21**: 269–284.

Meehl GA, Arblaster JM, Tebaldi C. 2005. Understanding future patterns of precipitation extremes in climate model simulations. *Geophysical Research Letters* **32**: L18719, DOI: 10.1029/2005GL023680.

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 and 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., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Avery, M. Tignor and H.L. Miller (eds). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Meehl GA, Arblaster JM, Tebaldi C. 2007a. Contributions of natural and anthropogenic forcing to changes in temperature extremes over the U.S. *Geophysical Research Letters* **34**: L19709, DOI:10.1029/2007GL030948.

Meehl GA, Washington WM, Santer BD, Collins WD, Arblaster JM, Hu A, Lawrence DM, Teng H, Buja LE and Strand WG Jr. 2006.. Climate change projections for the 21st century and climate change commitment in the CCSM3. *J. Climate* **19**: 2597–2616.

Meehl GA, Collins WD, Boville BA, Kiehl JT, Wigley TML, Arblaster JM. 2000. Response of the NCAR climate system model to increased CO2 and the role of physical processes. *Journal of Climate* **13**: 1879–1898.

Meehl GA, Tebaldi C. 2004. More intense, more frequent and longer lasting heat waves in the 21st century. *Science* **305**: 994–997.

Meehl GA, Tebaldi C, Nychka D. 2004. Changes in frost days in simulations of 21st century climate. *Climate Dynamics* **23**: 495–511.

Nicholls N, Drosdowsky W, Lavery B. 1997. Australian rainfall variability and change. *Weather* **52**: 66–72.

Perkins SE, Pitman AJ, Holbrook NJ, McAneney J. 2007. Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature and precipitation over Australia using probability density functions. *Journal of Climate* **20**: 4356–4376.

Pitman AJ, Narisma GT. 2005. The role of land surface processes in regional climate change: a case study of future land cover change over south western Australia. *Meteorology and Atmospheric Physics* **89**: 235–249.

Plummer N, Salinger MJ, Nicholls N, Suppiah R, Hennessy KJ, Leighton RM, Trewin B, Page CM, Lough JM. 1999. Changes in climate extremes over the Australian region and New Zealand during the twentieth century. *Climatic Change* **42**: 183–202.

Power SF, Tseitkin F, Torok S, Lavery B, Dahni R, McAvaney B. 1998. Australian temperature, Australian rainfall and the Southern Oscillation, 1910–1992: coherent variability and recent changes. *Australian Meteorological Magazine* **47**: 85–101.

Rotstayn LD, Cai W, Dix MR, Farquar GD, Feng Y, Ginoux P, Herzog M, Ito A, Penner JE, Roderick ML, Wang M. 2007. Have Australian rainfall and cloudiness increased due to the remote effects of the Asian anthropogenic aerosols? *Journal of Geophysical Research-Atmospheres* **112**: D09202, DOI:10.1029/2006JD007712.

Santer BD, Wigley TML, Gleckler PJ, Bonfils C, Wehner MF, AchutaRao K, Barnett TP, Boyle JS, Brüggemann W, Fiorino M, Gillett N, Hansen JE, Jones PD, Klein SA, Meehl GA, Raper SCB, Reynolds RW, Taylor KE, Washington WM. 2006. Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis regions. *Proceedings of the National Academy of Sciences* **103**: 13905–13910, DOI:10.1073/pnas.0602861103.

Scaife AA, Folland CK, Alexander LV, Moberg A, Knight JR. 2008. European climate extremes and the North Atlantic Oscillation. *Journal of Climate* **21**: 72–83.

Shi G, Cai W, Cowan T, Ribbe J, Rotstayn L, Dix M. 2008. Variability and trend of the northwest Western Australia Rainfall: observations and coupled climate modelling. *Journal of Climate* (in press).

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

Smith I. 2004. An assessment of recent trends in Australian rainfall. *Australian Meteorological Magazine* **53**: 163–173.

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA. 2006. Going to the extremes: an intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change* **79**: 185–211.

Timbal B, Arblaster JM. 2006. Land cover change as an additional forcing to explain the rainfall decline in the south west of Australia. *Geophysical Research Letters* **33**: L07717, DOI:10.1029/2005GL025361.

Timbal B, Arblaster JM, Power SB. 2006. Attribution of the late 20th century rainfall decline in South-West Australia. *Journal of Climate* **19**: 2046–2065.

Trenberth KE, Jones PD, Ambenje P, Bojariu R, Easterling D, Klein Tank A, Parker D, Rahimzadeh F, Renwick JA, Rusticucci M, Soden B, Zhai P. 2007. Observations: surface and atmospheric climate change. 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, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, New York.

Trewin BC. 2001. The development of a high-quality daily temperature data set for Australia. In *11th Symposium on Meteorological Observations and Instrumentation*, Albuquerque, 14–18 January 2001, 279–284.

Tryhorn L, Risbey J. 2006. On the distribution of heat waves over the Australian region. *Australian Meteorological Magazine* **55**: 169–182.

Wang XL, Swail VR. 2001. Changes in extreme wave heights in Northern Hemisphere oceans and related atmospheric circulation regimes. *Journal of Climate* **14**: 2204–2220.

Wardle R, Smith I. 2004. Modelled response of the Australian monsoon to changes in land surface temperatures. *Geophysical Research Letters* **31**: L16205, DOI:10.1029/2004GL020157.

Wilks DS. 1997. Resampling hypothesis tests for autocorrelated fields. *Journal of Climate* **10**: 65–82.

Zhang XB, Hegerl G, Zwiers FW, Kenyon J. 2005. Avoiding inhomogeneity in percentile-based indices of temperature extremes. *Journal of Climate* **18**: 1641–1651.