Understanding the role of sea surface temperature-forcing for variability in global temperature and precipitation extremes


It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.wace.2018.06.002

Publisher: Elsevier

www.reading.ac.uk/centaur
CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
Understanding the role of sea surface temperature-forcing for variability in global temperature and precipitation extremes

Andrea J. Dittus\(^a,b,c,*\), David J. Karoly\(^a,b\), Markus G. Donat\(^b,d\), Sophie C. Lewis\(^b,e\), Lisa V. Alexander\(^b,d\)

\(a\) School of Earth Sciences, The University of Melbourne, Parkville, Victoria, Australia  
\(^b\) Australian Research Council Centre of Excellence for Climate System Science, Australia  
\(^c\) School of Earth, Atmosphere and Environment, Monash University, Clayton, Victoria, Australia  
\(^d\) Climate Change Research Centre, UNSW Sydney, New South Wales, Australia  
\(^e\) School of Physical Environmental and Mathematical Sciences, University of New South Wales, Canberra, ACT, Australia

**ABSTRACT**

The oceans are a well-known source of natural variability in the climate system, although their ability to account for inter-annual variations of temperature and precipitation extremes over land remains unclear. In this study, the role of sea-surface temperature (SST)-forcing is investigated for variability and trends in a range of commonly used temperature and precipitation extremes over the period 1959 to 2013. Using atmospheric simulations forced by observed SST and sea-ice concentrations (SIC) from three models participating in the Climate of the Twentieth Century Plus (C20C+) Project, results show that oceanic boundary conditions drive a substantial fraction of inter-annual variability in global average temperature extreme indices, as well as, to a lower extent, for precipitation extremes. The observed trends in temperature extremes are generally well captured by the SST-forced simulations although some regional features such as the lack of warming in daytime warm temperature extremes over South America are not reproduced in the model simulations. Furthermore, the models simulate too strong increases in warm day frequency compared to observations over North America. For extreme precipitation trends, the accuracy of the simulated trend pattern is regionally variable, and a thorough assessment is difficult due to the lack of locally significant trends in the observations. This study shows that prescribing SST and SIC holds potential predictability for extremes in some (mainly tropical) regions at the inter-annual time-scale.

1. Introduction

It is well known that the oceans are major drivers of internal climate variability, affecting climate around the globe (e.g. Shukla, 1998). Many modes of climatic variability are coupled ocean-atmosphere phenomena, such as the El Niño-Southern Oscillation (ENSO), which is the most important mode of variability for global climate on inter-annual timescales (e.g. McPhaden et al., 2006). The occurrence of an El Niño episode affects many regions worldwide through atmospheric teleconnections and increases global average temperatures in that year (e.g. Trenberth et al., 2002). ENSO also affects temperature and precipitation extremes in many regions across the world (Kenyon and Hegerl, 2008, 2016; Alexander et al., 2009; Arblaster and Alexander, 2012). Local SSTs can also be important for regional climate, for example, a recent study showed that SSTs to the northeast of Australia could enhance or dampen the ENSO-rainfall response in south-eastern Australia (van Rensch et al., 2015).

On a global scale, an example where natural variability in the Pacific Ocean is thought to have played an important role for global climate is through the so-called “hiatus” or slowdown in the rate of increase in global mean surface temperature (GMST). Studies have shown that the hiatus may have been driven by low-frequency internal variability in the Pacific Ocean (Kosaka and Xie, 2013), associated with the negative phase of the Interdecadal Pacific Oscillation (Meehl et al., 2014; Watanabe et al., 2014) or prolonged La Niña-like conditions (Kosaka and Xie, 2013). This phase of the IPO is characterised by stronger trade winds (e.g. England et al., 2014) and cooler than normal sea-surface temperatures in the eastern Pacific, which have increased heat uptake by the oceans and thus contributed to the hiatus (Meehl et al., 2011).

While it is clear that natural variability in the oceans greatly influences global climate in different ways, to date, few studies have systematically investigated the role of SST variability for temperature and precipitation extremes on land and their representation in global atmospheric climate models. Donat et al. (2016) compared century-long SST-driven runs from one atmospheric GCM with gridded observations and reanalysis data. They found that SST variability could explain about 50% of the inter-annual variability of globally averaged temperature extremes and about 15% for precipitation extremes. Regionally, significant correlations between SST-driven runs and...
observations were mainly found in the Tropics. However, understanding and estimating internal variability is equally important to understanding the role of anthropogenic forcing, as it is a combination of both that drives climatic events in the real world. Climate variability can act to temporarily enhance (e.g. heatwaves in Australia during El Niño) or dampen (e.g. GMTG hiatus) the effects of a warming climate. In this study, the potential of atmospheric simulations forced by observed SST and SIC conditions for understanding variability in temperature and precipitation extremes is investigated. Unlike in coupled ocean-atmosphere general circulation models (CGCMs), these atmospheric simulations will, by definition, have coupled ocean-atmosphere modes of variability in phase with the “real world” climate. Hence, there is potential for atmospheric simulations to be able to reproduce the timing of specific features of natural variability that coupled models may not.

In this study, a multi-model ensemble of atmospheric simulations coordinated through the CLIVAR Climate of Twentieth Century Plus Project (Folland et al., 2014) and an observational dataset of extreme indices are used to assess the effect of SST-forcing in a suite of widely used indices of temperature and precipitation extremes. First, the effect of SST-forcing for inter-annual variability in the global land average of the temperature and precipitation extremes is assessed. The contribution of SST-forcing to inter-annual variability is then quantified at the grid box level. Next, the relationship with ENSO is investigated in observations and models. Finally, long-term trends in observations and climate models are compared, to identify if prescribing SSTs and SIC improves the agreement between observed and modelled trends compared to coupled models. The paper is structured as follows. Section 2 provides an overview of the extremes indices, data and methods used in this study. Results are shown in section 3, and potential applications and limitations of this study are discussed in section 4. The conclusions of this paper are summarised in Section 5.

2. Data and methods

2.1. Extreme indices

The extreme indices used in this paper are a subset of the 27 indices recommended by the Expert Team of Climate Change Detection and Indices (e.g. Zhang et al., 2011). Here, we exclude absolute threshold indices (such as e.g. the number of frost days or summer days) and duration indices, as we focus on indices that are meaningful across the globe. The 11 remaining indices are either relative threshold indices (i.e. percentile-based indices) or indices of annual maxima or minima. Additionally, we have added a normalized heavy precipitation index (R95pTOT) where the amount of heavy precipitation is normalized by the total precipitation at each grid box. This ensures that values from all grid boxes are comparable, rather than being dominated by regions of high precipitation. This index is used in Fig. 1 (global averages), while the subsequent figures are all using the heavy precipitation amount (R95p). These indices are summarised in Table 1. Note that some of these indices measure relatively “moderate” extremes, as they occur several times per year by definition in the case of the percentile indices.

2.2. Data

The observational data in this study were obtained from the GHCNDEX dataset (downloaded in February 2015; Donat et al., 2013). GHCNDEX was chosen for its length of record compared to other similar datasets, to better match the time period available in model simulations. Three global climate models contributing to the C20C+ project with various ensemble sizes are analysed in this study. The CAM5-1-1degree model (Neale et al. 2012) provides the largest ensemble with 50 simulations available from 1959 while for MIROC5 (Watanabe et al. 2010) and ACCESS1.3 (Bi et al., 2013), 10 and 5 runs were available respectively. These simulations are available for the period from 1959, 1950 and 1955 to 2013 or 2014 respectively, and cover a common period from 1959 to 2013. The ACCESS1.3 and CAM5-1-1degree simulations were forced with observed SSTs and sea-ice (SICs) from the Hurrell et al. (2008) dataset. MIROC5 employed the Hadley Centre Sea Ice and Sea Surface Temperature data set (Rayner et al., 2003) as their forcing dataset. These historical simulations include natural and anthropogenic forcings: solar and volcanic forcings (natural), greenhouse gas forcings (including CO2, N2O, CH4 and major CFCs), ozone, sulphate aerosols, black and organic carbon. An exception is the CAM5-1-1degree ensemble, which does not include volcanic forcings. Note that these models were the only ones available in the C20C+ archive at the time of analysis (December 2016), but more have been added since.

3. Methods

The required ETCCDI indices were calculated using the climdex.p-pcic.ncdf R package provided by the Pacific Climate Impacts Consortium, Canada (https://github.com/pacificclimate/climdex.p-pcic.ncdf). This is the same software that was used to calculate the ETCCDI indices for CMIP5 models produced by the Canadian Community Climate Modelling Centre used in other studies (Stilman et al., 2013), and required daily maximum and minimum temperature and precipitation data. This ensures that there are no methodological differences between the indices calculated here and the CMIP5 indices used in other studies. All indices were calculated on the models’ native grid, and subsequently re-gridded to a common grid for the purposes of this study as described further down.

In a large enough ensemble, the ensemble mean represents the forced response to external forcings and prescribed boundary conditions such as greenhouse gases, aerosol, volcanic and solar forcings as well as SSTs and SIC in the model configuration of the C20C+ ensemble. The differences of the individual ensemble members from the ensemble mean therefore correspond to unforced atmospheric variability, since the forced response is accounted for in the ensemble mean. However, in any finite ensemble, some residual unforced variability may remain in the ensemble mean.

As we are primarily interested in the forced component of interannual variability rather than long-term forcing, we first remove the linear trend from the time series for all analysis with the exception of section 3.3 on trends and where stated otherwise (e.g. Fig. 1). Removing this long-term trend also likely removes some long-term effects of solar and volcanic forcings and some decadal variability. The interannual variability in the de-trended time series is then primarily due to volcanic forcing and SST variability, or stochastic atmospheric variability. Major volcanic eruptions in the second half of the Twentieth Century include Mt Agung in 1963, El Chichón in 1982, Pinatubo in 1991. Unless stated otherwise, the term “forced response” in this study refers to the forced component of inter-annual variability and not long-term forcing. The “common variance” across ensemble members is measured as the variance in the de-trended ensemble mean divided by the variance of a single de-trended ensemble member. Specifically, at each gridbox the ensemble member with the median variance in the ensemble is used, although the choice of the ensemble member has little overall effect on the results. The “fraction of common variance” can be interpreted as the fraction of variability forced by oceanic boundary conditions relative to stochastic atmospheric variability.

The model-simulated temperature and precipitation extremes were re-gridded to the resolution of the GHCNDEX observational dataset (2.5°×2.5°) for all analyses using a first order conservative remapping scheme and masked to the available observational coverage when calculating global averages. Only grid boxes where GHCNDEX contained valid data for at least 80% of all years were considered to minimise possible bias from too incomplete time series. Note the same masking
Fig. 1. Global land average time series of each of the 11 annual indices considered here, in order from left to right and top-to-bottom: cool nights, warm nights, cool days, warm days, hottest day, coldest day, coldest night, hottest night, contribution of very wet days to total precipitation, max. 5-day precipitation amount, max. 1-day precipitation amount. Observations are shown in red, each model ensemble is shown in a different colour. The correlation coefficient between the de-trended observations and individual ensemble members are given in the top right-hand corner. The correlation coefficient is first calculated for each ensemble member then averaged across all ensemble members of a single model, then across all models. The second number (“RensM”) corresponds to the correlation coefficient calculated between observations and each model’s ensemble mean, then averaged across all three models. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
and completeness criteria apply for the pattern correlations calculated in Figs. 3 and 4. Trends were calculated using a modified least squares method as outlined in the Intergovernmental Panel on Climate Change Fifth Assessment Report Supplementary Material (Hartmann et al., 2013), taking into account lag-1 autocorrelation in the residuals. Statistical significance was assessed at the 5% level throughout this study, and adjusted for multiple hypothesis testing issues using the “false discovery rate” following Wilks (2016) and using a control level of 0.1 (i.e. double the statistical significance level used), as recommended by Wilks (2016) for spatially correlated data. For climate model data, black stippling indicates areas where at least 80% of realisations agree on sign and at least 50% are statistically significant using the criteria described above, while gray stippling indicates where observations lie within the model range. Finally, please note that the Niño3.4 index was calculated from the Hurrell et al. (2008) SST dataset. All analysis was performed for calendar years from January to December, due to some precipitation indices currently only being available for the calendar year.

4. Results

The results are divided into three subsections. First, the ability of the atmospheric model simulations to reproduce observed year-to-year variability in the global land average of temperature and precipitation extremes is assessed. Then, the relationship of the extreme indices with atmospheric model simulations to reproduce observed year-to-year variability in the global land average of temperature and precipitation extremes is assessed. Then, the relationship of the extreme indices with atmospheric variability is averaged out in the ensemble mean. It is therefore not surprising that correlation coefficients are higher using the ensemble mean than when first calculated on each model realisation prior to averaging.

The fraction of “common variance” across ensemble members represents the fraction of the total variance (as represented by the variance of a single ensemble member) that is common to all ensemble members (i.e. variance of the ensemble mean). The “common variance” thus represents the fraction of variability that is found in all ensemble members and is likely forced by common boundary conditions, primarily SSTs and SIC but also including volcanic aerosols (except in the case of CAM, which does not include volcanoes) and solar forcing. The remainder represents unforced atmospheric variability that may be different in each ensemble member but averaged out in the ensemble mean. The fraction of common inter-annual variance (after de-trending) to all ensemble members is shown for each index for the CAM5-1-degree model in Fig. 2. This model was chosen because it has the largest number of ensemble members available and results are thus more robust than those based on smaller ensembles. The results are very similar for the other models, with slightly larger fractions occurring outside the tropics, likely due to the smaller number of ensemble members used to produce the ensemble mean. The variance common to all ensemble members is largest for the percentile temperature indices, and is highest in the tropics where the fraction of SST-forced variance can reach up to 50 percent. SST forcing is also important for variability in the absolute temperature indices, although the common variance is lower for these extremes. There is some influence of SST-forcing on very wet days, but little to no influence on the annual wettest one and five-day events. When analysed at a grid box level, the highest values of “common variance” are found in the tropics for all indices (Fig. 2). For the precipitation indices in particular, the common variance outside of

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Definition</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN10p</td>
<td>Cool Nights</td>
<td>Annual percentage of days when daily min temperature &lt; 10th percentile</td>
<td>%</td>
</tr>
<tr>
<td>TN20p</td>
<td>Warm Nights</td>
<td>Annual percentage of days when daily min temperature &gt; 10th percentile</td>
<td>%</td>
</tr>
<tr>
<td>TX10p</td>
<td>Cool Days</td>
<td>Annual percentage of days when daily max temperature &lt; 10th percentile</td>
<td>%</td>
</tr>
<tr>
<td>TX20p</td>
<td>Warm Days</td>
<td>Annual percentage of days when daily max temperature &gt; 10th percentile</td>
<td>%</td>
</tr>
<tr>
<td>TXx</td>
<td>Max Tmax</td>
<td>Annual maximum value of daily max temperature</td>
<td>°C</td>
</tr>
<tr>
<td>TXn</td>
<td>Max Tmin</td>
<td>Annual maximum value of daily min temperature</td>
<td>°C</td>
</tr>
<tr>
<td>TXn</td>
<td>Min Tmax</td>
<td>Annual minimum value of daily max temperature</td>
<td>°C</td>
</tr>
<tr>
<td>TNn</td>
<td>Min Tmin</td>
<td>Annual minimum value of daily min temperature</td>
<td>°C</td>
</tr>
<tr>
<td>R95p</td>
<td>Very wet days</td>
<td>Annual total precipitation from days &gt; 95th percentile</td>
<td>mm</td>
</tr>
<tr>
<td>R95pTOT</td>
<td>Contribution of very wet days to total precipitation</td>
<td>Fraction of annual total precipitation from days &gt; 95th percentile</td>
<td>mm</td>
</tr>
<tr>
<td>Rx1day</td>
<td>Max 1-day precipitation amount</td>
<td>Annual maximum 1-day precipitation</td>
<td>mm</td>
</tr>
<tr>
<td>Rx5day</td>
<td>Max 5-day precipitation amount</td>
<td>Annual maximum 5-day precipitation</td>
<td>mm</td>
</tr>
</tbody>
</table>
Fig. 2. The fraction of common variance across 50 CAMS-1-1 degree simulations is shown at the grid box level for all indices considered in this study. The fraction of common variance is calculated as the de-trended variance of the ensemble mean divided by the de-trended variance of a single ensemble member. At each grid box, the variance of the ensemble member with the median variance was used.
the tropics drops off rapidly, which is in line with the earlier observation that the correlations between model-simulated heavy precipitation and observations are greater for the R95p index, which is dominated by high-rainfall areas, than R95pTOT, where all grid boxes contribute equally. Note that the assumption that other forcings are negligible is justified, since the “common variance” between ensemble members of a coupled model is close to zero everywhere (not shown).

4.2. Relationship with ENSO

Next, the effects of ENSO on the variability in the extreme indices and their representation in the C20C+ ensemble are investigated.
ENSO being the most important coupled atmosphere-ocean mode of variability on inter-annual timescales, it is likely that the year-to-year variability of temperature and precipitation extremes in some regions is influenced by ENSO, as has previously been shown in observational and modelling studies (e.g. Kenyon and Hegerl, 2008, 2010; Alexander et al., 2009; Arblaster and Alexander, 2012).

The correlation coefficients between the Niño3.4 index and the annual number of warm days as well as precipitation on very wet days respectively are shown in Fig. 3 for the average of the correlations calculated for each ensemble member individually. We only show one temperature index and one precipitation index here for better readability, additional indices can be found in the Supplementary Material (S1-S9).

The simulations reproduce the positive correlation between warm...
temperature extremes and Niño3.4 in north-western North America, and the corresponding negative correlation coefficients in south-eastern North America (Fig. 3). Black stippling over these areas in the models indicates that this is a robust model response across different ensemble members. They further also simulate the correct sign over Australia and south-east Asia, which indicate an increase in warm days during El Niño conditions, although the correlations are too weak in the models over Australia.

Furthermore, the models all show a strong and coherent positive correlation over northern South America including the Amazon region, however observations are not available over this region. The CAM5 model produces robust positive correlations between warm days and Niño3.4 over tropical latitudes of Africa, that are much weaker in the other models. Again, observations are not available over this region. A strong negative correlation in the Himalayas is found in the ACCESS model (and to a lesser extent in MIROC5) that is not seen in CAM5. Despite these differences, overall the model simulations reproduce the observed correlation pattern over the regions with observational coverage remarkably well, with pattern correlations between 0.51 and 0.59 depending on the model. Furthermore, the gray stippling indicates that where observations are available, the models are largely consistent with observations. The lesser consistency for MIROC5 and ACCESS1.3 compared to CAM5 is likely due to the reduced ensemble size. Note that these pattern correlations are calculated for each individual ensemble member then averaged, and are therefore smaller than if they were calculated on the average correlations directly (not shown).

For precipitation extremes, the pattern correlation of the ENSO-correlation fields is lower, ranging from 0.21 to 0.30 depending on the model. The models reproduce the correlation sign over Australia, the eastern coast of South America and Western Europe, albeit most of these local correlations are not statistically significant in either the observations or the climate models. All models show a negative correlation in the Amazonas region (albeit with different magnitudes and levels of statistical significance) and similar patterns of correlations over Africa, however observational data are not available in these regions. It is worth noting that the spatial pattern is much more heterogeneous than for temperature extremes in both observations and models, and that most regions do not exhibit statistically significant correlations with Niño3.4. It is therefore not surprising that the pattern agreement between observed and modelled correlation coefficients is substantially lower than for temperature extremes. The gray stippling indicates that the simulated correlation coefficients are largely consistent with observations where they are available, confirming that the lack of pattern agreement is likely due to little impact of ENSO on annual rainfall from events exceeding the 95th percentile in many regions of the world, rather than poor model performance.

4.3. Trend patterns of temperature and precipitation extremes

Finally, this section examines whether regional trends in temperature and precipitation extremes as simulated in the C20C+ ensemble can be better captured than in coupled climate model simulations. Again, only two temperature and precipitation indices are shown here, additional indices can be found in the Supplementary Material, Figures S10-18. The average trend across all ensemble members is shown here.

The simulated trend patterns are broadly similar to the observed trends (pattern correlations ≥ 0.78), with widespread increases in the frequency of warm days in both observations and models. However, there are also some marked differences, for example, the observations show an area of decreases in the frequency of warm days in parts of southern South America, which is not reproduced in any of the models (Fig. 4). This decrease in observed warm day frequency lies outside of the model range, indicating that the models are either biased or that the ensemble is not large enough to reflect the "true" unforced variability. Additionally, the simulated increase in warm day frequency over North America is very robust across the models but inconsistent with observations over a large area of the United States. In this area, observations show a weaker increase that is not statistically significant. Interestingly, all realisations (not shown) of all models show small areas of decreases in warm days in different locations of Asia that are not observed, albeit in different locations, which is likely associated with the response to aerosol forcing in this region in different models (see Fig. 4 for the model average).

It is difficult to compare trends in the simulated precipitation extremes with observations since trends in this variable are weak and the variability that is not SST-driven is large. Therefore, pattern correlations are close to zero, as expected, due to weak local trends that are not statistically different from zero. It is interesting to note that for most regions, observed trends lie within the model spread. Note however that studies have found that atmospheric simulations consistently produced mean and extreme precipitation changes of the wrong sign over Australia and East Asia (Dong et al., 2017). Consistent with these studies, trends in heavy precipitation of the opposite sign to the observations (albeit not significant) are found over northern Australia in the ensemble mean for some models (Fig. 4). Furthermore, the pattern correlations for the precipitation extremes are substantially higher for the coupled MIROC5 model (up to 0.41 for R95p, Figure S19) than for the C20C+ models. Therefore, caution is strongly recommended when using AMIP-style simulations for the analysis of regional trends in precipitation extremes.

5. Discussion

The results presented in this study show that SST and SIC-forcing plays an important role for temperature and precipitation extremes. In particular, inter-annual variations of globally averaged percentile temperature indices and annual maximum and minimum temperatures are driven to a large extent by variability in SSTs. It is likely that the effect of SSTs, and ENSO in particular, affects extreme indices to different extents in different seasons. Hence, by analysing annual indices of extremes and Niño3.4 we expect some of the signal to be smoothed out, especially as Niño3.4 SST anomalies usually switch sign during the first half of the calendar year. Nevertheless, this analysis shows that the climate models are able to broadly reproduce the pattern associated with the El Niño-Southern Oscillation in the observations. Although not shown here, an analysis of composites of climate extremes for El Niño and La Niña conditions separately has shown that the responses to ENSO are simulated similarly well for both phases.

Another important point to note is that the estimates provided in this study are likely to be conservative estimates of the influence of SSTs on temperature and precipitation extremes. For example, it is possible that local SSTs (e.g. unusually warm SSTs in a certain year) can be important for a particular event over a region, without this region necessarily showing a strong influence of SSTs or ENSO every year. This type of influence on individual events would not necessarily be picked up in this analysis. Nevertheless, this study identifies regions with potential “added predictability” compared to coupled models, stemming from the fact that oceanic boundary conditions are prescribed. Therefore, for regions where SSTs force a substantial part of atmospheric variability, atmospheric model simulations might provide substantial added value for investigating mechanisms or detection and attribution of specific events. However, as other studies have found, ocean-atmosphere coupling is particularly important for precipitation trends, and therefore these types of simulations may not be particularly suitable for detection and attribution of precipitation extremes. Our results however show that these simulations may be more valuable in particular for investigating variability and trends in temperature extremes.

6. Conclusions

In this study, the influence of SST-forcing on inter-annual variations
and trends in eight temperature and three precipitation extremes is examined. Using atmospheric simulations from three climate models forced by observed SSTs and SIC, the respective roles of SST-forcing and stochastic atmospheric variability are assessed for global land averages of these extremes indices. For warm temperature extremes, SST-forcing drives a large portion of inter-annual variability, particularly in the tropics. For annual precipitation extremes, the effect of SST-forcing on global average time series is smaller, and its influence is limited to the tropics, where in climate model simulations it can explain between 10 and 50% of the year-to-year variability in certain locations, depending on the index considered. The models used here are able to capture most key features in the relationship between ENSO and temperature and precipitation extremes.

Finally, our analysis shows no obvious improvement in SST-driven atmospheric models over coupled ocean-atmosphere models in the simulated regional trends in temperature extremes and suggests that the trends in precipitation extremes may agree less with observations in the C20C+ models than in coupled models. Regional details such as a decrease in the frequency of warm days over a region in southern South America are not simulated by the models, which suggests that either this observational feature is not forced by SST-variability, or that the models are unable to represent a process necessary to reproduce this feature. Since none of the ensemble members capture a cooling in this area, this would suggest a deficiency or missing process in the models. However, since the ensemble are finite, it is equally possible that our ensemble may be too small to capture the ‘true’ range of unforced variability. Similarly, the models are inconsistent with observations in large parts of North America, where the simulated increase in warm days is too strong compared to observations. The trend patterns in precipitation extremes are hard to assess in the absence of strong statistically significant local trends in the observations. However, it is worth noting that in all model simulations, the sign of the trend over North America and Australia is opposite to that found in observations, which is reflected in pattern correlations close to zero. Pattern correlations from one coupled model suggest that this problem is worse in atmosphere-only simulations than in coupled model simulations.

Acknowledgements

Authors are supported by Australian Research Council grant CE110001028. AJD also acknowledges the David Lachlan Hay Memorial Fund. SCL is supported by Australian Research Council Grant DE160100092. MGD is supported by Australian Research Council Grant DE150100456. LVA is supported by Australian Research Council Grant DP160103439.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.wace.2018.06.002.

References