Author/s:
Moise, A;Smith, I;Brown, JR;Colman, R;Narsey, S

Title:
Observed and projected intra-seasonal variability of Australian monsoon rainfall

Date:
2020-03-30

Citation:
Moise, A., Smith, I., Brown, J. R., Colman, R. & Narsey, S. (2020). Observed and projected intra-seasonal variability of Australian monsoon rainfall. International Journal of Climatology, 40 (4), pp.2310-2327. https://doi.org/10.1002/joc.6334.

Persistent Link:
https://hdl.handle.net/11343/286547

Colman Robert (Orcid ID: 0000-0002-6046-3176)

Observed and projected intra-seasonal variability of Australian monsoon rainfall.

Aurel Moise[1], Ian Smith[1], Josephine R. Brown[1,2], Robert Colman[1] and Sugata Narsey[1]

[1]Australian Bureau of Meteorology, Melbourne, Australia

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

For submission to *International Journal of Climatology*

Key points:

- Many climate models provide a reasonable simulation of present-day features of 'bursts and breaks' in the Australian monsoon
- However there is little consensus on significant changes to seasonal rainfall totals across the full ensemble
- A subset of models with detectable MJO skill shows evidence for an increase in the seasonal total rainfall amount and most other monsoon metrics, except a slight decrease in the number of burst events.

Address for correspondence:

Robert Colman

Australian Bureau of Meteorology,

GPO Box 1289,

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/joc.6334

Melbourne, Victoria, 3001,

Australia

robert.colman@bom.gov.au

ABSTRACT

Indices derived from daily rainfall time series are used to measure "burst" features of the northern Australia monsoon, corresponding to one or more days of heavy rainfall. These indices include numbers of burst days, numbers and durations of burst events, and average intensity. The results using observational data show how these features can vary from one year to the next, and how they can vary from the station scale (Darwin) to the regional scale (northern Australia). The results from CMIP5 climate model simulations under both historical and future greenhouse gas conditions have also been analysed and indicate how well models can capture these features and how they might change by the end of the 21st century under a high emissions scenario. While most models provide a reasonable simulation of present-day burst features, there is little consensus for a significant change to seasonal rainfall totals when looking at the full CMIP5 ensemble. A subset of models with detectable skills with respect to the Madden-Julian Oscillation shows evidence for an increase in the seasonal total rainfall amount and most other monsoon metrics, except a slight decrease in the number of burst events. This is consistent with a basic thermodynamic response to warming and consistent with findings elsewhere. However, the Australian monsoon is strongly influenced by the large scale circulation and there remains some doubt about whether we can confidently diagnose all the changes to monsoon bursts that could occur given the limited ability of many of the current generation of models to simulate tropical cyclones, the Madden Julian Oscillation and other relevant features.

Keywords: monsoon, variability, climate models

1. INTRODUCTION

The northern Australia monsoon is typically characterized by excessively wet conditions usually referred to as "bursts" interspersed with relatively dry "breaks". Such "bursts" and "breaks" can sometimes extend for several weeks (Drosdowsky, 1996) and, because these can have a range of impacts including flooding and recharge of freshwater ecosystems (e.g. Smith and McAlpine, 2014), it is arguably just as important to investigate these intra-seasonal features as it is to investigate seasonal totals. The intensity, frequency and duration of burst events can vary considerably, mainly because of variability associated with various rainfall-related mechanisms. For example, Madden-Julian Oscillation (MJO) events tend to reoccur on 30- to 60-day time scales and can be associated with significant rainfall anomalies in far northern coastal regions (Hendon and Liebmann, 1990) while tropical cyclones, sporadic thunderstorms, or the passage of mid-latitude synoptic-scale troughs are important at shorter time scales (Davidson et al., 2007; Wheeler and McBride, 2005, 2011). Narsey et al. (2017) found that the initiation of monsoon bursts was associated with increased cyclonic circulation in the monsoon region, influenced by mid-latitude front-like features.

Key inter-annually varying features, such as monsoon onset dates, have been estimated using a combination of rainfall and dynamic indices, typically westerly wind shifts (e.g. Hendon and Liebmann, 1990; Drosdowsky, 1996). Other studies have focussed on just the rainfall component using a variety of indices to define various key features (Nicholls et al., 1982; Smith et al., 2008; Berry and Reeder, 2016). For example, the Australian Bureau of Meteorology currently defines the northern rainfall onset date at any location as the day on which accumulated rainfall since the 1st of September reaches 50 mm, since this is considered to be approximately the amount of rainfall required to stimulate plant growth (Drosdowsky and Wheeler, 2014). Smith et al. (2008) noted long-term trends in both onset and end dates of the rainy season across northern Australia and found that increases in summer rainfall totals over the period from 1950 to 2005 (Smith, 2004) were accompanied by positive trends in average intensity. At the same time, there is evidence that the positive trends in rainfall totals have been accompanied by similar trends in the frequency of extreme rainfall events (Suppiah and Hennessy, 1996, 1998). Catto et al. (2012) suggested that changes in the large-scale circulation were more important to these trends compared to thermodynamic changes (i.e. changes due to increased temperatures and unchanged relative humidity).

If large-scale circulation changes were to occur in the future due to projected increases in greenhouse gases, we might therefore expect to see a range of changes to intra-seasonal features including numbers of rainy days, the intensity of rain that falls on rainy days, and the sequencing of rainy days including frequency and duration. These could occur independently of any changes to seasonal rainfall totals.

While some studies of climate model projections suggest changes in variability or extremes (e.g. Menon et al. 2013; Brown et al., 2017), most studies have dealt with projected changes to totals. For example, projections based on Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al. 2012) models suggest that northern Australia may experience an increase in average summer rainfall, but with a large uncertainty due to the wide model spread (Christensen et al. 2013; Jourdain et al. 2013; Brown et al. 2016). Brown et al. (2016) found that those CMIP5 models that simulated little change or increased rainfall were more plausible than those which simulated drying, as the latter tended to have larger cold biases in the western equatorial Pacific.

More recently, Brown et al. (2017) focussed on the simulations of monsoon rainfall variability and noted that variability increased on a very broad range of time scales from daily to decadal. This increase in variability in the majority of models was consistent with the increased amount of rainfall associated with typical rainfall events in a warmer climate (i.e. a "thermodynamic" response), but there was a wide range of responses between different models – from large increases to decreases in variability at all timescales. Because few model-based studies include a detailed analysis of projected changes in the sub-seasonal characteristics of Australian monsoon rainfall, it remains unclear how such increases in variability are manifest – e.g. as more frequent wet events, or no change in frequency but simply an increase in the rainfall for each event – nor how such changes vary across models.

In this study, we address this issue by focussing on intra-seasonal features of the Australian summer monsoon, including the magnitude, frequency and duration of rainfall events. It is important to note that the concepts of bursts and breaks are based on experiencing monsoon weather at the point/station scale and attempts to quantify these features can involve some degree of subjectivity. We firstly describe and demonstrate a methodology for quantifying these features using Darwin station observations. We then apply the same methodology to large-scale average values and note the differences that can be associated with such changes of scale. Thirdly, we apply the methodology to large-scale average values based on simulated values for historical (i.e. 1969 to 1999) conditions and projected future (2069 to 2099)

conditions. Specifically, we investigate the nature of the projected changes due to enhanced greenhouse gas concentrations (the RCP8.5 emissions scenario). A high emission scenario is chosen to explore the implications of a "business as usual" future pathway and to examine the impact of relatively large temperature changes.

2. METHODS

The northern Australia rainy season typically starts some time between September and December and ends rapidly in March or April (Suppiah and Hennessy, 1996). Here we focus on a fixed four-month (121 day period) covering December to March, a period during which most, but not all, of the monsoon rain falls. Thus, for example, the 1974 season runs from Dec 1 1973 to March 31 1974. Other periods could be chosen (e.g. October to April, Berry and Reeder, 2016), which affects the absolute values derived in the analysis below, but it has a minor effect on the derived differences that are the main focus of this study.

For each season, we take the 121 daily rainfall values and apply a simple 1-2-1 weighted moving average to arrive at a smoothed time series (P). The purpose of the running average is to reduce the noise in the daily values to better focus on multi-day and longer variability. The seasonal total (T) is simply the sum over all days (i, i=1 to 121), viz.:

$$T = \Sigma_i P^i$$

To remove the effect of the seasonal variation in rainfall, we calculate a smooth long-term average value for rainfall for each day throughout the season. We then define a burst index (B) that distinguishes those days (B=1) when the rainfall P exceeds these long-term rainfall values from those that do not (B=0). The total number of "burst days" in a season (N_b, which is less than the number of rainy days) is defined as:

$$N_b = \mathbf{\Sigma_i} B^i$$

In theory, N_b can range from zero (every day less than the long-term average for that day) to 121 (every day greater). It is more useful to sometimes express N_b as a percentage of the total days in the season (N_b%). The seasonal burst total (T_b) is defined as:

$$T_b = \Sigma_i B^i * P^i$$

which can also be expressed as a percentage of the seasonal total (T_b%). This helps distinguish between a season comprising a string of relatively similar rain days where the contribution from burst days may be close to 50%, from one comprising relatively few, but very intense rain days, where the contribution could approach 100%.

The average burst intensity (I_h) is defined as:

$$I_b = T_b/N_b$$

A burst "event" is defined here to be underway when at least three out of five days (centered on that day) qualify as burst days, i.e. median (B^{i-2}: B^{i+2}) =1. A sequence of days when this occurs defines a single burst event. The total number of these events in a season is denoted by N_e and the average duration (D) is simply the total number of burst days in the sequences divided by N_e. Note that there can exist quite a few "sporadic" or "short-lived" bursts in addition to burst events. Furthermore, while N_e tends to increase with the number of burst days, past a certain point the sequences begin to merge and N_e can begin to decrease. In the extreme, if every day was a burst day, there would only be a single burst event with a duration of 121 days.

The effect of the smoothing can be appreciated by considering the case of a single, isolated rainfall day (i.e. with two preceding and two succeeding completely dry days). After the smoothing, in order to be classified as a burst *day*, the raw total must exceed twice the long-term average for that day. Further, to be classified as a burst *event*, the total must exceed four times the long-term average, so that the preceding and following days also lie above average, following smoothing. This means that smoothing the data tends to focus attention on the more significant rainfall days at the expense of the low-rainfall days, and permits contributions to "burst" events of a single days rainfall, provided it is very large.

If we consider a time series of artificially generated random rainfall amounts (with no significant temporal correlation), we find that the number of "random burst days" would be (as expected) 50% and that these days would be expected to contribute about 70% to the "random seasonal" total. In addition, based on an analysis of random time series, there are about eight "random burst events" every 121 days and they tend to last about 7.5 days. These values are useful when comparing derived indices from observations and models since they provide a statistical "baseline" of burst/break characteristics. In practice we would expect significant deviation from random behaviour as we expect serial correlation of rainfall events due to synoptic variability in both observations and models.

"Break" events can be defined as all non-burst events since they correspond to periods marked by three out of five days with no burst days. Break events will not be considered in detail in this study.

Another approach to defining monsoon bursts and breaks is described by Berry and Reeder (2016) who used gridded daily rainfall data for northern Australia and define a burst as when "… the area-averaged rain transitions from at least 0.5 standard deviations below the seasonal average to at least 0.5 standard deviations above the seasonal average in less than a 7-day period". Their method is similar to that employed here, since a daily total of about twice the long-term average qualifies as a burst day under both definitions. It is worth noting that Berry and Reeder (2016) found that their results and conclusions were not substantially affected by the use of different standard deviation thresholds.

Changes in the frequency, intensity or duration of burst events can occur from one year to the next, or, in the long-term, due to climate change. In the case of long-term changes we need to consider the fact that the background climatology can also change which then affects the definitions for rainfall anomalies. If we calculate the indices using distinct climatological values from two periods (say, as simulated by climate models) we can investigate "relative" (normalized) changes – or changes in the nature of the monsoon rainfall relative to changes to the climatological (~30 year) mean. Results (not shown) indicate very little in the way of any such changes. This is an important finding in itself, since it implies that the simulated mechanisms for rainfall production, *relative to the mean,* do not change much over the period of the climate projections. On the other hand, by normalizing the data relative to the historical climatology we find substantial "absolute" changes do occur; i.e. changes that would be experienced by a historical observer who continues observing well into the future. These are of immediate interest since any such absolute changes could imply potential impacts on ecosystems, hydrology, agriculture etc.

The station data analysed here are for Darwin Airport (WMO Station 94120) while the regional average values (that encompass Darwin) refer to the “Australian monsoon domain” defined here by (land areas within) 10^{o}S-20^{o}S, 110^{o}E-150^{o}E for both observations and models, consistent with previous studies e.g. Brown et al. (2016). The Australian monsoon domain and Darwin locations are shown in Figure 1. In addition to station rainfall, we use observed rainfall based on the Australian Water Availability Project (AWAP) data set (Jones et al., 2009) which is gridded rainfall data produced by interpolating station observations to a fixed 0.05^{o} x 0.05^{o} grid. The statistics for both Darwin station and AWAP data were calculated for the period 1969-1999. We also use the monsoon domain box to calculate regional average

values from the results of CMIP5 simulations under the RCP8.5 emissions scenario (van Vuuren et al., 2011). The model-based values were extracted for both 1969 to 1999 (hereafter referred to as a "historical" period) and 2069 to 2099 (a "projection" period).

3. RESULTS FROM OBSERVATIONS

3.1 Characteristics of Darwin station rainfall

Figure 2 shows an example of Darwin station data from four different monsoon seasons that illustrate the differences in intra-seasonal rainfall variability that can occur from year to year. For example, the relatively wet 1974 season (total = 1877 mm) was characterized by many burst days (51), and seven burst events that lasted (on average) 7.3 days including many high intensity events (average 33 mm per day). The relatively dry 1985 season (total=1303 mm) was characterized by fewer burst days (38), slightly fewer burst events (6) that tended to be of shorter duration (5.9 days) and slightly less intense (30 mm per day). The next season (1986) was even drier (total =1197 mm) with even fewer burst days (35) and fewer burst events (3) even though these were relatively long-lasting (average duration 10 days). Finally, the 1999 season was another relatively wet season (total=1717 mm) that comprised many burst days (58), and six, relatively prolonged (average duration 9 days) burst events.

Time series of key indices over the 48 years 1969 to 2016 are displayed in Figure 3. The amount of rain falling during burst events ranges from 436 mm (1991) to 2069 mm (1997) and contributes, on average, 85% to the seasonal total -- dashed versus solid lines in Fig 3a (which ranges from 630 mm to 2192 mm during those years). The number of burst days ranges from as low as 17 (or 14%) in 1991 to as many as 58 (or 48%) in 1999. The number of burst events has been as low as one (1991) and as high as 10 (1997). The average duration of burst events ranges from a low of 3.7 days (1984) to a high of 13 days (2009) while the average burst intensity ranges from 20 mm per day (2001) to 45 mm per day (2010). Over the 48 years there is very little evidence for any significant long-term trends in any of the indices.

Table 1 summarizes some of the key statistics for Darwin derived for the 31-year period 1969 to 1999, to enable comparison with the model historical period. This indicates that on burst days, the average rainfall is nearly 30 mm. On average, there are about 6 bursts with a typical duration of almost 7 days.

Compared to random data, Darwin experiences fewer burst days (35% c.f. 50%) and fewer burst events (6 c.f. 8) that tend to be slightly shorter (6.7 c.f.7.5 days). This is consistent with synoptic influence on the bursts/breaks, as expected. Table 1 also shows the direct correlations between the seasonal total and other indices. Considering the variations in seasonal means across the period 1969-99, the seasonal total is closely and significantly ($p<0.01$) linked to the burst totals ($r = +0.995$, rounded to 1.00 in Table 1), to the number of burst days ($r = +0.87$), burst intensity ($r = +0.76$) and to the number of burst events ($r=+0.69$) but only weakly and insignificantly ($p>0.1$) linked to event duration ($r = +0.22$).

3.2 Characteristics of observed regional rainfall

Here we repeat the same analysis but instead use the monsoon domain average values for daily rainfall based on the observed AWAP gridded rainfall data set since these are relevant for comparison with climate model average values. This exercise is also useful since it reveals some of the differences in the statistics when changing from the station scale to the large scale. One major difference from the Darwin station values is that there are few completely dry days. i.e. during the northern Australia monsoon season it is likely to rain at some location within the domain in a 24 hour period. The area average values will also be compared with area averages from climate models in the next Section.

The monsoon domain average values are also shown in Table 1 and indicate some significant differences with the Darwin values. Monsoon domain total rainfall (683 mm) is much less than at Darwin (1459 mm) and it is characterized by slightly less burst days (31% vs 35%) which tend to contribute less (77% c.f. 85%) to the seasonal totals. The burst events (5.1 vs 6.0) are fewer at the regional scale but tend to last longer (7.4 compared to 6.7 days).

In all four cases (1974, 1985, 1986 and 1999) the contrast with Darwin (Figure 4 vs Figure 2) highlights the very different nature of rainfall time series at regional and station scales. The AWAP data shows that 1974 was an extremely wet year (total 1046 mm) with a high proportion of burst days (67%) that contributed 77% to the total. The relatively dry 1985 season (total 302 mm) was characterized by relatively few burst days that contributed only 7% to the total and by only two burst events. The following year (1986) was much wetter (758 mm) and notable for a relatively large number of burst events (7) while the even wetter 1999 season (total 1038 mm) comprised only three, but significantly protracted events – the longest lasting for 45 days.

It is worth noting that Berry and Reeder (2016) analyzed the same regional data over the period 1979 to 2010 and also focused on burst events, although using a different definition to that used here. While they estimated a much higher average number (7 versus 4.5 events) per season, this is mainly because they aggregated over 7 months (October to April) compared to the just the 4 months (December to March) considered here. They also noted that the distribution of return periods indicates a substantial percentage (38%) occurred between 10 and 20 days with a secondary peak (18%) occurring between 30 and 40 days. This most likely reflects the differences in the months used in the analyses since it is apparent that a substantial number of bursts occur between October and November. The different periods considered may also influence results, given the large year to year variability in burst characteristics (see Figure 2).

Other major differences between monsoon domain and station rainfall are indicated by the different correlations between seasonal totals (T) and: burst totals (r=+0.91 vs +1.00), the number of burst events (r=+0.41 versus +0.69) and duration (r=+0.71 versus +0.22). This last difference, plus the longer average duration of regional scale burst events (7.4 days versus 6.7 days) clearly reflects the influence of large scale phenomena such as tropical cyclones and intra-seasonal synoptic variability (e.g. MJO). A tropical cyclone, for example, can affect the rainfall at a station for a few days whereas it could affect the monsoon domain for up to a week or more depending on its speed and trajectory across the region. An example is TC Steve, which spent 13 days traversing Queensland, the Northern Territory and Western Australia in February-March 2000 and caused widespread heavy rainfall, evident in the extended peak in Figure 4d. In this situation, the longevity of the system becomes an important determinant at the regional scale but less so for a single station. This will also likely be the case for rainfall events associated with mid-latitude disturbances (Berry and Reeder, 2016, Narsey et al., 2017) or the Madden Julian Oscillation (Hung and Yanai, 2004, Hung et al., 2013, Ahn et al., 2017).

4. RESULTS FROM CMIP5 MODELS

4.1 Historical intra-seasonal rainfall variability

Figure 5 compares the major indices from each of 35 CMIP5 model simulations (one simulation per model, see Table 2 for a list of models) for the monsoon average region (see Fig.1) for historical (1969 to 1999) conditions. The models tend to overestimate the AWAP seasonal total of monsoon domain wet

season rainfall since 15 out of the 35 values fall above the AWAP uncertainty range (defined here as the average plus/minus one standard deviation) and only four models lie below the range (Figure 5a). On the other hand, most models yield percentage burst totals that lie within the (plus/minus one standard deviation) uncertainty band (Figure 5b). As could be expected, most model values lie closer to the AWAP value (77%) than the higher Darwin station value. With two exceptions (models 29 and 30) the percentage number of burst days (31%) is also reasonably simulated by the models (Fig. 5c). The two exceptions correspond to rainfall distributions that can be described as too sporadic.

While the models tend to overestimate the average number of burst events relative to AWAP observations (5.1), consistent with Narsey et al. (2018), most provide reasonable values that lie within one standard deviation of the AWAP value (Figure 5d).

Figure 5e shows that all models simulate event durations that are longer than the AWAP average value (7.4 days). In fact a third of all models appear to yield events that are far too long in duration compared to the events that characterize both Darwin data and a random process. Only one models (10) lies outside the uncertainty band.

Finally, the results for burst intensity (Figure 5f) indicate that the majority of the models yield values comparable with the AWAP value (17 mm per day), despite the underlying wet bias in the seasonal totals. One interpretation is that the simulated "dry" events tend to be more similar to the AWAP events than are "wet" events.

In general, many models simulate the monsoon features within the uncertainty range of AWAP (see Fig. 5). Stratifying model projections based on their ability to simulate present-day features of the climate does not always lead to less uncertainty (e.g. Smith and Chandler, 2010, Brown et al., 2016) but we can attempt to do this by identifying obvious outliers and reducing the model ensemble. If we focus on values that both represent the top two extreme values (minima and maxima) and also lie outside the AWAP uncertainty range, we can identify ten models (3,4,9,10,14,25,26,28,29 and 30 - see Table 2) that could be described as outliers. Certainly, at least one model (3) can be regarded with caution since, while it overestimates the contribution from burst days, it underestimates both the number of burst days and the number of burst events while severely overestimating burst intensity. By excluding these 10 models we can refer to a smaller (N=25) stratified ensemble which may be regarded as more representative of the AWAP values and which might provide less uncertain projections. While other criteria for excluding

models could be used, we restrict our analysis to the entire ensemble (N35), the reduced ensemble from the metrics described above (N25) and an additional subset with respect to MJO skill (N7, see below).

If we calculate the monsoon burst indices at each grid point, we can display the spatial variations in the indices from both the models and the AWAP data (see Supplementary Figure S1). This is also useful in illustrating the coarse scale of current climate models relative to the fine scale of the features of interest. Supplementary Figure S1 shows that the models are able to capture the spatial patterns of the indices reasonably well. In addition, the indices are spatially homogenous over the region of the monsoon domain, providing support for the use of area average values.

Table 3 summarizes the ensemble statistics and shows that there is not a great deal of difference between the 35-member ensemble (N35) and the stratified (N25) ensemble mean values including the inter-model correlations between the totals and the indices. Other than the expected strong association with burst totals, the only other significant association is with intensity. i.e. the model ensembles suggest that changes in seasonal totals are more likely to arise from changes to intensity rather than changes in other features of the rainfall series. The smaller N25 ensemble (excluding 10 outlier models) shows a stronger correlation with burst intensity compared to the full N35 ensemble. We also note that Narsey et al. (2018) found a weak positive correlation in CMIP5 models between the seasonal rainfall total and number of bursts, but using the full October-March wet season period. The AWAP-derived values indicate moderate to strong correlations between seasonal rainfall totals and numbers of burst days, number of events and duration which are not replicated in the CMIP5 model results.

The lack of significant correlation between burst event number or duration and seasonal rainfall totals in the models may reflect the influence of small-scale phenomena (e.g. tropical cyclones) in observations that may not be well represented by global climate models (Flato et al. 2013, Ng et al., 2015) due to their relatively coarse resolution. More recently, Narsey et al. (2018) evaluated the performance of CMIP5 models at simulating burst features of the monsoon region. They found that midlatitude influences were reasonably well captured but found that the representation of MJO influences was poor. It should be noted that coupled models often struggle to represent many aspects of MJO variability skilfully even though most CMIP5 models are capable of producing some MJO-like dynamical features (e.g. Hung et al., 2013 Ahn et al., 2017). Here we analysed CMIP5 models with respect to their skill on simulating

processes relevant to the MJO, using two well-known MJO metrics from Sperber and Kim (2012) to assess this skill: (a) the lead-lag relationship of the first 2 EOF's of filtered outgoing longwave radiation (OLR) data; and (b) the east-ward propagating power as measured by the frequency-spectrum analysis from daily rainfall data. Figures S2 and S3 describe the results in more detail. Most of the models simulate a lag correlation structure similar to that of the observed OLR (AVHRR, thick black line), although there are several models whose correlation structures are profoundly different from observations. The first metric consists of the maximum positive correlation and the time lag at which it occurs. The maximum positive correlation is a measure of how coherent and/or dominant is the propagation of convective anomalies from the Indian Ocean to the Maritime continent. The time lag is the time that it takes for the system to transition from EOF-2 to EOF-1. From observations the maximum positive correlation is 0.63, which occurs at a time lag of 11 days.

For the second metric the frequency-wavenumber power spectra of 10N – 10S averaged GPCP and CMIP5 model daily precipitation for November to April 1997–2008 is calculated. This level-2 diagnostic from the CLIVAR MJO Working Group (CLIVAR, 2009) shows the spectral power for eastward *versus* westward frequencies (positive frequencies correspond to eastward propagation) for wavenumbers 0–8. For rainfall, eastward propagating power is strongest in the 30–80 day band for wavenumbers 1–3, indicative of the MJO. The East/West power ratio, calculated by dividing the sum of the eastward propagating power by the westward propagating counterpart for the aforementioned MJO frequencies and wavenumbers, is the second metric used here to assess if eastward propagating intraseasonal variability dominates in the MJO band. It is dimensionless. Most CMIP5 models underestimate MJO amplitude, especially when outgoing longwave radiation (OLR) is used in the evaluation, and exhibit too fast phase speed while lacking coherence between eastward propagation of precipitation/convection and the wind field (Ahn et al., 2017). We identified a small group of CMIP5 models which showed the highest skill along these two metrics and Table 2 then identifies these 7 models which were diagnosed with the most realistic MJO-like features. Additionally, we also identified 7 CMIP5 models with relatively poor MJO-like features measured by these two metrics. It is worth noting that none of the outlier models exhibit realistic MJO-features while two outlier models (4 and 23) exhibit poor MJO features.

Figure S4 shows the comparison of the 5 monsoon metrics for all ensembles mentioned (N35, N27, N7good, N7bad) against the observed value from AWAP.

4.2 Projected changes to intra-seasonal rainfall variability

We consider here changes simulated for the high emissions RCP8.5 scenario and compare the indices for the historical period (1969-1999) relative those calculated for the late 21st century (2069-2099). Figs. 6a to 6e show the projected changes in the indices from each of the 35 CMIP5 models while Table 4 summarizes the results for all ensembles. Note that, in each case in Figure 6, most of the 10 outlier models are identified, irrespective of how they perform at reproducing each individual index.

Several models project very large increases or decreases in total rainfall. Model 14 projects the largest decrease (-42%) which has been noted previously to be related to warming in the western equatorial Pacific where the model also has a large negative sea surface temperature (SST) bias (Brown et al., 2016). The present analysis finds that this decrease is associated with a decrease in the percentage of rain that falls on burst days (-48%, Fig 6b), a decrease in burst days (-58%, Fig.6c) and a large decrease (-42%) in the number of burst events (Fig.6d). Model 5 projects the largest increase in seasonal totals (+32%), due to a relatively large increase in burst totals (+46%) number of burst days (+30%) and also number of burst events (+6%). Model 14 (an outlier) yields a relatively large decrease in event duration (-25% days), while model 9 (another outlier) yields a relatively large increase in duration (+10%). Model 11 yields the largest increase in intensity (+23%). Overall, most of the models simulate more often decreases in monsoon metrics than increases.

Table 4 indicates that the full ensemble yields a slight decrease (-1%) in seasonal rainfall totals but given that inter-model spread is fairly high (Table 3) or that the interannual standard deviation of the observed totals is 23% (Table 1) this is insignificant. The stratified ensemble N25 shows similar values for mean changes compared to the full ensemble (see Table 4), most likely because even though outlying models are all decreasing, the large ensemble size prevents this from showing overall impact. However, there is a strong separation in the mean changes when the stratification is along MJO skill: the N7good ensemble simulates an increase in total seasonal rainfall of 5% while the N7bad ensemble shows a mean decrease of 7%. This difference in total seasonal rainfall changes is significant, which is also the case for all other

metrics listed in Table 4 comparing the N7good ensemble to the N7bad ensemble (even at that smaller sample size).

To emphasize this, we show in Figure 7 not only the mean ensemble changes across all ensembles for all monsoon metrics, but also the spread in the ensemble change (expressed as the difference between the 19th and 90th percentile). Compared to the full ensemble, the N7good ensemble shows an increase in total seasonal DJFM rainfall which is also in accordance with the analysis of Brown et al. (2016) who concluded that models with projected decrease in rainfall had very strong biases in historical simulations and are therefore deemed less confident. Figure 7 also shows the corresponding stronger increase in total burst rainfall, burst duration and burst intensity compared to the full ensemble. It also highlights the opposite direction of the changes simulated by the N7bad ensemble: models without any MJO skill in their historical simulations simulate a drying of seasonal DJFM rainfall, while models with detectable MJO skill (see figure in supplement section) clearly favor increased DJFM rainfall over Northern Australia.

Another way of understanding the projected changes is to form a scatter plot of projected versus historical values including the historical AWAP-based values for comparison (Figure 8). In each case a line corresponding to a one-to-one correspondence (i.e. no projected change) has been drawn to delineate projected increases (above the line) from projected decreases (below the line). These plots firstly indicate any relationship between the nature of the projected changes and how wet/dry the models are and, secondly, whether there is any evidence of an improved consensus depending on how similar the model values are to the observed/AWAP values. The models from the N7good ensemble are identified in all scatterplots.

In the case of seasonal totals, (Figure 8a) there is little evidence for either effect from the full ensemble, but it does suggest that the results from the two wettest models (5 and 14) may need to be treated with caution since they appear to deviate significantly from the consensus. These models are part of the group of outlier models shown in Figure 5. Figure 8b suggests that there are several models projecting strong decreases in burst totals some of which are also outlier models.

Is there any evidence to indicate that the historical values affect the projected values ? For example, do relatively "wet" models get wetter and vice-versa? Figures 8a and 8b show that there is little preference

for increases or decreases to cluster towards the right and left of the diagrams considering the full ensemble. There is some suggestion that models which simulate relatively large numbers of burst days (Figure 8c) tend to project increases in these indices but, at the same time, most models project an increase in intensity (Figure 8f). Secondly, is there any evidence that historical values that cluster near the AWAP values (i.e. more realistic models) differ from the ensemble of results? Apart from changes to duration (Figure 8e) there is little evidence for this since there is no preference for either increases or decreases in the vicinity of values that best match the AWAP values (the red dots).

Does the ability to represent MJO-like features have an effect on the projections? Focussing on the green circles in Figure 8a-f, we see the tendency to be above the 1-1 line for seasonal burst totals, duration and intensity changes (Figure 8d), suggesting more confident projections in increases in these monsoon metrics for future warming under the RCP8.5 scenario.

5. DISCUSSION AND CONCLUSIONS

In this study, we have examined daily rainfall time series that are associated with "bursts" of the Australian monsoon and have defined some relatively simple indices that specifically measure features such as numbers of burst days, numbers and durations of burst events, and average intensity. We have analysed observed daily rainfall data at both station (Darwin) and regional (northern Australia) scales and have performed the same analysis of data, at the regional scale, from CMIP5 model simulations under both historical and future greenhouse gas conditions.

The results show how these features of daily rainfall time series can vary from one year to the next, and how they can vary from the point scale to the regional scale. In some respects, Darwin rainfall is more similar to a random time series compared with the regional scale rainfall – especially with regard to the relative contribution of rainfall bursts to the seasonal total, the number of burst events and their duration. While generally drier at the regional scale, there are more burst days but fewer, and longer burst events.

In general, the models provide a reasonable simulation of the present-day rainfall characteristics at the regional scale. While seasonal rainfall totals are underestimated by most models relative to AWAP, the burst-break representation of the monsoon is captured quite well since most model values tend to lie within the range of uncertainty associated with the AWAP observed values. In projections under the high

emission RCP8.5 scenario, the models yield a wide range of results with little consensus for a change in seasonal totals of either sign by the end of the century. Stratifying the models by omitting those which appear as outliers in the historical simulations has little effect on the ensemble-mean results for future changes. Neither is there much evidence for any difference in the simulated changes based on whether the models are relatively wet or dry.

While most CMIP5 models underestimate MJO amplitude and exhibit too fast phase speed while lacking coherence between eastward propagation of precipitation/convection and the wind field (Ahn et al. 2017), our evaluation revealed a small group of CMIP5 models with above average skill along two MJO metrics (see Supplementary section). A separate analysis of these more skilful models reveales significant differences in projected changes of the monsoon metrics between the more skilful models and a similar size group of least skilful models.

Our results suggest that while there is little consensus for a significant change to rainfall totals in the full CMIP5 ensemble, there is some evidence for a slight increase when sub-setting the ensemble along metrics assessing MJO skill. We also see an increase in burst totals and intensities that are counterbalanced by a slight decrease in the number of burst events. The most likely interpretation being that, in a warmer world, there will be fewer but more intense rainfall events. This will mainly represent a thermodynamic response to warming and is consistent with findings elsewhere that indicate future rainfall will be more intense over a range of timescales (e.g. Allan and Soden 2008; O'Gorman and Schneider 2009; Menon et al. 2013; Sharmila et al. 2015; Brown et al. 2017). Nevertheless, the results suggest that any such changes may be relatively small, particularly when compared with the interannual variability seen in the observations.

It is also worth noting again that, for the northern Australia monsoon region, interannual variability strongly reflects the influence of large-scale phenomena such as tropical cyclones, MJO events and mid-latitude disturbances. As already noted, the ability of many models to represent these features is somewhat limited and it is quite possible that this contributes to the relatively small overall response seen in the model projections. Consequently, the projections may only represent part of the response that could occur in the future. In this context, the results imply that the development and testing of higher spatial

resolution climate models, or models with revised physical parametrisations may be more useful than simply increasing the size of multi-model ensembles.

.

ACKNOWLEDGEMENTS

We acknowledge constructive comments on earlier versions of this paper by Drs Linden Ashcroft and Simon Grainger, and two anonymous reviewers. The research presented in this paper was jointly supported by the Australian Bureau of Meteorology and the Australian Government's National Environmental Science Programme. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups (listed in Table 2 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. All CMIP5 model output used in this study can be obtained from the online database as listed at the PCMDI website: http://cmip-pcmdi.llnl.gov/mips/cmip5/availability.html. For further information regarding model and observational data, analysis methods and scripts contact the corresponding author.

REFERENCES

Ahn,, M. S., Kim, D., Sperber, K. R., Kang, I. S., Maloney, E., Waliser, D. and Hendon, H (2017) MJO simulation in CMIP5 climate models: MJO skill metrics and process-oriented diagnosis, *Clim Dyn*, doi.org/10.1007/s00382-017-3558-4.

Allan, R.P., and Soden B.J., (2008) Atmospheric warming and the amplification of precipitation extremes. *Science* 321:1481–1484.

Berry, G.J and Reeder, M.J (2016) The dynamics of Australian monsoon bursts. *J. Atmos. Sci.*, 73, 55-69.

Brown, J.R, Moise, A. F., Colman R., and Zhang, H. (2016), Will a warmer world mean a wetter or drier Australian monsoon? *J. Clim.*, 29, 4577-4596, doi:10.1175/JCLI-D-15-0695.1.

Brown, J.R., Moise, A. F., Colman R. (2017), Projected increases in daily to decadal variability of Asian-Australian monsoon rainfall. *Geophys. Res. Lett.*, 44, 5683–5690, doi.org/10.1002/2017GL073217.

Catto, J. L., Jakob,C. and Nicholls, N. (2012), The influence of changes in synoptic regimes on north Australian wet season rainfall trends, *J. Geophys. Res.*, 117, *D10102, doi:10.1029/2012JD017472.*

Christensen, J.H., and Coauthors (2013), Climate phenomena and their relevance for future regional climate change. Climate Change *2013: The Physical Science Basis*. T. F. Stocker et al., Eds., Cambridge University Press, 1217–1308.

CLIVAR MJO Working Group (2009) MJO Simulation Diagnostics. Journal of Climate, 22(11), 3006-3030, doi.org/10.1175/2008JCLI2731.1

Davidson, N.E., Tory, K.J. Reeder, M.J. and Drosdowsky W.L. (2007), Extratropical–tropical interaction during onset of the Australian monsoon: Reanalysis diagnostics and idealized dry simulations. *J. Atmos. Sci.*, 64, 3475–3498, doi:10.1175/JAS4034.1

Drosdowsky, W. (1996). Variability of the Australian summer monsoon at Darwin: 1957–1992. *J. Clim.*, 9(1), 85-96.

Drosdowsky, W., and Wheeler M.C. (2014) Predicting the onset of the north Australian wet season with the POAMA dynamical prediction system. *Wea. Forecasting*, 29, 150-16

Flato, G. and Coauthors (2013), Evaluation of Climate Models. Climate Change *2013: The Physical Science Basis*. T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

Hendon, H. H., and Liebmann B. (1990), The intraseasonal (30–50 day) oscillation of the Australian summer monsoon. *J. Atmos. Sci.*, 47, 2909–2924.

Hung, C. W., and Yanai, M. (2004). Factors contributing to the onset of the Australian summer monsoon. *Quart. J. of the Royal Met. Soc.*, 130(597), 739–758.

Hung, M-P, Lin J-L, Wang W, Kim D, Shinoda T, Weaver SJ (2013), MJO and convectively coupled equatorial waves simulated by CMIP5 climate models. *J.Clim.* 26:6185–6214

Jones, D. A., Wang, W. and Fawcett R. (2009), High-quality spatial climate data-sets for Australia. *Aust. Met. and Oceanog. J.*, **58**, 233–248.

Jourdain, N. C., A. Sen Gupta, A. S. Taschetto, C. C. Ummenhofer, A. F. Moise, and K. Ashok (2013), The Indo-Australian monsoon and its relationship to ENSO and IOD in reanalysis data and the CMIP3/CMIP5 simulations. *Clim. Dyn.*, 41, 3073–3102, doi:10.1007/s00382-013-1676-1.

Menon, A., Levermann A. and Schewe G. (2013) Enhanced future variability during India's rainy season, *Geophys. Res. Lett.*, 40, 3242–3247, doi:10.1002/grl.50583.

Narsey, S., Reeder, M.J. Ackerley D. and Jakob C. (2017), A midlatitude influence on Australian monsoon bursts, *J. Clim*, 30, 5377-5393, doi:10.1175/JCLI-D-16-0686.1

Narsey, S., Reeder, M. J., Jakob C. and Ackerley D. (2018), An evaluation of Northern Australian wet season rainfall bursts in CMIP5 models. *J. Clim*, 31, 7789-7802, doi: 10.1175/JCLI-D-17-0637.1.

Ng, B., Walsh K. and Lavender, S. (2015), The contribution of tropical cyclones to rainfall in northwest Australia, *Int. J. Climatol*, 35, 2689-2697.

Nicholls, N., McBride, J.L. and Ormerod R.J. (1982), On predicting the onset of the Australian wet season at Darwin. *Mon. Wea. Rev.*, 110, 14–17, doi:10.1175/1520-0493(1982)110<0014:OPTOOT>2.0.CO;2

O'Gorman, P.A., and Schneider, T. (2009), The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences* 106.35 (2009): 14773-14777.

Sperber, K. R. and D. Kim (2012) Simplified metrics for the identification of the Madden–Julian oscillation in models, *Atmos. Sci. Let.* 13: 187–193 (2012) DOI: 10.1002/asl.378.Sharmila, S., Joseph, S. Sahai, A.K. Abhilash S. and Chattopadhyay R. (2015) Future projection of Indian summer monsoon variability under climate change scenario: An assessment from CMIP5 climate models. *Glob. and Planetary Change* 124 (2015) 62–78

Smith, I. (2004). An assessment of recent trends in Australian rainfall. *Aust. Met. Mag.*, 53(3), 163-173.

Smith I, and McAlpine C (2014) Estimating future changes in flood risk: case study of the Brisbane river, *Australia. Clim Risk Manag* 6:6–17. doi:10.1016/j.crm.2014.11.002

Smith, I. N., Wilson, L., and Suppiah, R. (2008). Characteristics of the northern Australian rainy season. *J. Clim.* 21(17), 4298-4311.

Smith, I., and Chandler, E. (2010), Refining rainfall projections for the Murray Darling Basin of south-east Australia —The effect of sampling model results based on performance, *Clim. Change*, 102, 377–393, doi:10.1007/s10584-009-9757-1

Suppiah, R., and Hennessy, K. J. (1996), Trends in the intensity and frequency of heavy rainfall in tropical Australia and links with the Southern Oscillation, *Aust. Met. Mag.*, 45, 1-17.

Suppiah, R., and Hennessy, K. J. (1998), Trends in total rainfall, heavy rain events and number of dry days in Australia, 1910–1990, *Int. J. Climatol.*, 10, 1141–1164.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A. 2012: An Overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.* **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

van Vuuren et al., (2011) The representative concentration pathways: an overview. Clim. Change, 109, 5–31, doi:10.1007/s10584-011-0148-z.

Wheeler, M., Maloney, E., and MJO Task Force (2013) Madden-Julian Oscillation (MJO) Task Force: a joint effort of the climate and weather communities, CLIVAR Exchanges No. 61, Vol. 18, No.1, March 2013.

Wheeler, M.C. and McBride J. L., (2005) Australian-Indonesian monsoon. In: W.K.M. Lau and D.E. Waliser (eds), Intraseasonal Variability in the Atmosphere-Ocean Climate System. Praxis, Springer Berlin Heidelberg, pages 125-173.

Wheeler, M.C. and McBride, J.L. (2011) Australasian monsoon. In: W.K.M. Lau and D.E. Waliser (eds), Intraseasonal Variability *in the Atmosphere-Ocean Climate System* (2nd edition). Springer, p147-198.

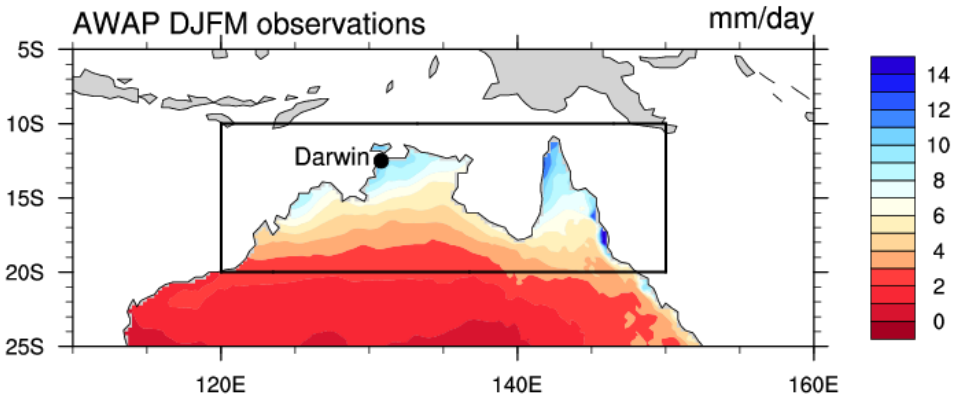

Figure 1. The location of Darwin station (WMO Station 94120) and the Australian monsoon domain. The contours show average December-March seasonal mean rainfall (mm/day) from AWAP data (1900-2005).

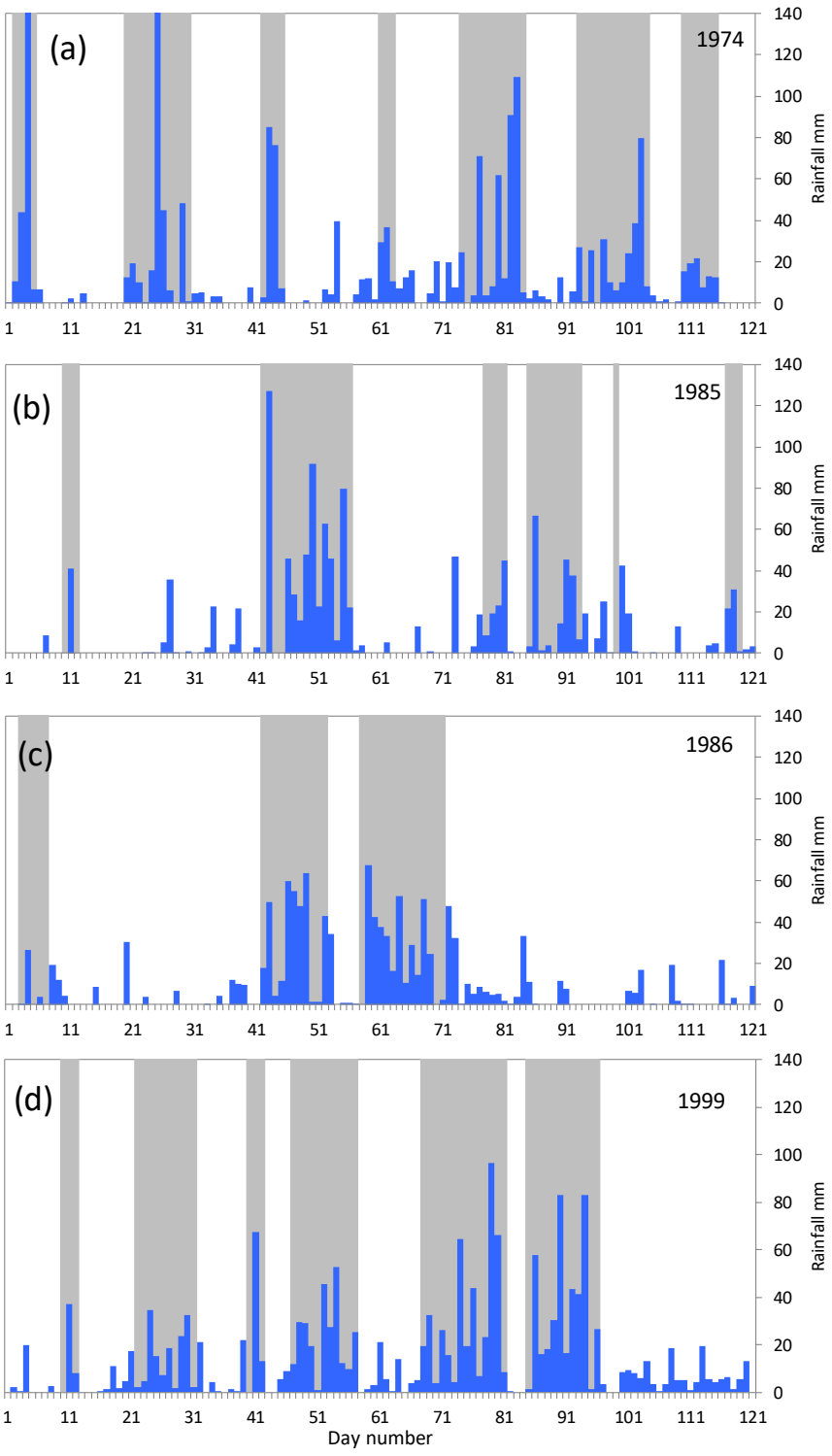

Figure 2. Daily rainfall time series for Darwin station from four different monsoon seasons: (a) 1974, (b) 1985, (c) 1986 and (d) 1999. The raw (i.e. non-normalized, unsmoothed) values are indicated in blue while the calculated burst events are indicated by the grey shading. The two values in 1974 on days 4 and 25 that are off scale are 151 and 277 respectively.

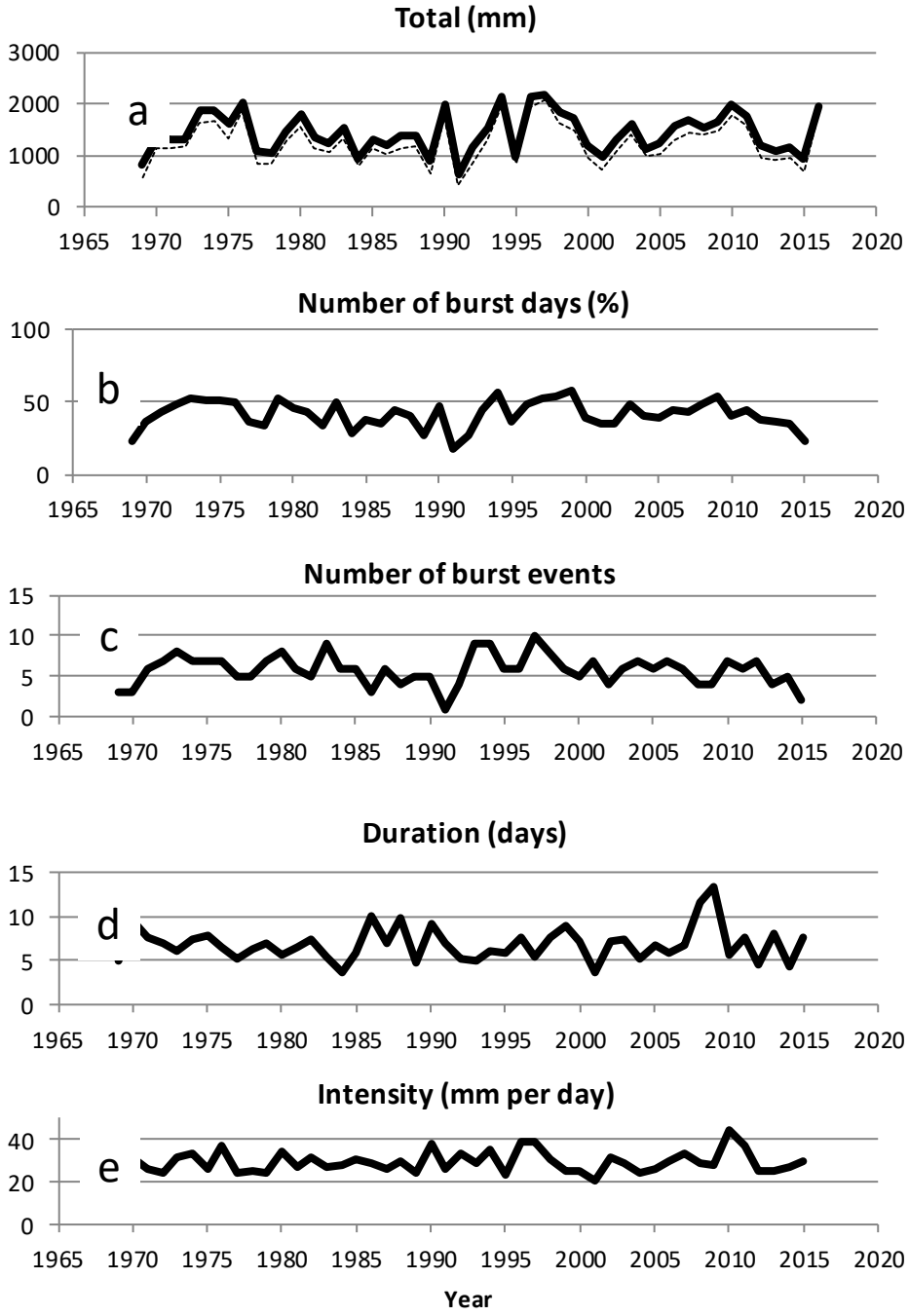

Figure 3. Time series of Darwin monsoon season indices over the 48 years 1969 to 2016: (a) seasonal total (solid line) and burst totals (dashed line), (b) percentage numbers of burst days, (c) number of burst events, (d) duration of burst events and (e) intensities of burst days.

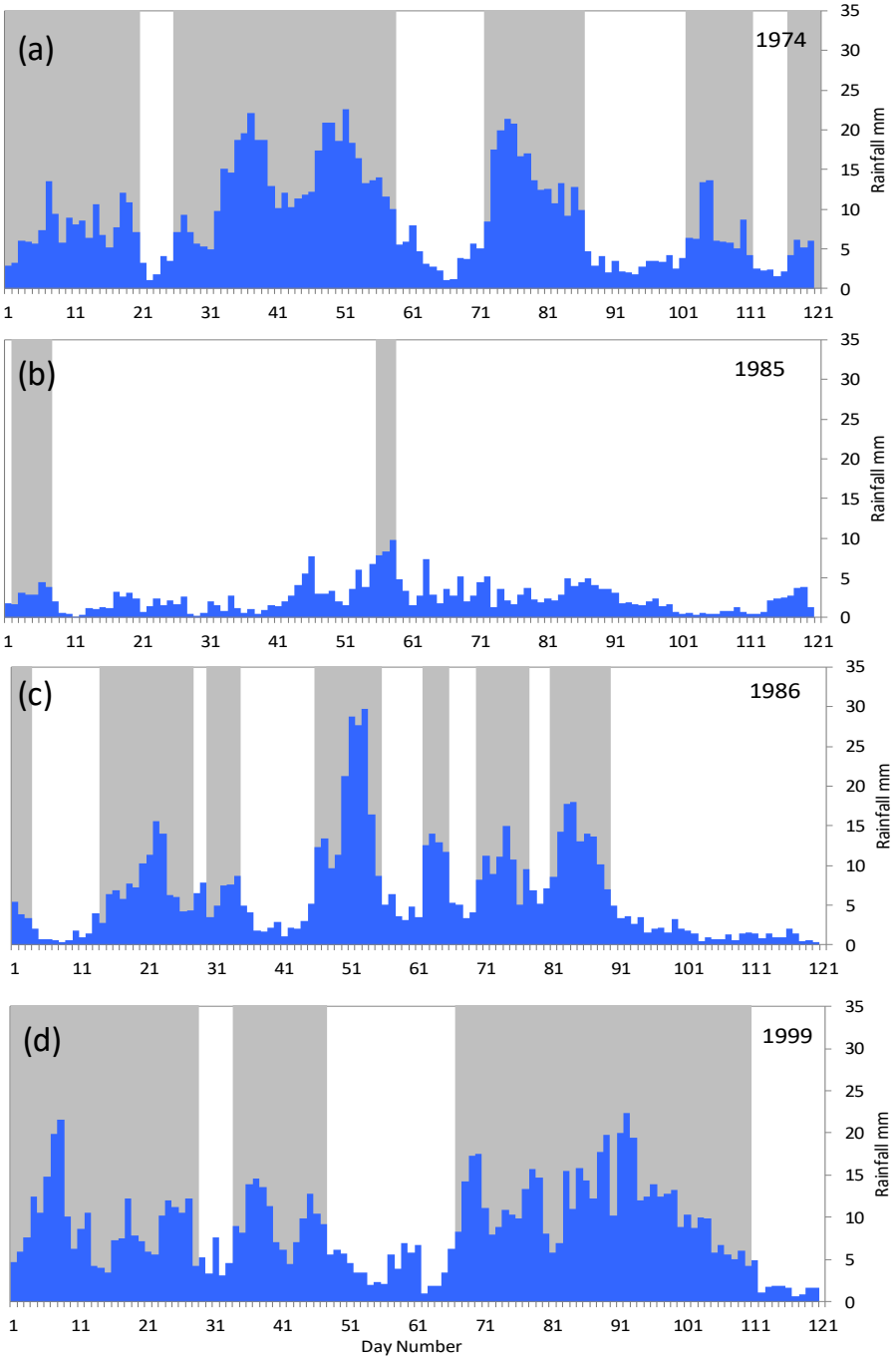

Figure 4. As for Figure 2 except the analysis is based on monsoon domain average values for daily rainfall (AWAP gridded data).

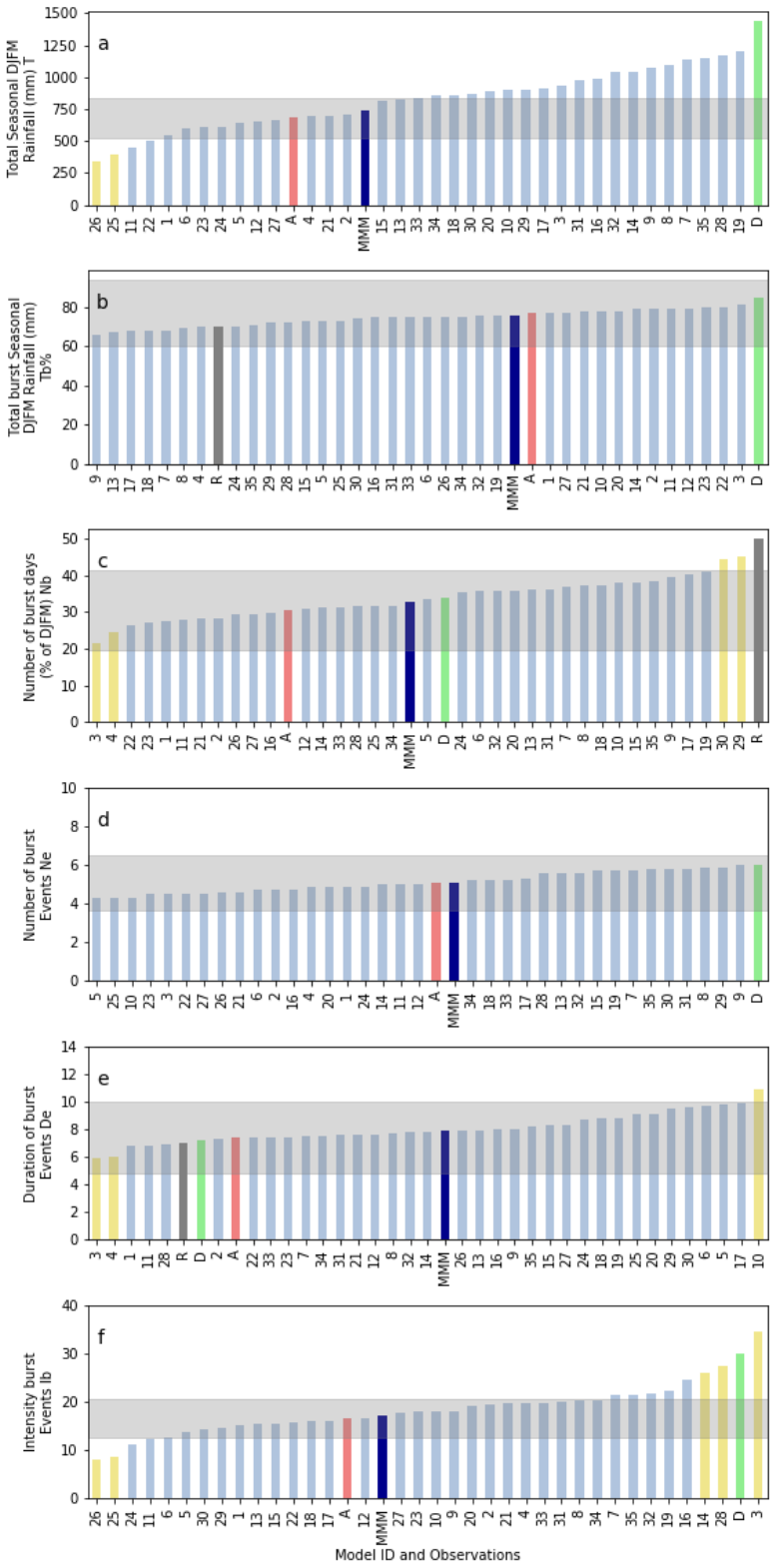

Figure 5. A comparison of the monsoon indices from each of the 35 CMIP5 model historical simulations (1969 to 1999) with the average from AWAP data (A, red bars), Darwin station data (D, green bars) and randomized data (R, grey bars): (a) DJFM seasonal total rainfall, (b) percentage burst total rainfall, (c) percentage burst days during DJFM, (d) number of burst events, (e) burst event durations and (f) burst event intensities. The grey shading represents the AWAP average values

(plus/minus one standard deviation). The most extreme model values that also lie outside this range are indicated in yellow and the ensemble mean model (MMM) is shown in dark blue.

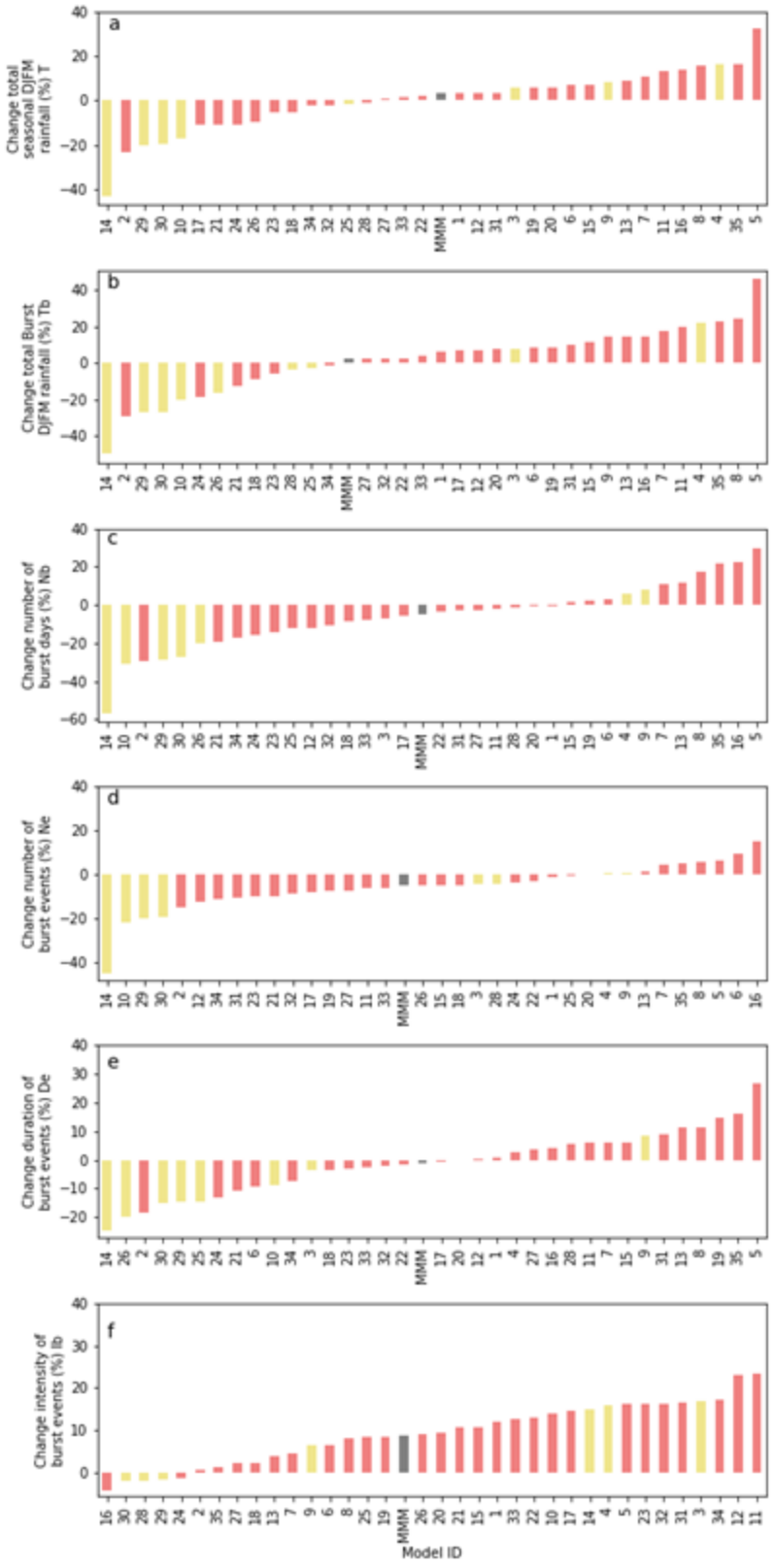

Figure 6. The projected percentage changes (Δ%) in monsoon indices from each of the 35 CMIP5 models: (a) DJFM seasonal total rainfall, (b) percentage burst total rainfall, (c) percentage burst days during DJFM, (d) number of burst events, (e) burst event durations and (f) burst event intensities. The

changes refer to average values for the late 21st century (2069-2099) compared to the historical period (1969-1999) under the high emission RCP8.5 scenario. Outlier models from Figure 5 (those that lie outside the grey ranges) are indicated in yellow and the ensemble mean model (MMM) change is shown in grey.

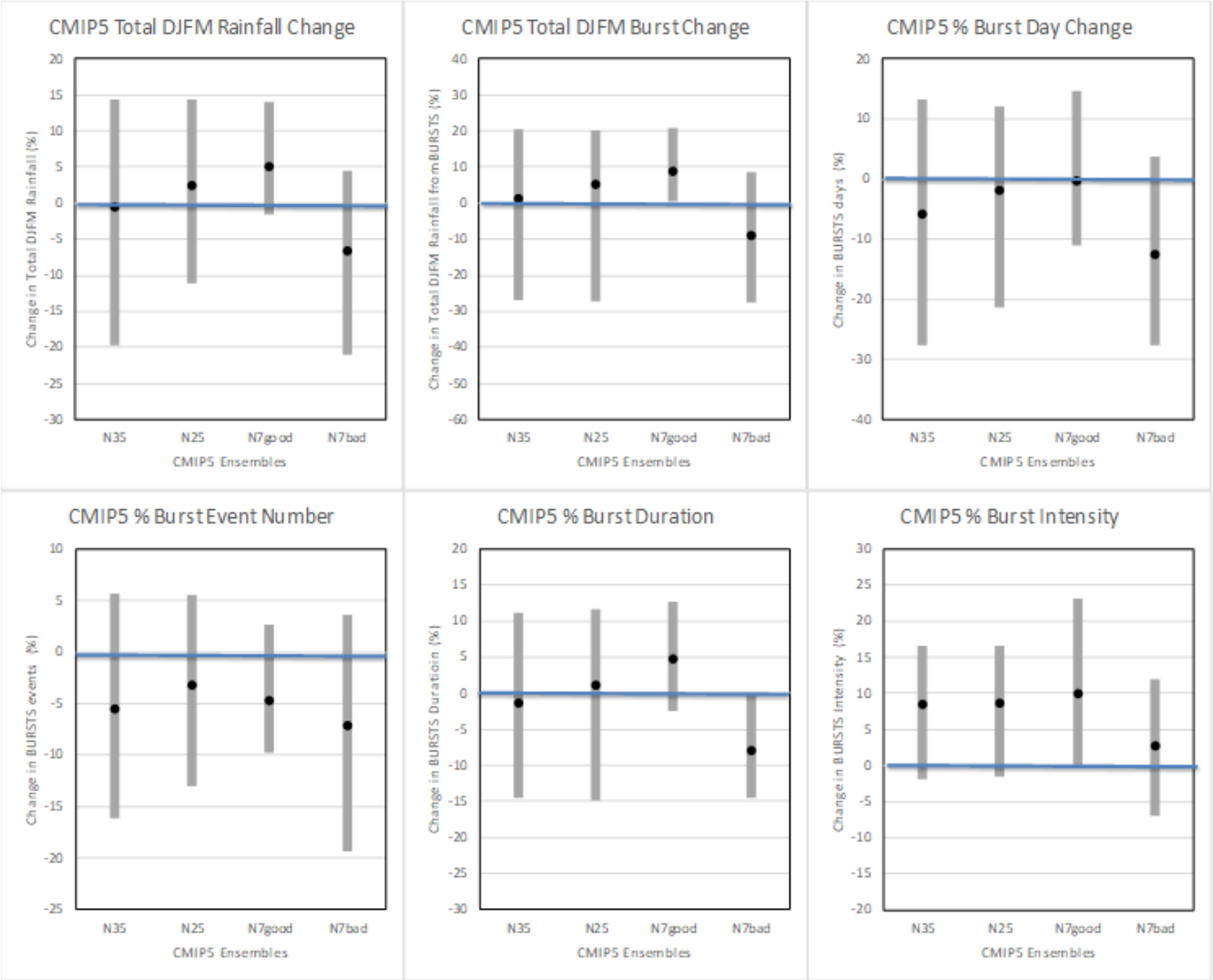

Figure 7. Change in monsoon metrics for the various ensembles (N35, N25, N7good and N7bad): (a) Change in DJFM seasonal rainfall totals (%), (b) Change in DJFM seasonal burst totals (%), (c) Change in number of burst days (%), (d) Change in number of burst events (%), (e) Change in burst event duration (%) and (f) Change in burst event intensities (%). Each bar shows the mean change and the 10^{th} and 90^{th} percentile of change within each ensemble.

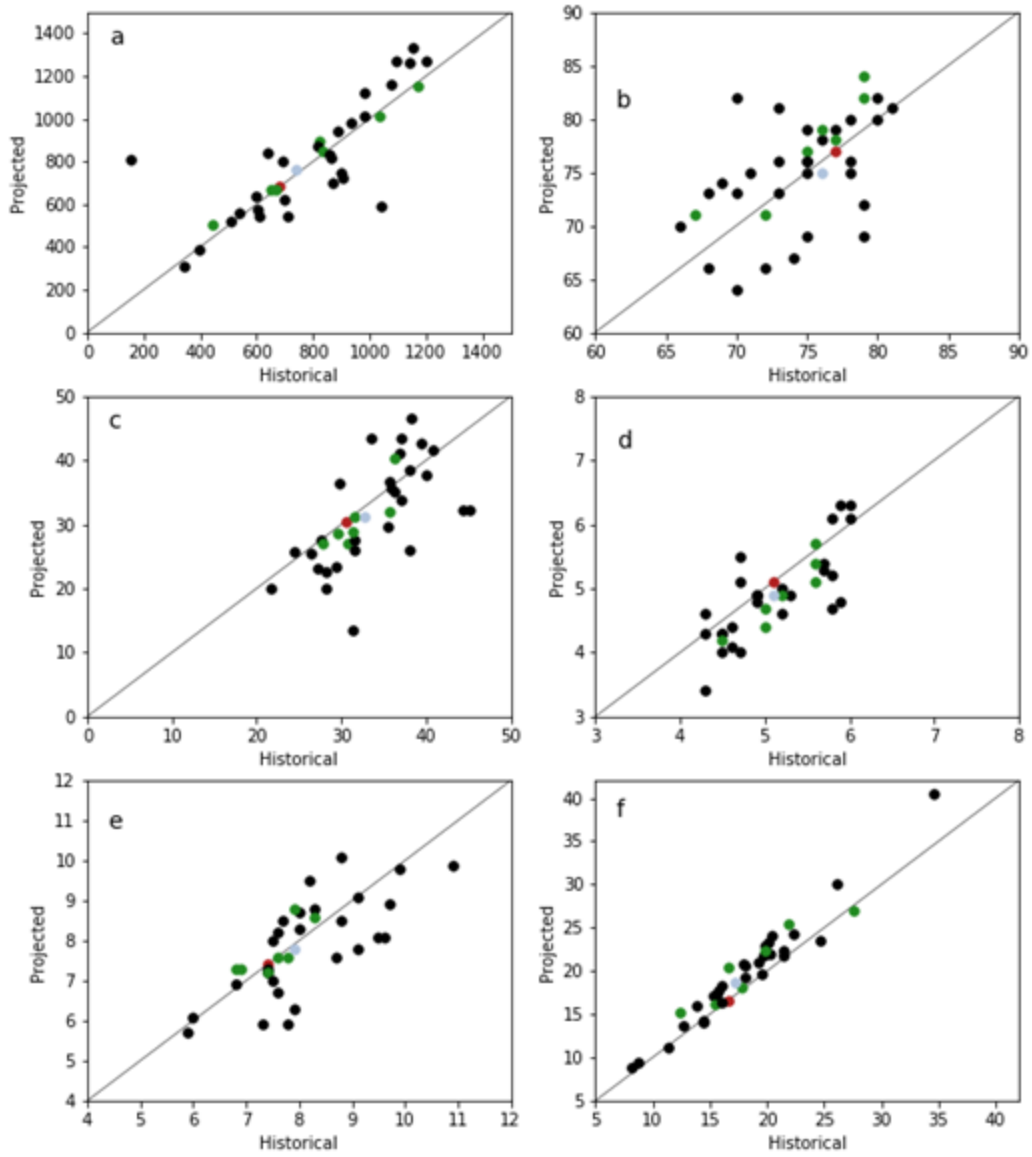

Figure 8. Scatter plots of projected values versus historical values from each of the 35 models: (a) DJFM seasonal rainfall totals (mm), (b) DJFM seasonal burst totals (%), (c) number of burst days (%), (d) number of burst events, (e) burst event duration (days) and (f) burst event intensities (mm/day). The AWAP values (red) are shown on the lines representing a one-to-one correspondence (i.e. no projected change). Projected increases are indicated by values above the line as opposed to projected decreases which fall below the line. The models with better MJO skill indicated are indicated in green and the model mean is shown in light blue.

TABLES

Table 1: Statistics of Darwin (station data) and monsoon domain (AWAP data) summer monsoon DJFM rainfall (1969 to 1999). Shown is the mean value, standard deviation, coefficient of variation (CV) and the correlation of each variable with total DJFM rainfall (T). Monsoon variables are described in Section 2.

		Total Rainfall T (mm)	Total Burst Rainfall T_b (%)	Fraction Burst Days N_b (%)	Number Burst Events N_e(days)	Burst Duration D (days)	Burst Intensity I_b (mm/d)
MEAN	Darwin	1459	86	34	6.0	6.7	29
	AWAP	683	77	31	5.1	7.4	17
STD	Darwin	389	27	9.6	1.9	1.9	5
	AWAP	158	17	14	1.4	4.6	3.9
	(the following statistics are dimensionless)						
CV	Darwin	0.27	0.32	0.23	0.31	0.28	0.17
	AWAP	0.23	0.22	0.35	0.28	0.35	0.23
CORREL	Darwin	1.00	1.00	0.87	0.69	0.22	0.76
	AWAP	1.00	0.91	0.93	0.41	0.71	0.83

TABLES

Table 2: CMIP5 models included in this study. Model names and model host institution summarised from https://pcmdi.llnl.gov/mips/cmip5/availability.html. The ten outlier models in Figure 5 are highlighted in bold while models deemed to provide a good/poor representation of MJO features are indicated by ticks/crosses.

Model name(s)	**MJO skill**	**Model host institution(s)**
1. ACCESS1.0 2. ACCESS1.3		Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Bureau of Meteorology
3. BCC-CSM1.1 **4. BCC-CSM1.1-M**	 X	Beijing Climate Center, China Meteorological Administration
5. BNU-ESM		College of Global Change and Earth System Science, Beijing Normal University
6. CanESM2	X	Canadian Centre for Climate Modelling and Analysis
7. CCSM4		National Center for Atmospheric Research
8. CESM1-BGC **9. CESM1-CAM5**		National Science Foundation, Department of Energy, National Center for Atmospheric Research
10. CMCC-CESM 11. CMCC-CM 12. CMCC-CMS	 ☐ ☐	Centro Euro-Mediterraneo per I Cambiamenti Climatici
13. CNRM-CM5	☐	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique
14. CSIRO-Mk3.6.0		Commonwealth Scientific and Industrial Research Organisation (CSIRO)
15. EC-EARTH		EC-EARTH consortium
16. FGOALS-g2		LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences; and CESS, Tsinghua University
17. FGOALS-s2		LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences
18. GFDL-CM3 19. GFDL-ESM2G 20. GFDL-ESM2M		Geophysical Fluid Dynamics Laboratory
21. HadGEM2-AO 22. HadGEM2-CC 23. HadGEM2-ES	 X X	Met Office Hadley Centre
24. INM-CM4	X	Institute for Numerical Mathematics
25. IPSL-CM5A-LR		Institut Pierre-Simon Laplace

26. IPSL-CM5A-MR 27. IPSL-CM5B-LR	☐	
28. MIROC5	☐	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
29. MIROC-ESM **30. MIROC-ESM-CHEM**	X X	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
31. MPI-ESM-LR 32. MPI-ESM-MR	 ☐	Max Planck Institute for Meteorology (MPI-M)
33. MRI-CGCM3 34. MRI-ESM1	☐	Meteorological Research Institute
35. NorESM1-M		Norwegian Climate Centre

TABLES

Table 3. Statistics of the monsoon domain rainy seasons from the CMIP5 35-model ensemble as well as the reduced N25 ensemble and the MJO-stratified ensembles N7good and N7bad. P10 and P90 is the 10th and 90th percentile within each ensemble and CORREL is the inter-model correlation with total (T) (values significant at the p<.01 level are bolded) within each ensemble.

	Ensemble	T (mm)	TB (%)	NB (%)	NE (days)	D (days)	IB (mm/d)
MEAN	N35	795	604	33	5	8	18
	N25	786	600	34	5	8	18
	N7good	**804**	**598**	32	5	8	19
	N7bad	**683**	**507**	34	5	8	15
P10	N35	470	412	27	4	7	13
	N25	425	386	28	4	7	12
	N7good	**568**	**450**	29	5	7	14
	N7bad	**560**	**420**	26	4	7	12
P90	N35	1124	804	40	6	10	24
	N25	1147	818	39	6	10	22
	N7good	**1091**	**810**	**36**	6	8	24
	N7bad	**884**	**647**	**45**	6	10	19
CORREL	N35	1	**0.89**	0.41	0.60	-0.01	**0.71**
	N25	1	**0.88**	0.47	0.70	-0.10	**0.81**
	N7good	1	**0.99**	0.61	0.72	-0.04	**0.92**
	N7bad	1	**0.97**	0.78	**0.96**	0.43	-0.02

Table 4. Multi-model projected (2069 to 2099 versus 1969 to 1999) changes to northern Australia rainy season total rainfall and burst indices (for the full CMIP5 ensemble and three subsets under the RCP8.5 emission scenario) and correlation between total rainfall change (ΔT) and other indices. The values corresponding to the stratified ensemble (N=25) are shown, as are those corresponding to the 7 models which display realistic MJO-like features and the 7 models which do not (values significant at the p<.01 level are bolded).

	Ensemble	ΔT (%)	ΔTb (%)	ΔNb (%)	ΔNe (days)	ΔD (days)	ΔIb (mm/d)
MEAN	N35	-1	1	-6	-5	-1	8
	N27	-1	1	-6	-6	-1	8
	N7good	**5**	**9**	**0**	**-5**	**5**	**10**
	N7bad	**-7**	**-9**	**-13**	**-7**	**-8**	**5**
P10	N35	-20	-27	-28	-16	-15	-2
	N27	-21	-27	-21	-13	-15	-2
	N7good	-2	1	-11	-10	-2	0
	N7bad	-21	-28	-28	-19	-17	-4
P90	N35	14	20	13	6	11	17
	N27	13	20	12	6	12	17
	N7good	14	21	15	3	13	23
	N7bad	9	13	4	4	0	16
CORREL	N35	1	**0.97**	**0.94**	**0.85**	**0.85**	0.15
	N27	1	**0.97**	**0.95**	**0.83**	**0.86**	0.17
	N7good	1	**1.00**	**0.98**	**0.85**	**0.84**	**0.77**
	N7bad	1	0.96	0.49	0.36	0.62	0.38

TABLES

Table 4. Multi-model projected (2069 to 2099 versus 1969 to 1999) changes to northern Australia rainy season total rainfall and burst indices (for the full CMIP5 ensemble and three subsets under the RCP8.5 emission scenario) and correlation between total rainfall change (ΔT) and other indices. The values corresponding to the stratified ensemble (N=25) are shown, as are those corresponding to the 7 models which display realistic MJO-like features and the 7 models which do not (values significant at the p<.01 level are bolded).

	Ensemble	ΔT (%)	ΔTb (%)	ΔNb (%)	ΔNe (days)	ΔD (days)	ΔIb (mm/d)
MEAN	N35	-1	1	-6	-5	-1	8
	N27	-1	1	-6	-6	-1	8
	N7good	**5**	**9**	**0**	**-5**	**5**	**10**
	N7bad	**-7**	**-9**	**-13**	**-7**	**-8**	**5**
P10	N35	-20	-27	-28	-16	-15	-2
	N27	-21	-27	-21	-13	-15	-2
	N7good	-2	1	-11	-10	-2	0
	N7bad	-21	-28	-28	-19	-17	-4
P90	N35	14	20	13	6	11	17
	N27	13	20	12	6	12	17
	N7good	14	21	15	3	13	23
	N7bad	9	13	4	4	0	16
CORREL	N35	1	**0.97**	**0.94**	**0.85**	**0.85**	0.15
	N27	1	**0.97**	**0.95**	**0.83**	**0.86**	0.17
	N7good	1	**1.00**	**0.98**	**0.85**	**0.84**	**0.77**
	N7bad	1	0.96	0.49	0.36	0.62	0.38