

# An Evaluation of Rainfall Frequency and Intensity over the Australian Region in a Global Climate Model

JOSEPHINE R. BROWN\*

*School of Geography and Environmental Science, Monash University, Melbourne, Victoria, Australia*

CHRISTIAN JAKOB AND JOHN M. HAYNES

*School of Mathematical Sciences, Monash University, Melbourne, Victoria, Australia*

(Manuscript received 14 December 2009, in final form 8 June 2010)

## ABSTRACT

Observed regional rainfall characteristics can be analyzed by examining both the frequency and intensity of different categories of rainfall. A complementary approach is to consider rainfall characteristics associated with regional synoptic regimes. These two approaches are combined here to examine daily rainfall characteristics over the Australian region, providing a target for model simulations. Using gridded daily rainfall data for the period 1997–2007, rainfall at each grid point and averaged over several sites is decomposed into the frequency of rainfall events and the intensity of rainfall associated with each event. Daily sea level pressure is classified using a self-organizing map, and rainfall on corresponding days is assigned to the resulting synoptic regimes. This technique is then used to evaluate rainfall in the new Australian Community Climate and Earth-System Simulator (ACCESS) global climate model and separate the influence of large-scale circulation errors and errors due to the representation of subgrid-scale physical processes. The model exhibits similar biases to many other global climate models, simulating too frequent light rainfall and heavy rainfall of insufficient intensity. These errors are associated with particular synoptic regimes over different sectors of the Australian continent and surrounding oceans. The model simulates only weak convective rainfall over land during the summer monsoon, and heavy rainfall associated with frontal systems over southern Australia is also not simulated. As the model captures the structure and frequency of synoptic patterns, but not the associated rainfall intensity or frequency, it is likely that the source of the rainfall errors lies in model physical parameterizations rather than large-scale dynamics.

## 1. Introduction

The distribution and variability of rainfall are among the most important climate variables for human society, agriculture, and natural ecosystems, but capturing observed rainfall characteristics remains a challenge for global climate models (e.g., Trenberth et al. 2003; Meehl et al. 2005; Randall et al. 2007). Evaluation of rainfall in global climate models is often focused on monthly and interannual time scales (e.g., Randall et al. 2007).

However, the simulation of mean rainfall and monthly to interannual variability requires that models capture the physical processes producing rainfall on the time scale of individual synoptic events (e.g., Trenberth et al. 2003; Dai 2006). Even if models are able to simulate the correct amount of rainfall at a given location, it is also important that this occurs for the right reason: does the model simulate the observed atmospheric circulation and moisture transport, and does the model simulate the dominant local rainfall processes, such as convective or stratiform precipitation? This study outlines an approach to characterizing daily rainfall over the Australian region, using both regional atmospheric circulation and grid-scale rainfall frequency and intensity, to evaluate global climate model performance.

The use of daily rainfall allows the separation into frequency (*fraction of rain days*) and intensity (*rainfall per rain day*) of observed and modeled rainfall (e.g., Dai 2006;

---

\* Current affiliation: Centre for Australian Weather and Climate Research, Melbourne, Victoria, Australia.

---

*Corresponding author address:* Christian Jakob, School of Mathematical Sciences, Monash University, Melbourne, VIC 3800, Australia.  
E-mail: christian.jakob@sci.monash.edu.au

Sun et al. 2006). Sun et al. (2006) apply this approach to a set of 18 coupled climate models, dividing rainfall into light and heavy categories and comparing the simulated rainfall characteristics with daily observations. They find that most models overestimate the frequency of light rainfall and underestimate the intensity of heavy rainfall. Thus, while models may simulate the correct mean rainfall, they may not capture the correct combination of frequency and intensity. The studies of Dai (2006) and Sun et al. (2006) consider the influence of model convective parameterization, but they do not explicitly evaluate the role of synoptic-scale circulation.

Several recent studies have used clustering techniques, such as self-organizing maps (SOMs), to investigate Australian rainfall variability (Hope 2006; Hope et al. 2006; Alexander et al. 2010) using observational records of sea level pressure (SLP) and rainfall. In those studies, SOMs were used to define a set of synoptic regimes or pressure patterns and their associated rainfall distributions, with the goal of identifying drivers of recent rainfall trends. However, these studies did not directly consider rainfall as a function of frequency and intensity, or consider different processes producing light and heavy rainfall. The studies were also limited to rainfall observations over land.

Here, we apply a combination of the approaches described above to better understand rainfall behavior over Australia and the adjacent ocean areas. Including both land and ocean areas in the analysis provides a more complete picture of regional rainfall processes and therefore allows a more comprehensive model evaluation. We carry out an analysis of synoptic regimes using SOMs, while at the same time separating rainfall into light and heavy categories, and considering rainfall frequency and intensity. We then apply our findings to evaluate rainfall characteristics in a simulation with the new Australian Community Climate and Earth-System Simulator (ACCESS; available online at [www.accessimulator.org.au](http://www.accessimulator.org.au)) model.

Section 2 describes the data and methods used. Section 3 outlines the seasonal rainfall climatology and daily rainfall characteristics in both the observations and the model. In section 4, the observed and simulated daily rainfall frequency and intensity are examined for each synoptic regime derived from the SOM analysis. A brief discussion of the results is given in section 5 and the main conclusions of the study are summarized.

## 2. Data and methods

### *a. Rainfall and SLP data*

The study makes use of gridded rainfall datasets from the Global Precipitation Climatology Project (GPCP).

Monthly rainfall was obtained from the GPCP version 2 dataset for the period 1979–2008 at  $2.5^\circ \times 2.5^\circ$  resolution (Adler et al. 2003). Daily rainfall was also obtained from GPCP 1 degree daily (1DD) version 1.1 dataset for the period 1997–2007 at  $1^\circ \times 1^\circ$  resolution (Huffman et al. 2001). The GPCP gridded rainfall datasets are constructed from merged satellite and rain gauge data, providing an estimate of rainfall amount over regions with sparse direct observations. As global, gridded datasets, GPCP rainfall products have a spatial scale appropriate for climate model evaluation. The GPCP datasets can be readily compared with gridded model rainfall with a similar grid size, whereas station data are not suitable for direct comparison. The GPCP rainfall is referred to in this study as “observed” rainfall, while it is noted that the GPCP dataset is only partially derived from direct observations.

Monthly and daily gridded sea level pressure data were obtained from the National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). These data are available for the period 1948–present; however, in this study, only data from the period 1979–2008 are used to allow comparison with the model simulation. Using only data from 1979 onward also excludes biases because of a lack of satellite data before the late 1970s.

### *b. Model simulation*

A 22-yr simulation from 1979 to 2000 was carried out with the atmospheric component of the ACCESS climate model. The model version used is based on the Hadley Centre Global Environmental Model version 2 (HadGEM2; Pope et al. 2007) with some modifications, including the use of the prognostic condensate and prognostic cloud fraction scheme (PC2; Wilson et al. 2008; see [http://www.hpsc.csiro.au/users/dix043/access\\_amip/](http://www.hpsc.csiro.au/users/dix043/access_amip/) for details of model version and boundary conditions). The horizontal resolution used is N96 ( $1.875^\circ \times 1.25^\circ$ ), with 38 vertical levels.

The experimental design followed the protocol of the Atmospheric Modeling Intercomparison Project Phase II (AMIP-II), with sea surface temperature (SST) and sea ice boundary conditions prescribed from the AMIP-II forcing data (Taylor et al. 2000). The experiment includes the full aerosol scheme with sulfate, soot, biomass, and mineral dust. Carbon dioxide, methane,  $N_2O$ , and chlorofluorocarbons (CFCs) were set at average values for the AMIP period. The  $CO_2$  concentration is 348 ppm, the methane concentration is 1650 ppb, and  $N_2O$  concentration is 306 ppb. The model uses the Met Office Surface Exchange Scheme (MOSES) with nine surface types.

### c. Regime analysis using SOMs

The study uses the technique of SOMs (Kohonen 2001) applied to daily sea level pressure maps over Australia and the adjacent oceans to investigate regimes of atmospheric circulation and associated rainfall. The SOM method is an artificial neural network algorithm that can be used to visualize and cluster large numbers of observations into a specified number of patterns or “nodes” to allow easier interpretation. The SOM algorithm matches each input vector to a reference vector (node) based on the minimum Euclidean distance between them. The matching node and those adjacent to it are then updated, leading to an array of nodes that span the input vectors in an ordered manner (e.g., Cassano et al. 2006; Hope et al. 2006). When creating the SOM, all data are assigned to a node, in contrast with some techniques where only a subset of the data may be used.

SOMs have been used in a wide range of climate applications, including classifying satellite images (Richardson et al. 2003), investigating rainfall extremes and trends (e.g., Cavazos 2000; Cassano et al. 2006; Hope et al. 2006; Uotila et al. 2007), and mapping variability of ENSO (Leloup et al. 2007) and the Indian Ocean dipole (Tozuka et al. 2008).

SOMs can be used to map climate variables, such as SLP, to construct patterns corresponding to synoptic types or regimes (e.g., Hewitson and Crane 2002; Tennant 2003). The SOM method preserves the probability density of the input data and produces a set of approximately equally probable nodes. The nodes are arranged with similar types located adjacent to each other in the two-dimensional ordered array, facilitating interpretation of the synoptic types and their relationship. The dominant synoptic patterns influencing a subregion can also be investigated (e.g., Schuenemann et al. 2009). A more detailed description of the SOM method as applied to synoptic classification is given in Cassano et al. (2006).

The SOM method is applied here to daily SLP patterns over the Australian region ( $0^{\circ}$ – $50^{\circ}$ S,  $110^{\circ}$ – $160^{\circ}$ E). The use of daily model and observed fields to construct SOMs allows us to identify patterns of circulation on synoptic time scales and investigate the corresponding daily rainfall variability. As the allocation of each daily model or observed state to a SOM node occurs in postprocessing, the SOM can be constructed from observed (NCEP–NCAR reanalysis) SLP and then model SLP can be mapped to the same SOM nodes. To facilitate model–data comparison, the SOM is constructed from observed SLP from the period 1979–2008, incorporating the period of the model simulation but also overlapping with the daily GPCP rainfall data period of 1997–2007.

A number of rainfall characteristics are examined. The *rainfall frequency* is the number of “rain days” (days with more than 1 mm of rainfall) divided by the total number of days. The *rainfall intensity* is the average rainfall amount per rain day, while the *mean rainfall* is the average rainfall amount over all days.

The decomposition of rainfall frequency and intensity is combined with clustering using SOMs to explore the model errors as a function of synoptic regime. The total rainfall over a given period ( $R_{\text{total}}$ ) can be considered as the sum over regimes with occurrence frequency  $F_i$  and mean regime rainfall  $R_i$ :

$$\begin{aligned} R_{\text{total}} &= \sum_{i=1}^{20} (F_i \times R_i) \\ &= \sum_{i=1}^{20} (F_i \times Q_i \times N_i). \end{aligned} \quad (1)$$

The mean rainfall for a given regime ( $R_i$ ) is the product of the rainfall frequency ( $Q_i$ ) for the regime multiplied by the rainfall intensity ( $N_i$ ) for the regime. Therefore, an error in total rainfall over a region can be considered to be the sum of the products of the error in the frequency of occurrence of the regime ( $F$ ), the error in the regime rainfall frequency ( $Q$ ), and the error in the regime rainfall intensity ( $N$ ). Synoptic regime frequency is a property of the large-scale atmospheric circulation, whereas the rainfall frequency and intensity associated with each regime may be associated with local or subgrid-scale processes. Using this approach, we aim to determine whether the model correctly represents both large-scale circulation and local rainfall processes for the region of interest.

## 3. Rainfall climatology

### a. Seasonal rainfall climatology

Before considering the daily rainfall characteristics, we examine the seasonal climatology of global rainfall, making use of the longer 22-yr record of monthly data. The difference between the seasonal average rainfall in the ACCESS AMIP simulation and the GPCP seasonal rainfall for the same period (1979–2000) is shown in Fig. 1. The GPCP seasonal rainfall climatologies for the periods 1979–2000 and 1997–2007 were also compared (not shown) to determine whether the use of a different period for GPCP daily data and ACCESS daily data will influence the results. The differences between the two periods of GPCP rainfall are much smaller than the difference between GPCP and ACCESS rainfall (less than  $1 \text{ mm day}^{-1}$ ).

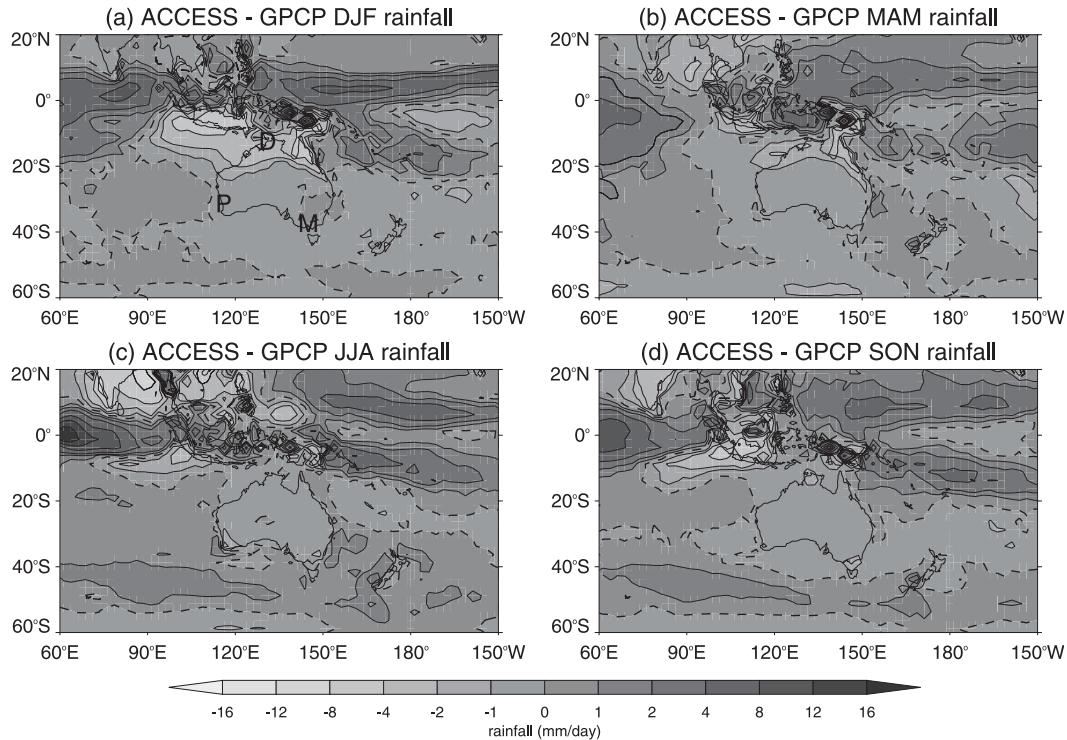


FIG. 1. Difference between seasonal rainfall climatology ( $\text{mm day}^{-1}$ ) from ACCESS AMIP simulation and GPCP monthly rainfall data (1979–2000). The locations of Perth (P), Melbourne (M), and Darwin (D) are indicated in (a). The dashed line indicates the zero contour.

Errors in model seasonal average rainfall are largest in the tropics, with too much rainfall over the western tropical Pacific and in the region of the South Pacific convergence zone (SPCZ). In June–August (JJA), the model Indian summer monsoon is too dry, with a rainfall maximum located over the tropical Indian Ocean rather than over the continent, a known bias in the HadGEM family of models (Martin et al. 2004). In December–February (DJF), the model simulates the intertropical convergence zone (ITCZ) located too far north over the Indonesian Maritime Continent region, resulting in too little summer monsoon rainfall over northern Australia. Over the Australian continent, the largest errors are seen over northern Australia in summer. Rainfall errors over southern Australia are less than  $1 \text{ mm day}^{-1}$  for all seasons, with the exception of larger dry biases of up to  $2 \text{ mm day}^{-1}$  over southwest Western Australia and southeast Australia during winter and spring, when the majority of rainfall is received. To explore the source of these biases, we now extend our investigation to daily time scales.

### b. Rainfall frequency and intensity

Daily rainfall characteristics can be analyzed as a combination of the frequency and the intensity of

rainfall events. The frequency and intensity of rainfall can also be separated into categories such as “light” ( $1\text{--}10 \text{ mm day}^{-1}$ ) and “heavy” ( $>10 \text{ mm day}^{-1}$ ), which may be considered to represent different classes of rainfall (Sun et al. 2006). Previous studies have found that climate models often overestimate the frequency of light rainfall and underestimate the intensity of heavy rainfall (Sun et al. 2006). We examine ACCESS and observed GPCP rainfall frequency over the Australian region in Southern Hemisphere summer (DJF) and winter (JJA) to determine whether the model is able to capture both aspects of rainfall variability. Daily model rainfall for the 22-yr AMIP simulation (1979–2000) is compared with daily GPCP precipitation data for the 11-yr period from 1997 to 2007. The two different periods are used to maximize the length of the daily rainfall record. Comparison of the shorter, overlapping period from 1997 to 2000 demonstrated that the rainfall characteristics were not dependent on the period analyzed (not shown).

Figure 2 shows the mean daily frequency of light and heavy rainfall over the Australian region for DJF and JJA. Following Sun et al. (2006), rainfall less than  $1 \text{ mm day}^{-1}$  is excluded, as drizzle will generally contribute little to rain gauge observations. We note that a slight discontinuity is visible in the GPCP rainfall data

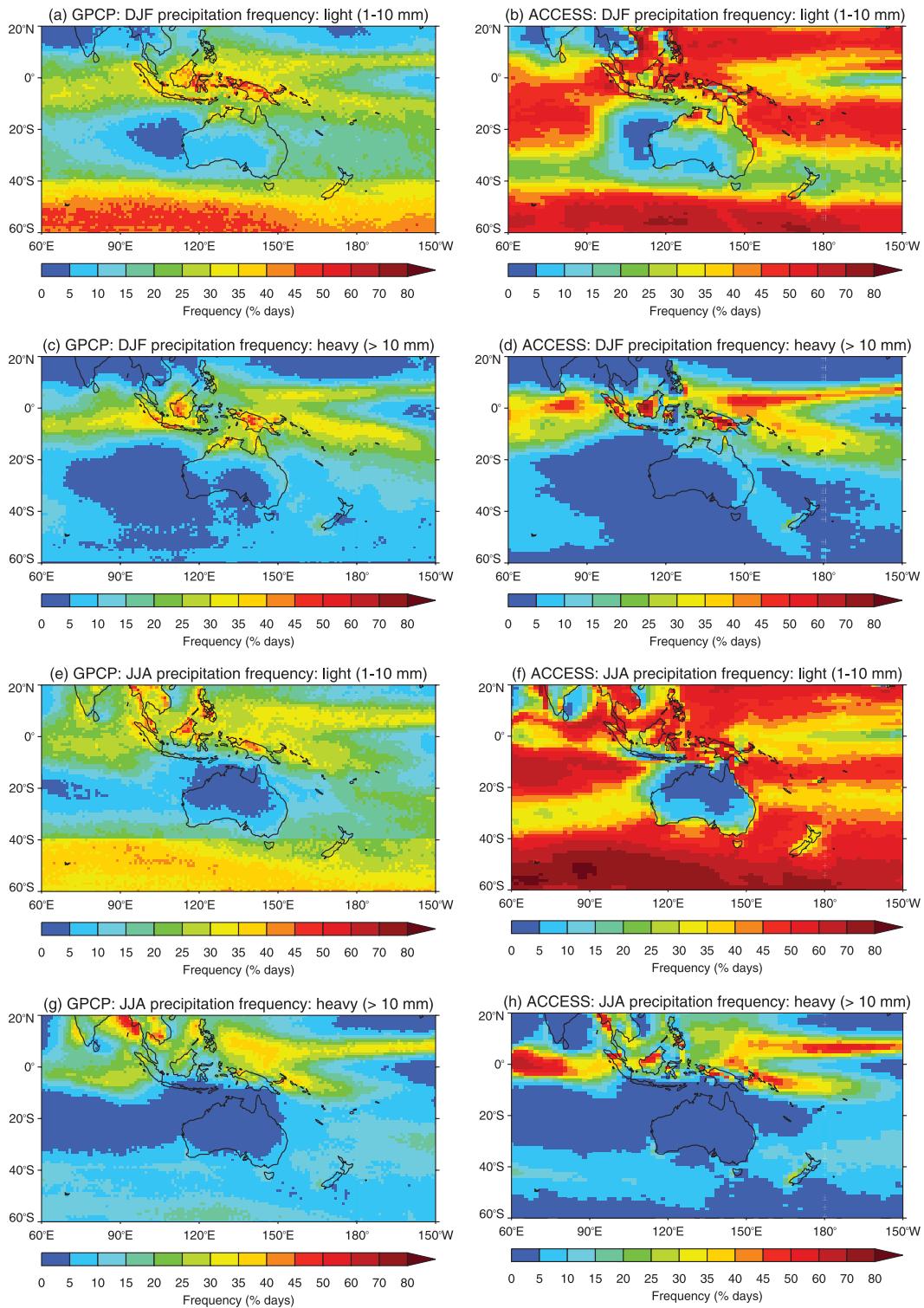


FIG. 2. Mean daily rainfall frequency (% days) at each grid point: (left) GPCP rainfall (1997–2007), (right) ACCESS rainfall (1979–2000), (top four plots) DJF light and heavy rainfall, and (bottom four plots) JJA light and heavy rainfall.

at 40°S (e.g., Fig. 2a) because of the use of different data sources north and south of this latitude (Huffman et al. 2001); however, this discontinuity does not alter the results. In DJF, the model simulates more frequent light rainfall than observed in the tropics, including over northern Australia, and also over the Southern Ocean. The frequency of heavy rainfall is in closer agreement between model and observations, with the most frequent heavy rainfall occurring within the tropical convergence zones. A notable exception is the underestimation of heavy rainfall over northern Australia. In JJA, the model simulates noticeably more frequent light rainfall than observed except over continental Australia, where the arid conditions are well captured. The frequency of heavy rainfall in winter is largely consistent with the observed distribution, although biases over the tropical Indian Ocean reflect mean rainfall errors. Heavy rainfall is underestimated over southern Australia in winter, with the model simulating a low frequency (<5%) of days with heavy rainfall over land.

Figure 3 shows the mean daily intensity of light and heavy rainfall over the Australian region for DJF and JJA. Rainfall of less than 1 mm day<sup>-1</sup> is again excluded, and grid points with no rainfall are shown in white. The spatial distribution of the intensity of light rainfall (1–10 mm day<sup>-1</sup>) is fairly uniform in the GPCP dataset, with most light rainfall occurring at the rate of 4–5 mm day<sup>-1</sup> in both DJF and JJA. In contrast, the model simulates intensities of 5–7 mm day<sup>-1</sup> in the tropical convergence zones, with the most intense rainfall over the islands of the Maritime Continent, and much lighter rainfall over the extratropics, especially southern Australia in DJF and northern Australia in JJA. The model's spatial distribution of heavy rainfall (>10 mm day<sup>-1</sup>) is in closer agreement with observations, although the model significantly underestimates the amount of heavy rainfall over land, including the Indian and Australian summer monsoons and winter rainfall over southern Australia.

### c. Rainfall at Perth, Melbourne, and Darwin

As the three regions of largest rainfall error over Australia correspond to the locations of the cities of Perth, Melbourne, and Darwin (shown as “P,” “M,” and “D,” respectively, in Fig. 1a), a more detailed analysis of daily rainfall variability is now carried out comparing model rainfall in the region around each city with GPCP daily rainfall. Note that while city names are used for orientation, the analysis represents larger regions around the cities themselves that are typical for southwestern, southeastern, and northern Australian rainfall processes. The nearest four land grid boxes to the cities of Perth

(32°S, 116°E), Melbourne (38°S, 145°E), and Darwin (12°S, 131°E) are used to calculate a time series of daily rainfall. While this time series will not correspond to rainfall at a given station, it is more appropriate for comparison between the two gridded datasets to use the average of several grid boxes. The 1° × 1° GPCP rainfall data was interpolated to the 1.875° × 1.25° ACCESS model grid for the calculation of rainfall time series, so model and observed rainfall time series represent the same area.

A box-and-whisker diagram for the daily rainfall distribution is shown for Perth (Fig. 4a), Melbourne (Fig. 4b), and Darwin (Fig. 4c) for all days with rainfall exceeding 1 mm. The upper and lower limits of the whiskers are the 95th and 5th percentiles, respectively. The median and interquartile ranges are also shown, and rain day frequency (percent of days with rainfall >1 mm) is indicated above each box. Seasonal distributions include all days in DJF, March–May (MAM), JJA, and September–November (SON), and all months (ALL). The number of days in the model is less than the observations, as the model uses a 360-day year (each month is 30 days).

Perth is located in the southwest of the Australian continent, receiving the majority of its annual rainfall in winter (JJA) and very little rainfall in summer (DJF). The main sources of winter rainfall are prefrontal air masses within the dominant westerly airstream (e.g., Hope et al. 2006). The model underestimates the rainfall for all seasons except summer, and it fails to simulate the strong observed seasonal cycle of rainfall at Perth. The GPCP Perth-area mean rainfall in JJA is 2.2 and the intensity is 9.3 mm day<sup>-1</sup>. In contrast, the model simulates JJA mean rainfall of 1.7 and intensity of 4.1 mm day<sup>-1</sup>, failing to reproduce the extreme winter rainfall events of above 20 mm day<sup>-1</sup> seen in the GPCP record.

Melbourne is located in southeast Australia, a region that receives moderate rainfall in all months, with larger totals in winter and spring. Southeast Australian rainfall can result from both frontal and cut-off low pressure systems in these seasons (e.g., Pook et al. 2006; Risbey et al. 2009). The model underestimates the rainfall for all seasons and simulates more rainfall in summer than winter, in contrast to the observations. The model fails to simulate the high-intensity rainfall events that contribute a large component of observed winter rainfall (JJA model intensity is 3.4 mm day<sup>-1</sup>, compared to observed intensity of 9.6 mm day<sup>-1</sup>). While the intensity of model rainfall is too low in MAM and JJA, the frequency is higher than observed in these seasons.

Darwin is located in the “Top End” of northern Australia and experiences the direct influence of the Australian summer monsoon system, with the majority of

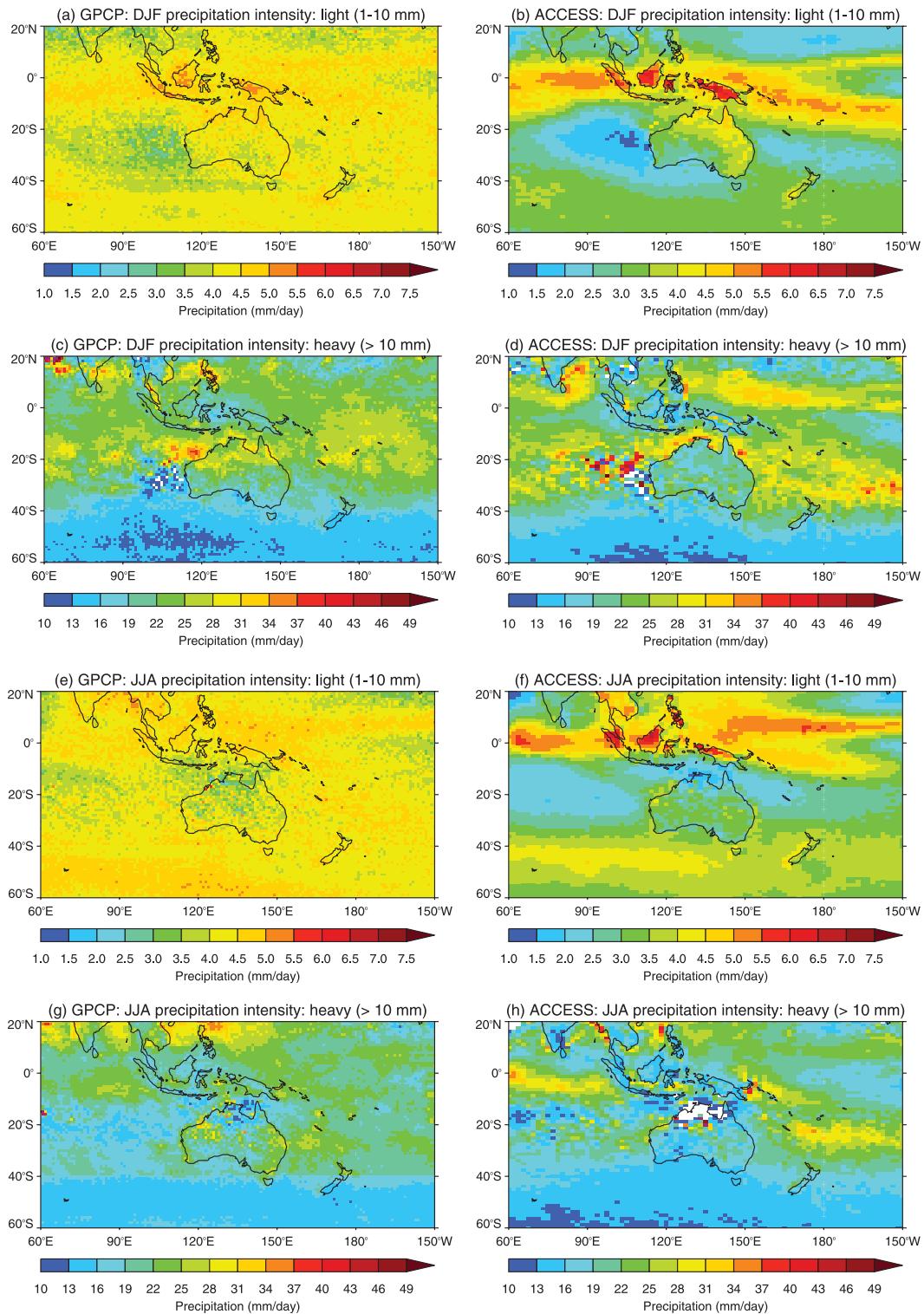


FIG. 3. As in Fig. 2, but for mean daily rainfall intensity. White areas indicate grid points with no heavy rainfall.

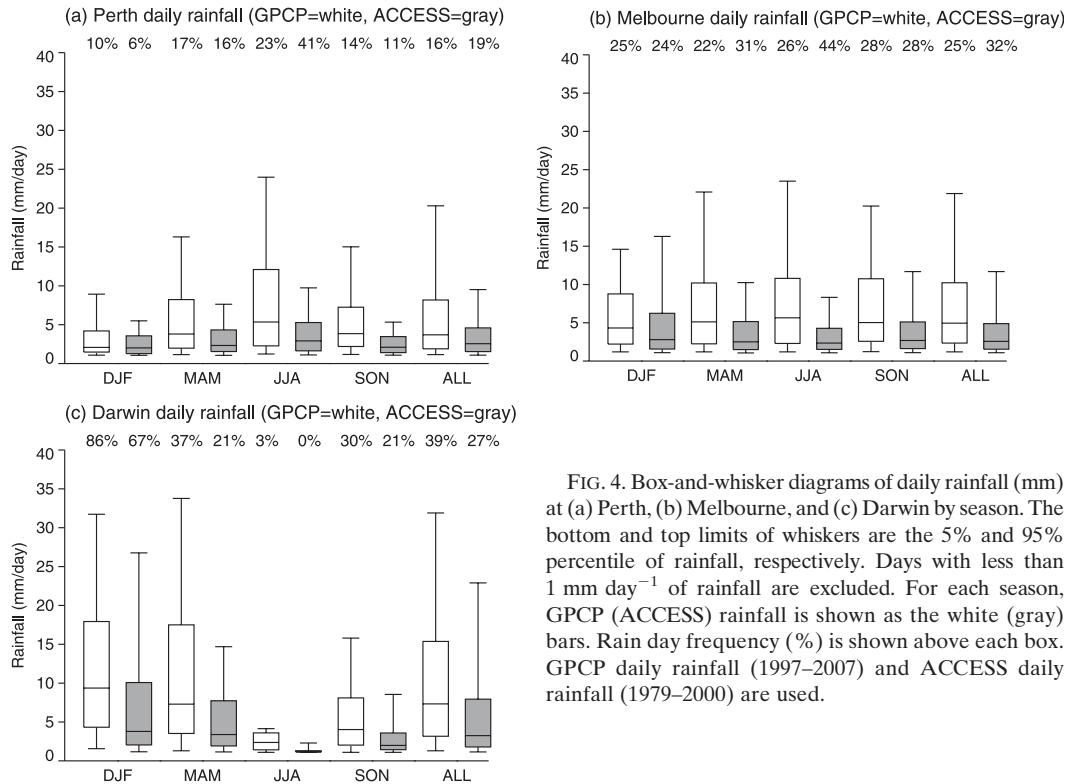


FIG. 4. Box-and-whisker diagrams of daily rainfall (mm) at (a) Perth, (b) Melbourne, and (c) Darwin by season. The bottom and top limits of whiskers are the 5% and 95% percentile of rainfall, respectively. Days with less than  $1 \text{ mm day}^{-1}$  of rainfall are excluded. For each season, GPCP (ACCESS) rainfall is shown as the white (gray) bars. Rain day frequency (%) is shown above each box. GPCP daily rainfall (1997–2007) and ACCESS daily rainfall (1979–2000) are used.

rainfall in the months from December to March and little rainfall during the dry season, corresponding to austral winter. The model is able to capture the observed strong seasonal cycle at Darwin, with the largest rainfall in summer and less than  $0.1 \text{ mm day}^{-1}$  mean rainfall in winter. However, the model underestimates the mean summer rainfall at Darwin ( $5.4 \text{ mm day}^{-1}$  compared to the observed value of  $10.8 \text{ mm day}^{-1}$ ), predominantly due to the lower intensity of model rainfall ( $8.1 \text{ mm day}^{-1}$  compared to  $12.5 \text{ mm day}^{-1}$  observed), as well as slightly lower rainfall frequency in the model compared with observations.

Having established the overall model errors in frequency and intensity of rainfall, an immediate question that arises is whether these errors are caused by the inability of the model to simulate the observed synoptic conditions over the regions of interest or if the model is unable to translate the correct large-scale conditions into correct rainfall behavior. We therefore extend our above analysis of frequency and intensity by studying rainfall behavior as a function of synoptic regime, with the aim to provide some further insight into the reasons for the model errors. Using this approach, we can determine whether the model captures the frequency of the main synoptic regimes, as well as the intensity and frequency of rainfall associated with each regime.

## 4. Synoptic regimes

### a. SOM

A SOM was constructed using gridded daily-average SLP data over the Australian region ( $0^{\circ}$ – $50^{\circ}\text{S}$ ,  $110^{\circ}$ – $160^{\circ}\text{E}$ ) from the NCEP–NCAR reanalysis for the period 1979–2008. A set of 20 SOM patterns ( $4 \times 5$ ) was chosen to span the range of synoptic conditions for all seasons. The choice of the number of SOM patterns is subjective; however, previous studies have found that this number is appropriate to capture the synoptic patterns influencing the Australian region (e.g., Hope et al. 2006; Nicholls et al. 2009; Alexander et al. 2010). It is important to note that the SOM will cover the probability space of SLP states; however, it may not represent particular conditions with a variable location, such as cut-off low pressure systems (Nicholls et al. 2009).

The 20 SOM patterns or nodes constructed from the NCEP SLP data are shown in Fig. 5. The frequency of occurrence ( $F$ ) of each SOM node is indicated above the plots. The nodes are arranged with similar patterns closer together and the most different patterns in opposite corners as a consequence of the SOM mapping algorithm (Hewitson and Crane 2002). In the top-left corner of the SOM are patterns with a strong trough to the south of the continent, while the patterns in the bottom-right corner

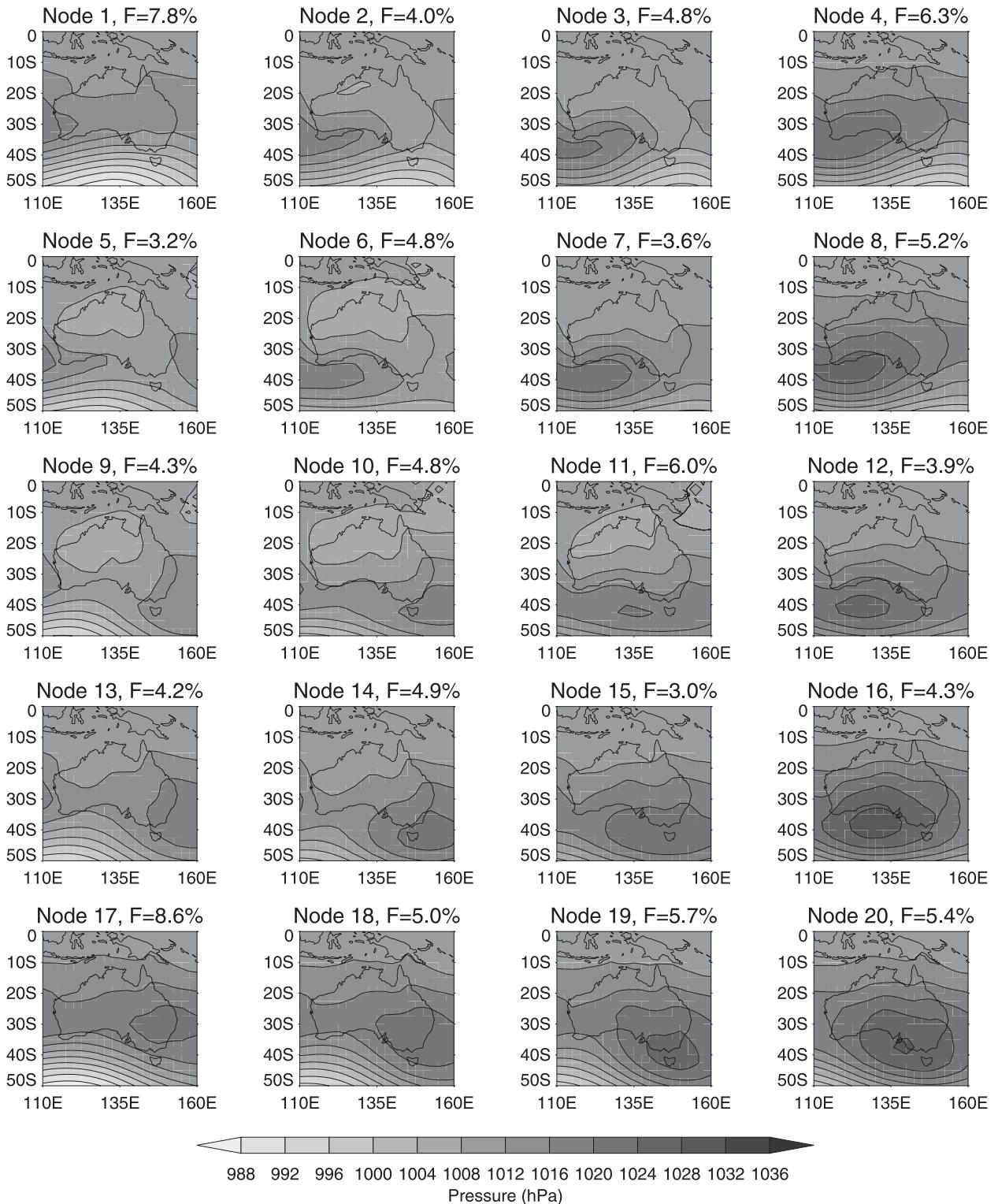


FIG. 5. SOM of SLP (hPa) over the Australian region ( $0^{\circ}$ – $50^{\circ}$ S,  $110^{\circ}$ – $160^{\circ}$ E) constructed using daily average NCEP SLP data from 1979 to 2008; frequency of occurrence ( $F$ ) is shown in percent.

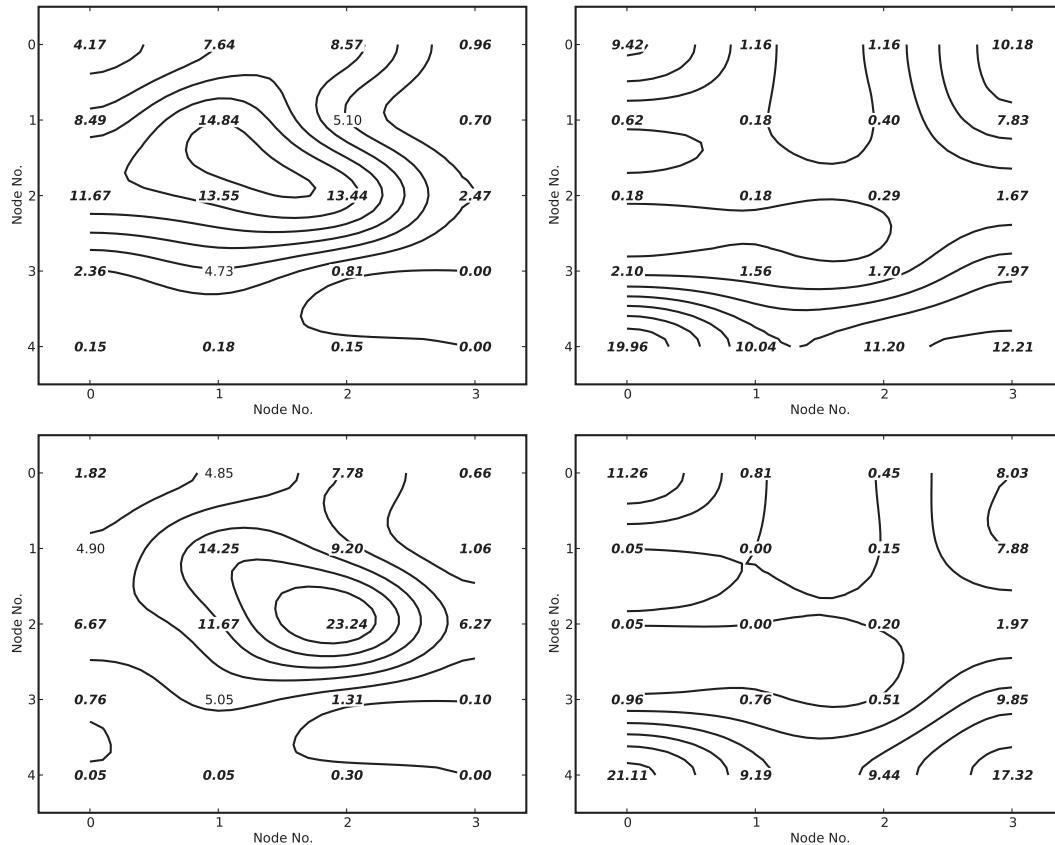


FIG. 6. Frequency map (% days) of SOM nodes for (top left) DJF and (top right) JJA for SOM derived from NCEP SLP from 1979 to 2008, and mapping of ACCESS (1979–2000) daily SLP to SOM nodes for (bottom left) DJF and (bottom right) JJA.

are dominated by a high pressure system centered over southern Australia. The adjacent nodes appear to capture the progression of low pressure systems and fronts to the south of the continent, as well as the intensification of high pressure systems and the development of the continental heat low over the monsoon tropics. As SLP variability is larger in extratropics than the tropics, the SOM nodes are grouped according to midlatitude SLP patterns, while tropical SLP patterns do not always vary smoothly, for example, between nodes 1 and 5.

The frequency of each pattern can be determined for the whole period of the SLP data or for individual seasons. Each daily model SLP field is also mapped onto one of these 20 nodes using its Euclidean distance from the node-average SLP field. The observed (upper) and modeled (lower) frequencies (% days) of the SOM nodes for DJF (left) and JJA (right) are shown in Fig. 6. The dominant nodes in summer (DJF) are in the top-left region of the SOM, corresponding to patterns with a developed heat low over the north of Australia. The dominant nodes in winter (JJA) are those along the

lower edge of the SOM, and particularly the bottom-left corner, as well as the top-left and top-right corners (but not surrounding nodes). The most common winter patterns are those with a trough to the southwest of the Australian continent (see Fig. 5), although states with a trough to the southeast also occur for close to 20% of days. The ACCESS daily SLP fields map to a similar pattern of nodes for summer and winter as the observed daily SLP, with a slightly different frequency of occurrence for some nodes.

As each day of the observed or model simulation period can be mapped to the closest SOM node, this mapping can be used to construct a composite of rainfall for each SOM node (e.g., Hope et al. 2006). The daily NCEP SLP states are allocated to 1 of the 20 nodes for the period 1997–2007, and the corresponding rainfall is composited from GPCP daily rainfall data. The model SLP is also allocated to 1 of the 20 SOM nodes for each day of the AMIP simulation (1979–2000), and a rainfall composite is constructed. The observed and model rainfall anomaly composites for the SOM nodes are

shown in Figs. 7 and 8. As the SLP patterns are dominated by the larger gradients in the extratropics while rainfall anomalies are largest in the tropics, some nodes have rainfall anomaly patterns that are not closely associated with their neighboring nodes, such as node 1. To obtain a more consistent set of regimes for the tropical region, it would be necessary to construct regimes based on alternative variables such as wind.

The ACCESS rainfall anomaly composites (Fig. 8) show that the model captures many of the same large-scale patterns of rainfall for each SOM node compared with the observed GPCP rainfall anomalies (Fig. 7). However, the model rainfall anomalies are generally of smaller magnitude, despite similar SLP composites (not shown) for each node. In particular, the heavy rainfall over northern Australia seen in nodes 3, 5, 6, 7, 9, 10, and 11 (“summer monsoon” nodes) is substantially reduced in the model composites and is located further north, indicating the monsoon does not penetrate far enough inland.

Nodes 1–4 represent pressure troughs bringing winter rainfall to southeast Australia (“wet southeast” nodes), while nodes 17–20 (“wet southwest” nodes) represent a trough or front to the southwest of the continent, bringing winter rainfall to southwest Australia. The model captures the meridional and zonal gradients of rainfall associated with these nodes, but it does not capture the amount of rainfall over land in both southeast and southwest Australia.

### *b. SOMs and frequency and intensity of rainfall*

The allocation of observed and model simulation days to SOM nodes is now used to decompose daily rainfall into frequency and intensity associated with each synoptic regime. In Fig. 9, the GPCP rainfall frequency for each SOM node is shown, while the model rainfall frequency by node is shown in Fig. 10. There is broad agreement between the observed and simulated spatial patterns of rainfall frequency over the Australian region, although differences are larger in the tropics and southern midlatitudes (see Fig. 2). In both observed and model rainfall, the highest frequency occurs over the tropics, including northern Australia, and over the Southern Ocean. The spatial patterns in the midlatitudes follow the SLP patterns in Fig. 5, with highest rainfall frequency over regions of low pressure. Over northern Australia, the model simulates frequent rainfall over land but not over adjacent ocean for the summer monsoon nodes (e.g., nodes 6 and 10), while the frequency over southeast and southwest Australia is also higher in the model than observed for the wet southeast and wet southwest nodes.

In Fig. 11, the GPCP rainfall intensity by SOM node is shown. It is evident that the most intense rainfall (heavier than  $20 \text{ mm day}^{-1}$ ) occurs over northern Australia in SOM nodes 3, 5, 6, 7, 9, 10, and 11 (summer monsoon). Relatively heavy rainfall over southeast Australia occurs in SOM nodes 1–4 (wet southeast), while the heaviest rainfall over southwest Australia occurs in SOM nodes 17–20 (wet southwest). In comparison, the model rainfall intensity by SOM node (Fig. 12) displays heavy rainfall in the summer monsoon nodes only over ocean and the islands of Papua and Indonesia but not over northern Australia. Note that this is in contrast to the model rainfall frequency in those nodes, which was highest over land. These errors are consistent with the seasonal mean rainfall errors over northern Australia and at the site of Darwin, as discussed in section 3. Rainfall intensity over southern Australia is uniformly light, with the wet southeast and wet southwest nodes failing to produce rainfall more than  $10 \text{ mm day}^{-1}$ , also consistent with the findings in section 3.

Sharp gradients in rainfall intensity at the edge of continents suggest a differentiation between maritime and continental rainfall regimes in the model. These sharp gradients are not seen in the observed GPCP rainfall, implying that the model may have an unrealistic division between rainfall triggers or processes over land and over ocean. The analysis approach used here reveals this bias is most extreme in cases of heavy rainfall over northern Australia during active summer monsoon regimes, and is accompanied by a sharp distinction in rainfall frequency between land and ocean, with more frequent (light) rain as well as a lack of heavy rain over land. This information provides useful guidance for model development, as it indicates an unrealistically strong dependence of rainfall on surface type and/or a possible lack of the influence of coastal effects on rainfall in the model.

### *c. Synoptic decomposition of rainfall at Perth, Melbourne, and Darwin*

Having successfully applied the SOM decomposition of SLP to the Australian region, we now make use of it to separate rainfall characteristics at a given location into components because of rainfall frequency and rainfall intensity associated with each node. The main purpose of doing so is to determine whether errors in model-simulated rainfall are due to the incorrect frequency of occurrence of a particular synoptic pattern or the incorrect rainfall intensity for this pattern. We describe here a comparison of the rainfall frequency and rainfall intensity at the locations of Perth, Melbourne, and Darwin for the 20 SOM nodes. The analysis compares daily rainfall for all seasons from the ACCESS simulation

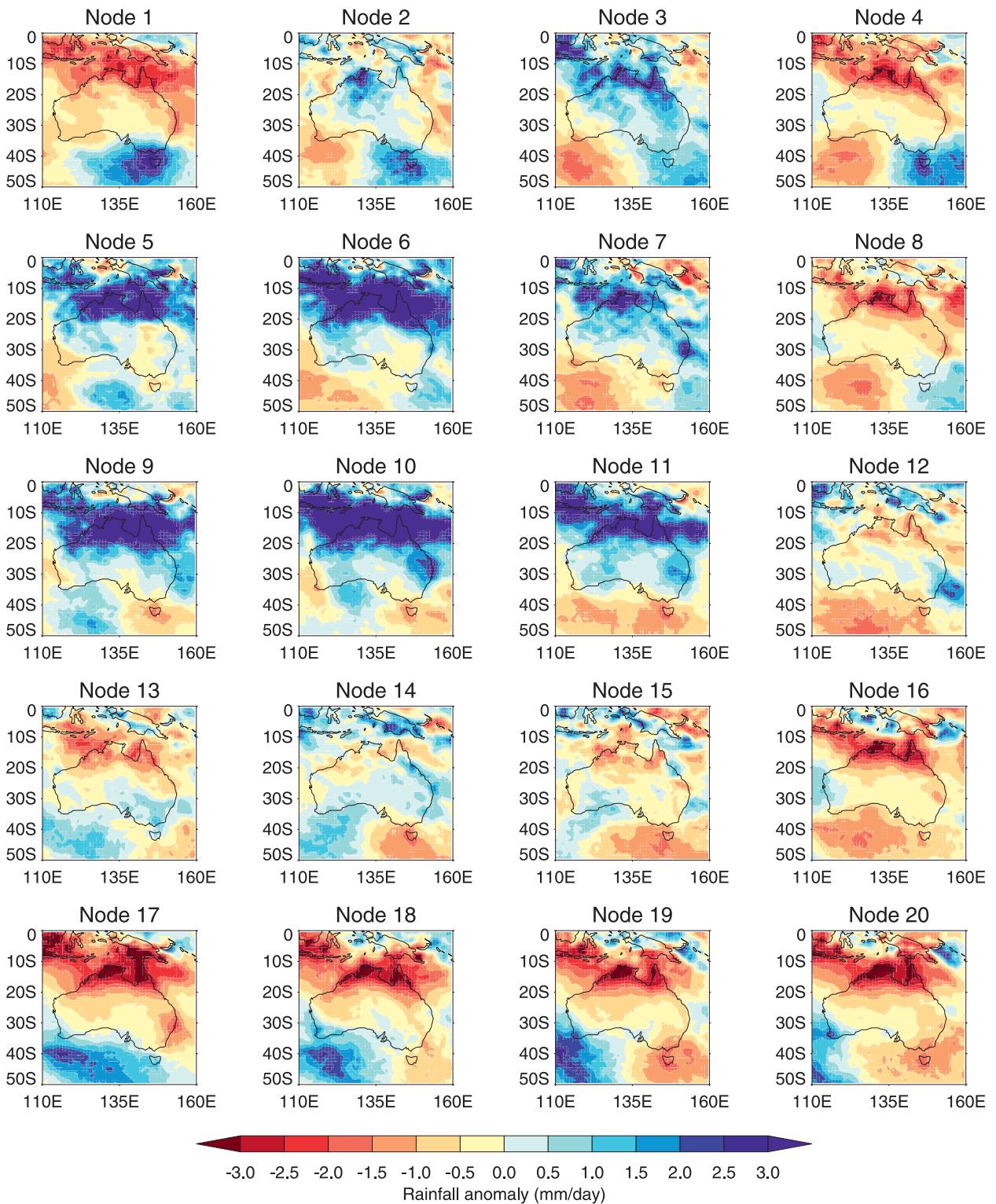


FIG. 7. Composite rainfall anomalies (mm day<sup>-1</sup>) for each SOM node derived from NCEP SLP: GPCP 1DD rainfall from 1997 to 2007.

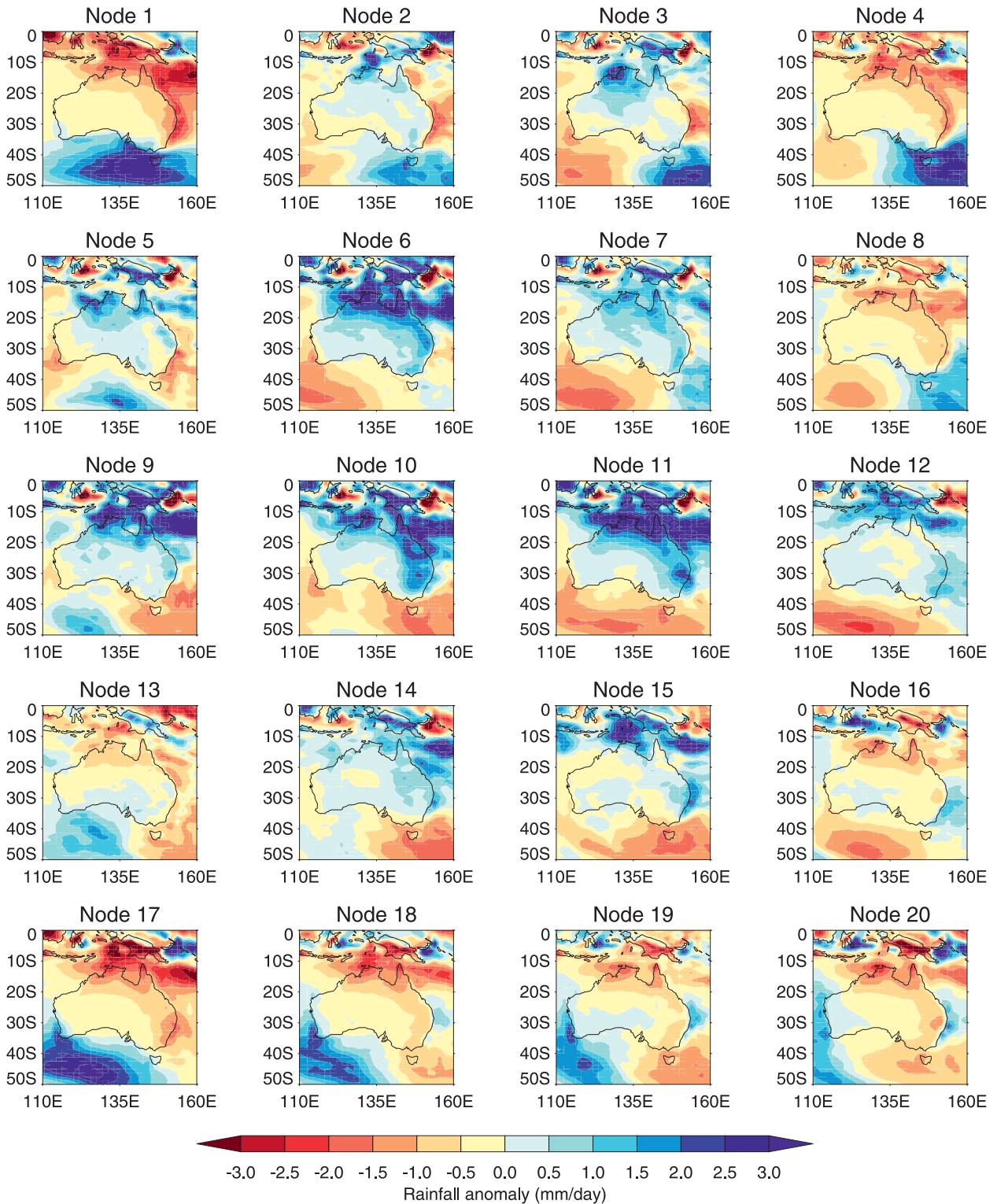


FIG. 8. Composite rainfall anomalies (mm day<sup>-1</sup>) for each SOM node derived from NCEP SLP: ACCESS AMIP simulation daily rainfall from 1979 to 2000 (daily model SLP is allocated to a SOM node, and rainfall composites constructed based on this allocation).

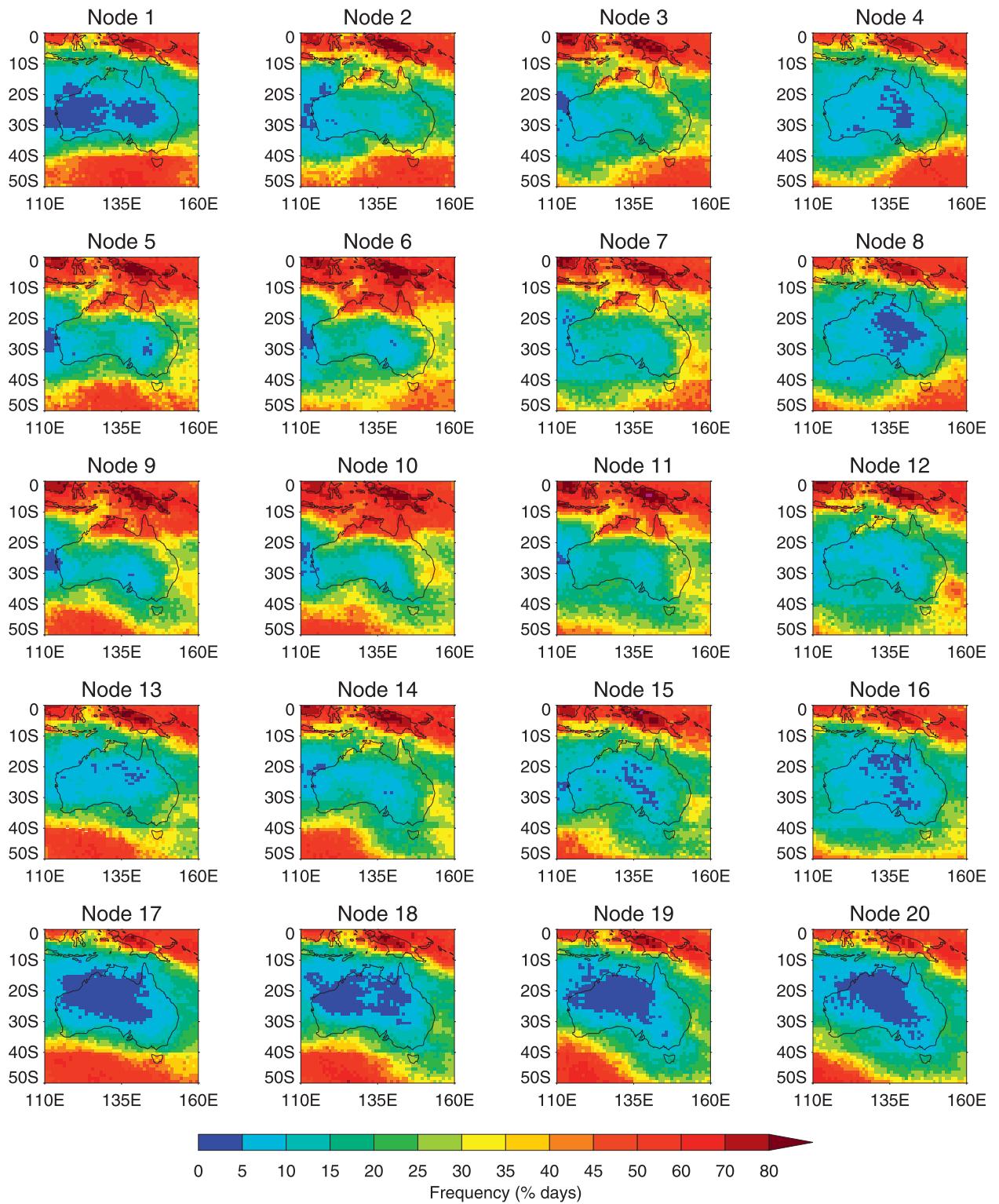


FIG. 9. Rainfall frequency (% days) by SOM node: GPCP 1DD rainfall from 1997 to 2007.

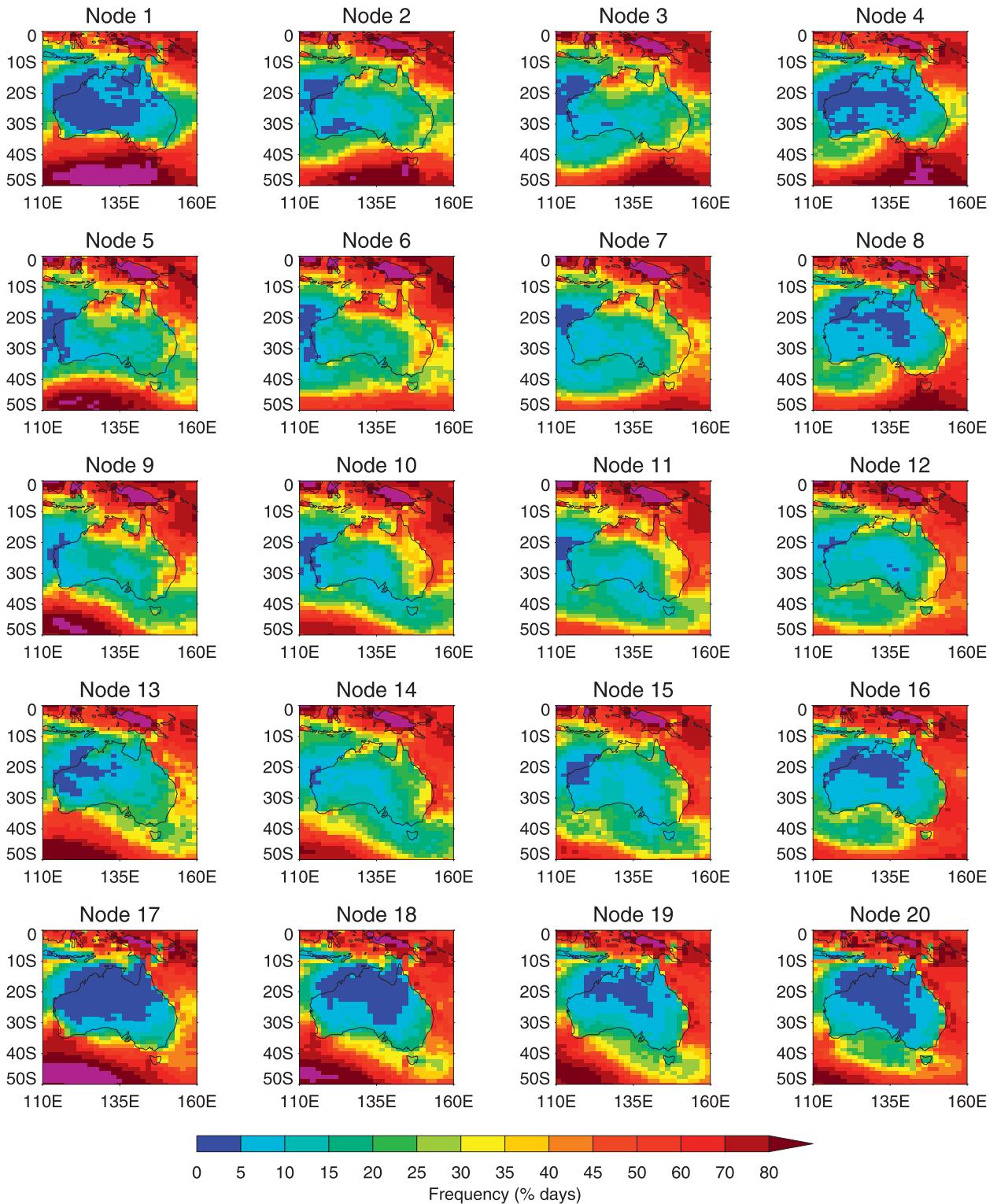


FIG. 10. Rainfall frequency (% days) by SOM node: ACCESS AMIP simulation daily rainfall from 1979 to 2000.

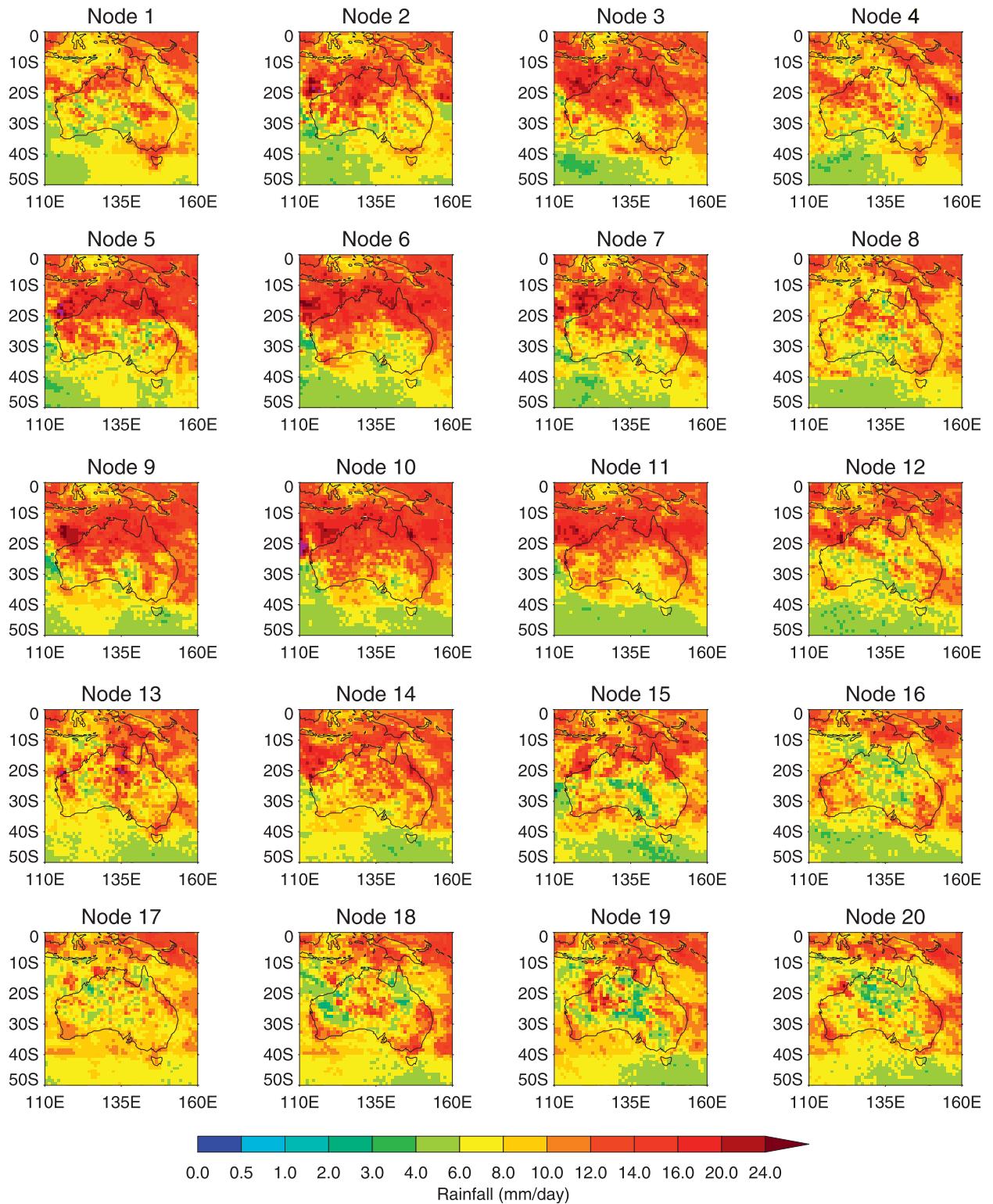


FIG. 11. As in Fig. 9, but for rainfall intensity ( $\text{mm day}^{-1}$ ).

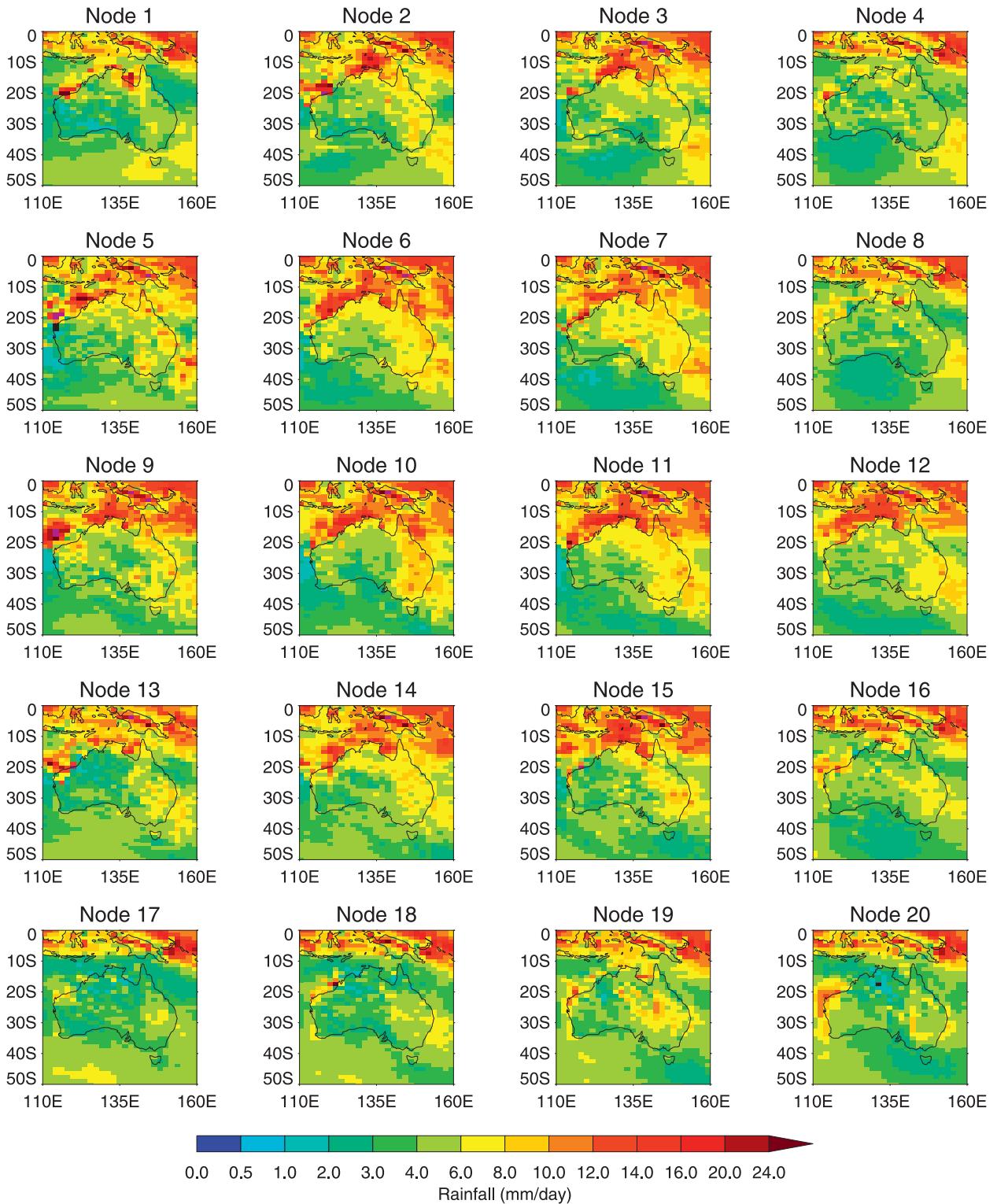
FIG. 12. As in Fig. 10, but for rainfall intensity ( $\text{mm day}^{-1}$ ).

TABLE 1. The calculation of  $R_{\text{total}}$  at Perth, Melbourne, and Darwin using Eq. (1). The model values of  $F$ ,  $Q$ , and  $N$  for each node are substituted in turn into Eq. (1), as shown in column 2, to investigate their relative influence on rainfall error at the three locations. The  $\Delta R_{\text{total}}$  ( $\text{mm day}^{-1}$ ) and percentage error (%) are shown in columns 4 and 5.

Location	Variables used to calculate $R_{\text{total}}$	$R_{\text{total}}$ ( $\text{mm day}^{-1}$ )	$\Delta R_{\text{total}}$ ( $\text{mm day}^{-1}$ )	Percentage error (%)
Perth	Observed $F, Q, N$	1.1	—	—
	Model $F, Q, N$	0.7	-0.4	-36
	Model $F$ , observed $Q$ and $N$	1.2	+0.1	+9
	Model $Q$ , observed $F$ and $N$	1.3	+0.2	+18
	Model $N$ , observed $F$ and $Q$	0.6	-0.5	-45
Melbourne	Observed $F, Q, N$	2.0	—	—
	Model $F, Q, N$	1.3	-0.7	-35
	Model $F$ , observed $Q$ and $N$	1.9	-0.1	-5
	Model $Q$ , observed $F$ and $N$	2.5	+0.5	+25
	Model $N$ , observed $F$ and $Q$	1.1	-0.9	-45
Darwin	Observed $F, Q, N$	4.3	—	—
	Model $F, Q, N$	1.8	-2.5	-58
	Model $F$ , observed $Q$ and $N$	3.9	-0.4	-9
	Model $Q$ , observed $F$ and $N$	3.3	-1.0	-23
	Model $N$ , observed $F$ and $Q$	2.6	-1.7	-40

and the GPCP 1DD rainfall dataset mapped onto the SOM derived from NCEP SLP, with the Perth, Melbourne, and Darwin records constructed from the nearest four land grid points as in section 3c. Using Eq. (1), we can decompose the error in total rainfall at the three locations

into error in the frequency of occurrence of each synoptic regime ( $F$ ), and error in the rainfall frequency ( $Q$ ) and rainfall intensity ( $N$ ) associated with each regime. As the observed and model values of  $F$ ,  $Q$ , and  $N$  are known for each node, we can substitute the observed values into the

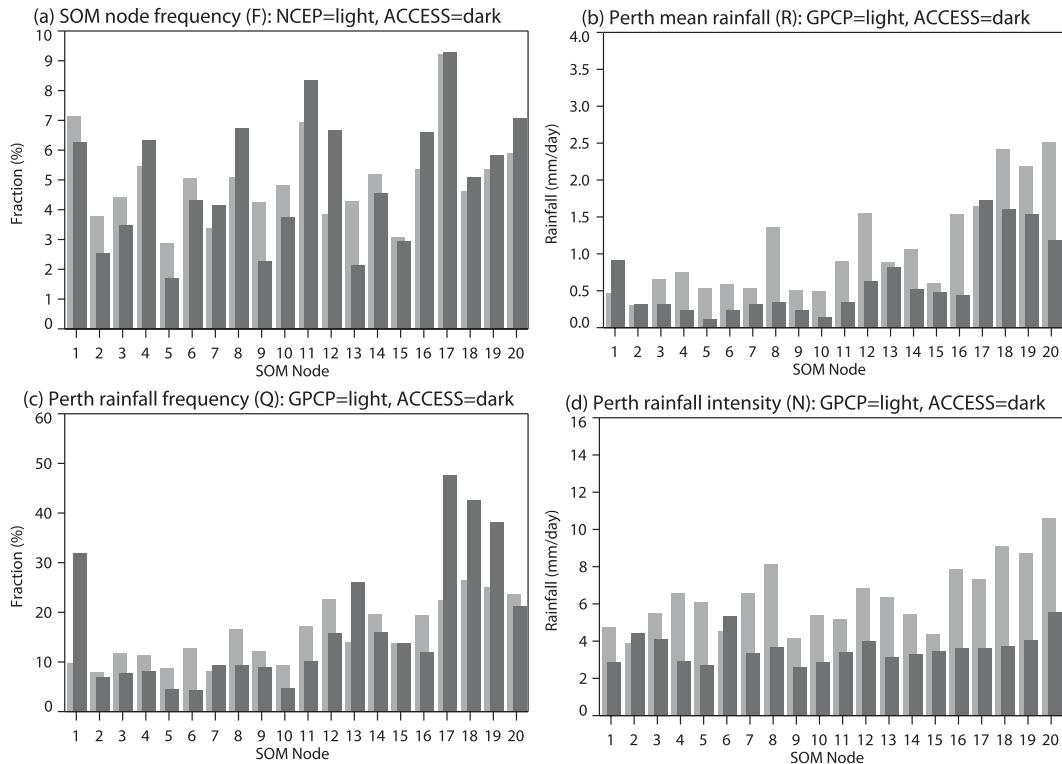


FIG. 13. (a) SOM node  $F$ , (b)  $R$ , (c)  $Q$ , and (d)  $N$  for grid points near Perth. In (a), SOM node  $F$  is calculated for the entire domain, using NCEP daily SLP (1979–2008), shown in light gray, and ACCESS daily SLP (1979–2000), shown in dark gray. In (b)–(d), GPCP daily rainfall (1997–2007) is shown in light gray and ACCESS daily rainfall (1979–2000) is shown in dark gray.

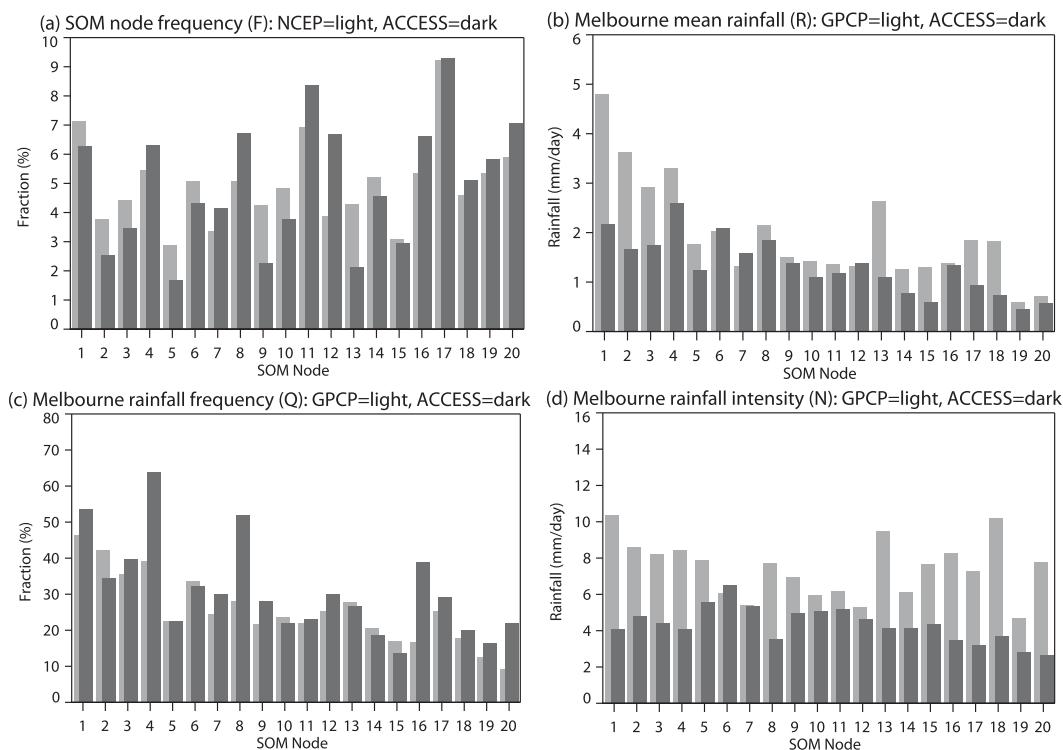


FIG. 14. As in Fig. 13, but near Melbourne.

equation for two out of three variables to identify the relative magnitude of error associated with the model  $F$ ,  $Q$ , and  $N$ , as shown in Table 1.

The rainfall frequency, rainfall intensity, and mean rainfall at Perth for each SOM node are shown in Fig. 13, as well as the frequency of occurrence for each node. The observed and model rainfall frequency is greatest for SOM nodes 17–20, which are predominantly winter patterns with a trough to the southwest of the continent. The model also simulates rainfall on more than 20% of days for node 1, in contrast with the drier observed pattern. It is clear that the underestimation of mean Perth rainfall by the model is primarily due to the low intensity of rainfall for all SOM nodes. The model rainfall intensity is almost uniform (average intensity =  $3.7 \text{ mm day}^{-1}$ ), whereas the observed intensity is almost 2 times greater (average intensity =  $7.0 \text{ mm day}^{-1}$ ), with more intense rainfall for nodes 17–20. The decomposition of errors in Table 1 indicates that the underestimation of  $N$  is the largest source of error at Perth, while this is partially offset by the overestimation of  $Q$ .

In Fig. 14, the rainfall frequency, rainfall intensity, and mean rainfall for each SOM node is shown for the grid points nearest to Melbourne. In the case of observations, the highest mean rainfall and highest rainfall frequency occur for SOM nodes 1–4, consistent with the rainfall anomalies in Fig. 7. The highest rainfall frequency in the

model occurs for nodes 1, 4, 8, and 16. The most important difference between observations and model appears to be the rainfall intensity, with the model simulating too low rainfall intensity for all SOM nodes, especially the wet southeast nodes 1–4. The mean rainfall in the model is lower for these same nodes. Therefore, the lower annual mean rainfall at Melbourne simulated by the model ( $1.3$  compared with  $2.0 \text{ mm day}^{-1}$  in the observations) is predominantly due to the low intensity of rainfall for all SOM nodes, and particularly the key wet southeast nodes 1–4 with a trough to the southeast of Australia. As shown in Table 1, model  $N$  is the largest source of error at Melbourne as at Perth, and again is partially offset by the overestimation of  $Q$ .

Figure 15 shows the rainfall frequency, rainfall intensity, and mean rainfall at Darwin by SOM node. The model broadly captures the rainfall frequency by node at Darwin, although the frequency of all the key summer monsoon nodes (3, 5, 6, 7, 9, 10, and 11) is somewhat underestimated. The most important difference between observed and model rainfall at Darwin is again rainfall intensity, as the model underestimates the intensity of the summer monsoon node rainfall by up to 50%. This bias results in substantially lower mean rainfall than observed for these key monsoon nodes, and summer (DJF) mean rainfall of less than half the observed value, as discussed in section 3. The low intensity of rainfall

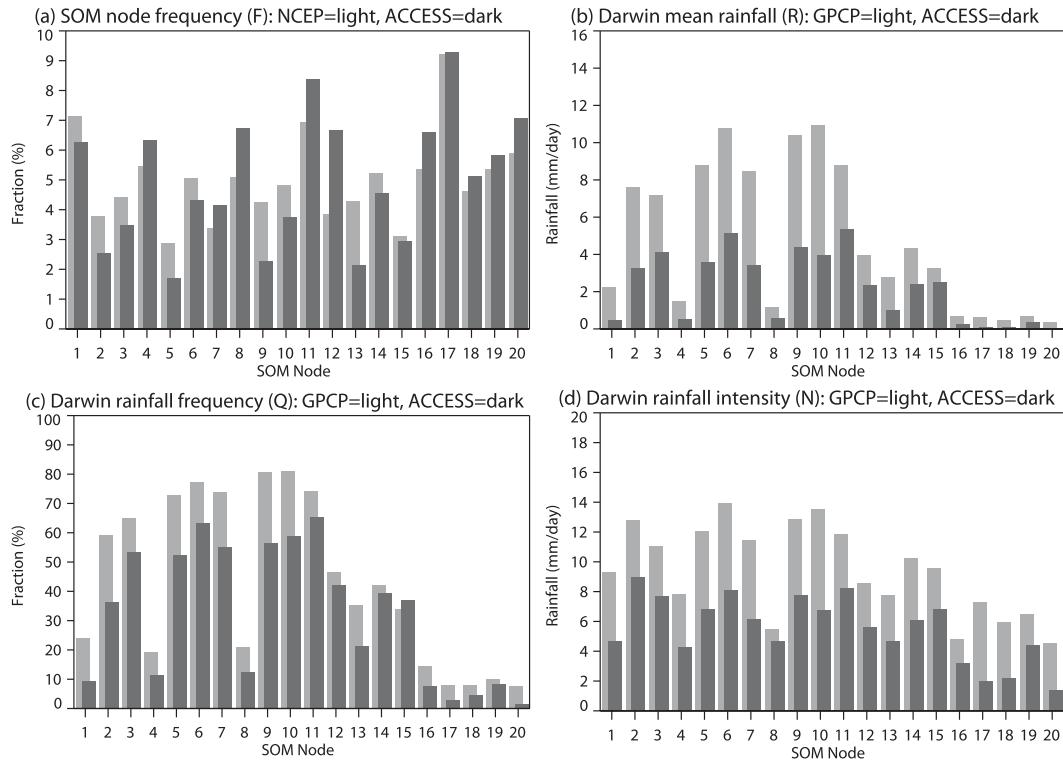


FIG. 15. As in Fig. 13, but near Darwin.

at Darwin is consistent with the suppression of heavy convective rainfall over land in the model, as seen for the summer monsoon nodes in Fig. 12. Errors calculated in Table 1 show that  $N$  is the largest source of error at Darwin but that  $Q$  is also underestimated, leading to a larger total percentage error in rainfall than for Perth or Melbourne.

## 5. Summary and conclusions

Previous studies have suggested that characteristics such as rainfall frequency and intensity are central to understanding and predicting the response of rainfall to climate change (e.g., Trenberth et al. 2003). Therefore, it is important to evaluate the ability of climate models to capture spatial and temporal patterns of rainfall frequency and intensity, and their relationship with atmospheric circulation regimes. We can then determine whether regional biases in model rainfall are due to the inability to simulate the dominant synoptic patterns, or the inability to capture the frequency and intensity of rainfall associated with such patterns. In this study, a combination of techniques was used to examine rainfall over the Australian continent and surrounding ocean. Daily rainfall from GPCP gridded observations was used to determine rainfall characteristics over the Australian

region, providing a target for the model simulations. Comparison between observations and ACCESS model output revealed a range of systematic biases in the model rainfall characteristics.

Comparison of the seasonal climatology of the model and GPCP rainfall over the Australian region revealed that the model was too dry over northern Australia during the summer monsoon, and that winter frontal rainfall over southern Australia was underestimated. An analysis of the frequency and intensity of light and heavy daily rainfall showed that ACCESS simulates too frequent light rainfall, both in the tropics and over southern Australia, while the frequency of heavy rainfall was more realistic. In addition, the model significantly underestimates the intensity of heavy rainfall over land, including the summer monsoon as well as winter rainfall over southern Australia, providing some explanation of the dry biases found in the model seasonal climatology.

Having identified a series of biases in the simulation of rainfall frequency and intensity over the region, the possible sources of the rainfall error were then investigated by considering the association between rainfall frequency and intensity and the dominant atmospheric circulation patterns. A self-organizing map was used to construct a set of SLP patterns or synoptic regimes, and the corresponding observed and model rainfall patterns

were examined. The rainfall for each synoptic regime was also decomposed into frequency and intensity to determine whether the model captured the distribution of rainfall associated with each regime. Three locations were considered in more detail: southwestern Australia (Perth), southeastern Australia (Melbourne), and northern Australia (Darwin).

It was found that ACCESS broadly captures the frequency of occurrence of the synoptic regimes represented by the SOM nodes [ $F$  in Eq. (1)], implying that the majority of errors in rainfall amount are not due to an inability to simulate the large-scale atmospheric circulation. In all three regions, the dominant source of error is the rainfall intensity ( $N$ ) of key SOM nodes or regimes, with a smaller contribution due to the rainfall frequency ( $Q$ ) associated with the nodes. In particular, the model does not capture the difference in rainfall intensity between nodes, failing to simulate more intense rainfall for regimes of intense convection or strong frontal systems.

Over northern Australia, ACCESS simulates too light rainfall over land, and too infrequent heavier rainfall over adjacent ocean for the “summer monsoon” SOM nodes. Over southern Australia, the model simulates light rainfall for all SOM nodes, failing to capture heavier rainfall ( $>10 \text{ mm day}^{-1}$ ) over the Perth and Melbourne regions in the “wet southeast” and “wet southwest” SOM nodes. The frequency of rainfall in these regions is overestimated by the model in the same nodes, partially offsetting the errors in model rainfall intensity. The occurrence of too frequent light rainfall is consistent with previous studies (e.g., Sun et al. 2006). The behavior of rainfall in the model as exposed by our analysis leads us to conclude that the most likely causes for the model errors lie in the representation of local subgrid-scale processes, such as convection and/or the representation of coastlines, rather than in the ability of the model to simulate the large-scale synoptic flow, while acknowledging that the two are, of course, not entirely independent of each other. That is, since the atmospheric circulation is relatively well simulated, we speculate that the error lies in the local processes that translate this large-scale circulation into rainfall. Further investigations are required to identify the details of the mechanisms responsible. These will include model sensitivity studies in particular for those regimes that exhibit large errors, with the ultimate aim to identify pathways to improved representations of subgrid-scale processes. As the main purpose of this study was to develop a methodology for a more insightful evaluation of model rainfall errors, such sensitivity studies are beyond the scope of this paper.

We have shown that the combination of techniques that were previously used independently provides a useful

set of tools for analyzing rainfall in observations and models, and for providing additional insight into the sources of model rainfall errors. The analysis could be refined further by considering the frequency and intensity of light and heavy rainfall by regime, as well as focusing on a single season. Further analysis of convective processes, moisture transport and sources, cloud properties, and microphysics is required to determine the detailed cause of the too frequent and too light rainfall simulated by the model, and the unrealistic sharp land–ocean gradients. This study has provided a useful new framework for doing so.

*Acknowledgments.* This study was supported by the Australian Research Council through the Linkage Project Grant LP0883961. Martin Dix and the ACCESS model team provided AMIP simulation output. Petteri Uotila provided software and guidance in the computation of self-organizing maps. The comments of three anonymous reviewers greatly improved this paper.

#### REFERENCES

- Adler, R. F., and Coauthors, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167.
- Alexander, L. V., P. Uotila, N. Nicholls, and A. Lynch, 2010: A new daily dataset for Australia and its application to the assessment of changes in synoptic patterns during the last century. *J. Climate*, **23**, 1111–1126.
- Cassano, J. J., P. Uotila, and A. H. Lynch, 2006: Changes in synoptic weather patterns in the polar regions in the 20th and 21st centuries, Part 1: Arctic. *Int. J. Climatol.*, **26**, 1027–1049.
- Cavazos, T., 2000: Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *J. Climate*, **13**, 1718–1732.
- Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate*, **19**, 4605–4630.
- Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res.*, **22**, 13–26.
- Hope, P., 2006: Projected future changes in synoptic systems influencing southwest Western Australia. *Climate Dyn.*, **26**, 765–780.
- , W. Drosowsky, and N. Nicholls, 2006: Shifts in the synoptic systems influencing southwest Western Australia. *Climate Dyn.*, **26**, 751–764.
- Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**, 36–50.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–470.
- Kohonen, T., 2001: *Self-Organizing Maps*. 3rd ed. Springer, 501 pp.
- Leloup, J. A., Z. Lachkar, J.-P. Boulanger, and S. Thiria, 2007: Detecting decadal changes in ENSO using neural networks. *Climate Dyn.*, **28**, 147–162.
- Martin, G. M., P. M. Inness, G.-Y. Yang, B. J. Hoskins, and J. M. Slingo, 2004: Tropical performance of HadGAM1. Met Office Hadley Centre Tech. Note 51, 13 pp. [Available online

- at [http://www.metoffice.gov.uk/publications/HCTN/HCTN\\_51.pdf](http://www.metoffice.gov.uk/publications/HCTN/HCTN_51.pdf)]
- Meehl, G. A., C. Covey, B. McAvaney, M. Latif, and R. J. Stouffer, 2005: Overview of the Coupled Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **86**, 89–93.
- Nicholls, N., P. Uotila, and L. Alexander, 2009: Synoptic influences on seasonal, interannual and decadal temperature variations in Melbourne, Australia. *Int. J. Climatol.*, 1372–1381, doi:10.1002/joc.1965.
- Pook, M. J., P. C. McIntosh, and G. A. Meyers, 2006: The synoptic decomposition of cool-season rainfall in the southeastern Australian cropping region. *J. Appl. Meteor. Climatol.*, **45**, 1156–1170.
- Pope, V., and Coauthors, 2007: The Met Office Hadley Centre climate modelling capability: The competing requirements for improved resolution, complexity and dealing with uncertainty. *Philos. Trans. Roy. Soc. London*, **A365**, 2635–2657, doi:10.1098/rsta.2007.2087.
- Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds. Cambridge University Press, 589–662.
- Richardson, A. J., C. Risien, and F. A. Shillington, 2003: Using self-organizing maps to identify patterns in satellite imagery. *Prog. Oceanogr.*, **59**, 223–239.
- Risbey, J. S., M. J. Pook, P. C. McIntosh, C. C. Ummenhofer, and G. Meyers, 2009: Characteristics and variability of synoptic features associated with cool season rainfall in southeastern Australia. *Int. J. Climatol.*, **29**, 1595–1613, doi:10.1002/joc.1775.
- Schuenemann, K. C., J. J. Cassano, and J. Finnis, 2009: Synoptic forcing of precipitation over Greenland: Climatology for 1961–99. *J. Hydrometeorol.*, **10**, 60–78.
- Sun, Y., S. Solomon, A. Dai, and R. Portmann, 2006: How often does it rain? *J. Climate*, **19**, 916–934.
- Taylor, K. E., D. Williamson, and F. Zwiers, 2000: The sea surface temperature and sea-ice concentration boundary conditions of AMIP II simulations. PCMDI Rep. 60, 20 pp.
- Tennant, W., 2003: An assessment of intraseasonal variability from 13-yr GCM simulations. *Mon. Wea. Rev.*, **131**, 1975–1991.
- Tozuka, T., J.-J. Luo, S. Masson, and T. Yamagata, 2008: Tropical Indian Ocean variability revealed by self-organizing maps. *Climate Dyn.*, **31**, 333–343.
- Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, **84**, 1205–1217.
- Uotila, P., A. H. Lynch, J. J. Cassano, and R. I. Cullather, 2007: Changes in Antarctic net precipitation in the 21st century based on Intergovernmental Panel on Climate Change (IPCC) model scenarios. *J. Geophys. Res.*, **112**, D10107, doi:10.1029/2006JD007482.
- Wilson, D. R., A. C. Bushell, A. M. Kerr-Munslow, J. D. Price, and C. J. Morcrette, 2008: PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description. *Quart. J. Roy. Meteor. Soc.*, **134**, 2093–2107.