

## Evidence for Using Lagged Climate Indices to Forecast Australian Seasonal Rainfall

ANDREW SCHEPEN

*Bureau of Meteorology, Brisbane, Queensland, Australia*

Q. J. WANG AND DAVID ROBERTSON

*CSIRO Land and Water, Highett, Victoria, Australia*

(Manuscript received 16 March 2011, in final form 18 July 2011)

### ABSTRACT

Lagged oceanic and atmospheric climate indices are potentially useful predictors of seasonal rainfall totals. A rigorous Bayesian joint probability modeling approach is applied to find the cross-validation predictive densities of gridded Australian seasonal rainfall totals using lagged climate indices as predictors over the period of 1950–2009. The evidence supporting the use of each climate index as a predictor of seasonal rainfall is quantified by the pseudo-Bayes factor based on cross-validation predictive densities. The evidence strongly supports the use of climate indices from the Pacific region with weaker, but positive, evidence for the use of climate indices from the Indian region and the extratropical region. The spatial structure and seasonal variation of the evidence for each climate index is mapped and compared. Spatially, the strongest supporting evidence is found for forecasting in northern and eastern Australia. Seasonally, the strongest evidence is found from August–October to November–January and the weakest evidence is found from March–May to May–July. In some regions and seasons, there is little evidence supporting the use of climate indices for forecasting seasonal rainfall. Climate indices derived from sea surface temperature anomalies in the Pacific region show stronger persistence in the relationship with Australian seasonal rainfall totals than climate indices derived from sea surface temperature anomalies in the Indian region. Climate indices derived from atmospheric variables are also strongly supported, provided they represent the large-scale circulation. Many climate indices are found to show similar supporting evidence for forecasting Australian seasonal rainfall, leading to the prospect of combining climate indices in multiple predictor models and/or model averaging.

### 1. Introduction

Probabilistic seasonal rainfall forecasts are important for users such as irrigators and water managers to assist in developing risk-management strategies and to inform decisions. Both statistical and dynamical climate prediction systems are widely used in practice to produce probabilistic seasonal rainfall forecasts up to a year in advance (Goddard et al. 2001). Statistical prediction systems are based on empirical relationships between observed variables and therefore rely on the availability of long data records and stationary relationships between the variables. Dynamical prediction systems are based on numerical simulations that directly model

physical processes, but they are more expensive to implement and operate than statistical climate prediction systems (Anderson et al. 1999). Despite substantial research efforts and technological advances, sophisticated dynamical prediction systems are still unable to consistently outperform simple statistical prediction systems for predicting ENSO and other climate variables (e.g., Barnston et al. 1999; Halide and Ridd 2008; Quan et al. 2006). Until dynamical prediction systems improve significantly, statistical prediction systems will continue to be improved and play a role in seasonal rainfall forecasting (e.g., Rajeevan et al. 2007).

There have been a number of studies that seek to explain relationships between oceanic and atmospheric circulation anomalies and Australian monthly or seasonal rainfalls (e.g., Meneghini et al. 2007; Murphy and Timbal 2008; Risbey et al. 2009; Wang and Hendon 2007). Concurrent relationships are typically quantified

Corresponding author address: Andrew Schepen, Bureau of Meteorology, GPO Box 413, Brisbane QLD 4001, Australia.  
E-mail: a.schepen@bom.gov.au

through linear regression analyses between rainfalls and a range of climate indices that represent anomalies in climate variables such as sea surface temperature, upper-ocean heat content, atmospheric pressure, and zonal wind. The strengths of the relationships vary with season and location.

For forecasting seasonal rainfall over the next season, climate indices that have a strong concurrent relationship with seasonal rainfall are natural candidates for use as predictors in a statistical prediction system. However, it is now the lagged relationship between seasonal rainfall and climate indices that matters. A strong concurrent relationship does not necessarily lead to a strong lagged relationship. Therefore, it is necessary to quantify the evidence supporting the use of various lagged climate indices for seasonal rainfall forecasting. Lagged indices of the Southern Oscillation have previously been found to be useful for forecasting Australian seasonal rainfall in some regions and seasons (Chiew et al. 1998; McBride and Nicholls 1983; Stone et al. 1996). Additionally, Kirono et al. (2010) evaluated correlations between various other climate indices averaged over the previous two months and Australian seasonal rainfalls for the four main seasons. Depending on the season and location they found significant correlations, measured by Pearson's correlation coefficient, between the lagged climate indices and seasonal rainfalls. However, detailed quantitative results were only presented for southeast Australia.

The objective of this study is to apply a rigorous method to quantify evidence supporting the use of lagged climate indices to forecast Australian seasonal rainfall.

To achieve this objective, we establish a sophisticated statistical prediction system based on a Bayesian modeling approach. For all 12 overlapping seasons and on a coarse grid that covers all of Australia, we compute the cross-validation predictive densities of seasonal rainfalls, which are then used to calculate pseudo-Bayes factors (PsBFs) (Geisser and Eddy 1979; Gelfand and Dey 1994; Robertson and Wang 2012). The PsBF is used as a measure of the statistical evidence supporting the use of lagged climate indices to forecast seasonal rainfall.

Because the PsBF is a direct measure of cross-validation predictive performance, it provides an estimate of the predictive ability of forecasting models for future events (Davis 1976; Michaelsen 1987). In this sense, it is a more desirable measure than Pearson's correlation coefficient that essentially evaluates how well a linear model fits data. Furthermore, Pearson's correlation coefficient is most suited for measuring linear relationships between variables that are normally distributed, while seasonal rainfall totals are generally highly skewed in distribution and their relationships with climate indices are often

nonlinear. In contrast, the PsBF evaluation here is based on forecasting models that allow for skewed distributions and nonlinear relationships (see section 2d).

The remainder of this paper is structured as follows: in the next section we introduce the climate indices and rainfall data to be analyzed, followed by the method used for this study. In section 3, we present figures and maps showing the evidence for using lagged climate indices to forecast Australian seasonal rainfalls and discuss the results. In section 4, we provide some additional discussion. Finally, section 5 completes the paper with a summary and conclusions.

## 2. Data and methods

### a. Overview

In this study, 13 monthly climate indices were investigated for evidence supporting their use as predictors of future Australian seasonal rainfall. The climate indices represent anomalies in large-scale oceanic and atmospheric variables. We apply a Bayesian joint probability (BJP) modeling approach to find the cross-validation predictive densities of seasonal rainfalls using the lagged climate indices as predictors in a set of single predictor models. We calculate PsBFs to estimate the evidence in favor of each model over a naïve no-predictor (or climatology) model.

### b. Climate indices

The climate indices investigated in this study have been chosen mostly because they have been shown to have a concurrent relationship with monthly or seasonal rainfall, although some of them have also been shown to have lagged relationships as discussed in the introduction. Table 1 summarizes the climate indices and provides brief descriptions of each. The climate indices are grouped into three geographical regions: Pacific, Indian, and extratropical. The period of record used for each climate index was from 1950 to 2009 unless stated otherwise.

The climate indices in the Pacific region represent anomalies in the oceanic and atmospheric variables over the Pacific Ocean and are associated with the El Niño–Southern Oscillation. The well-known Niño-3, Niño-3.4, and Niño-4 climate indices represent anomalies in near-equatorial sea surface temperatures. The ENSO Modoki Index (EMI) is a function of SST anomalies in the west, center, and east of the Pacific basin (Ashok et al. 2007). The Southern Oscillation index (SOI; Troup 1965) represents anomalies in the atmospheric pressure gradient between Tahiti and Darwin, Australia. The 20° isotherm climate index (D20) represents anomalies in

TABLE 1. Climate indices investigated as predictors of Australian seasonal rainfall: description and geographical region.

Predictor (alias)	Description	Region
SOI	Pressure difference between Tahiti and Darwin as defined by Troup (1965) and provided by BOM	Pacific
Niño-3	Average SST anomaly over 150°–90°W and 5°N–5°S	Pacific
Niño-3.4	Average SST anomaly over 170°–120°W and 5°N–5°S	Pacific
Niño-4	Average SST anomaly over 150°–160°E and 5°N–5°S	Pacific
EMI	$C - 0.5 \times (E + W)$	Pacific
	Where the components are average SST anomalies over C: 165°E–140°W and 10°N–10°S E: 110°–70°W and 5°N–15°S W: 125°–145°E and 20°N–10°S	
WPI	Average SST anomaly over 50°–70°E and 10°N–10°S	Indian
EPI	Average SST anomaly over 90°–110°E and 0°N–10°S	Indian
DMI	WPI–EPI	Indian
II	Average SST anomaly over 120°–130°E and 0°N–10°S	Indian
TSI	Average SST anomaly over 150°–160°E and 30°–40°S	Extratropical
B140	$0.5 \times [U(25) + U(30) - U(40) - 2 \times U(45) - U(50) + U(55) + U(60)]$ Where $U(X)$ is the 500-hPa zonal wind at latitude $X$	Extratropical
D20	Central Pacific 20°C isotherm depth anomaly as provided by BOM	Pacific
AAO	Antarctic Oscillation (Southern Annular Mode) anomaly as provided by NOAA	Extratropical

the heat content of the upper ocean over the central equatorial Pacific (Ruiz et al. 2006). Because D20 has a comparatively short record length (1980 onward), we will comment only briefly on the results.

The climate indices in the Indian region represent anomalies in oceanic and atmospheric variables over the Indian Ocean and the Indonesian Seas. The Indian Ocean West Pole index (WPI), Indian Ocean East Pole index (EPI), and Indian Ocean dipole mode index (DMI) (Saji et al. 1999) represent sea surface temperature anomalies in the northern Indian Ocean. The Indonesia index (II; Verdon and Franks 2005) represents sea surface temperature anomalies over the Indonesian Seas.

The climate indices in the extratropical region represent anomalies in oceanic and atmospheric variables south of the tropics. The Tasman Sea index (TSI; Murphy and Timbal 2008) represents sea surface temperature anomalies in the Tasman Sea. The 140°E blocking index (B140; Wright 1993) represents anomalies in the zonal wind across southern Australia at longitude 140°E. The Antarctic Oscillation climate index (AAO) (Mo 2000) represents anomalies in annular atmospheric circulation about the South Pole. Because AAO has a comparatively short record length (1979 onward), we will comment only briefly on the results.

Sea surface temperature climate indices were derived from the National Center for Atmospheric Research (NCAR) Extended Reconstruction of Sea Surface Temperature version 3 (Smith et al. 2008). B140 was derived from the National Centers for Environmental Prediction (NCEP)–NCAR reanalysis data (Kalnay et al. 1996). SOI and D20 were sourced from the

Australian Bureau of Meteorology (BOM). AAO was sourced from the National Oceanic and Atmospheric Administration (NOAA).

### c. Rainfall data

The rainfall data used in this study were derived from the Australian Water Availability Project (AWAP) (Jones et al. 2009)  $0.05^\circ \times 0.05^\circ$  gridded dataset of monthly rainfall for the period 1950–2009. The accuracy of these data is poorest in central-west Australia where the rain gauge network is sparse. Seasonal rainfall totals were calculated by aggregating monthly rainfall totals. For the purposes of this study, the rainfall totals were upscaled to  $2.5^\circ \times 2.5^\circ$  resolution. The grid used in this study is therefore closely aligned with the grid used by the Predictive Ocean Atmosphere Model for Australia (POAMA), a dynamical climate prediction system used by the Australian Bureau of Meteorology. There are 122 grid cells covering Australia.

### d. Analysis

In this study, we used a Bayesian joint probability modeling approach to find the cross-validation predictive densities of seasonal rainfalls using climate indices as predictors. Wang et al. (2009) and Wang and Robertson (2011) describe in detail the mathematical formulation and implementation of the BJP modeling approach. The application of the approach to the problem of finding cross-validation predictive densities is described in full in Robertson and Wang (2012). Only a few relevant aspects of the approach are summarized here.

TABLE 2. Scale of evidence for the  $\ln(\text{PsBF})$ .

$\ln(\text{PsBF})$	Evidence for $M_1$ against $M_0$
<2	Little
2–4	Positive
4–6	Strong
>6	Very strong

There are a number of hindrances to jointly modeling climate indices and rainfall. First, rainfall data are often highly skewed in distribution and their relationship with climate indices can be nonlinear. Second, periods of zero seasonal rainfall occur in some seasons in some parts of Australia. Third, the joint distributions can be difficult to determine analytically. These problems are addressed by the BJP modeling approach. Skewed data are transformed to normality by building the Yeo–Johnson transform into the approach. Zero values in seasonal rainfall records are treated as censored data, allowing the use of multivariate normal distributions. The posterior distributions of the parameters of the Yeo–Johnson transforms and multivariate normal distributions are inferred numerically through Markov Chain Monte Carlo (MCMC) simulations.

The predictive density of an event  $t$  is the probability density of the predictand variables conditional on known predictor values, given the distribution of model parameters. By using the distribution of model parameters inferred from the data record after excluding the event  $t$ , the cross-validation predictive density for event  $t$  is derived.

Mathematically, if  $\mathbf{y}^t$  is a vector holding the predictor variables  $\mathbf{y}^t(1)$  and predictand variables  $\mathbf{y}^t(2)$  for event  $t$ ,  $\mathbf{Y}_{\text{OBS}}^{(t)}$  is a matrix containing observed values of predictor and predictand variables for the remaining events, and  $\boldsymbol{\theta}$  represents the multivariate normal distribution parameters and the Yeo–Johnson transform coefficients, then the cross-validation predictive density of the predictand variables for event  $t$  is given by

$$f[\mathbf{y}^t(2)|\mathbf{y}^t(1)] = \int p[\mathbf{y}^t(2)|\mathbf{y}^t(1); \boldsymbol{\theta}] p[\boldsymbol{\theta}|\mathbf{Y}_{\text{OBS}}^{(t)}] d\boldsymbol{\theta}. \quad (1)$$

The PsBF compares competing models  $M_1$  and  $M_0$  in terms of cross-validation predictive densities evaluated at the observed predictand values for all events  $t = 1, 2, \dots, n$ . Therefore, we derive  $n$  cross-validation predictive densities for each candidate predictor model for determination of the PsBF. Mathematically, the PsBF can be given in logarithmic form by

$$\ln(\text{PsBF}) = \ln \prod_{t=1}^n \frac{f_{M_1}[\mathbf{y}^t(2) = \mathbf{y}_{\text{OBS}}^t(2)]}{f_{M_0}[\mathbf{y}^t(2) = \mathbf{y}_{\text{OBS}}^t(2)]}. \quad (2)$$

In our study,  $f_{M_1}$  is the cross-validation predictive density with a single climate index as a predictor variable and  $f_{M_0}$  is the cross-validation predictive density with no predictor variable. The PsBF assesses model  $M_1$  against model  $M_0$  in terms of their abilities to predict events that have not been used in the sampling of the model parameters. When  $\ln(\text{PsBF}) > 0$ , data evidence is in favor of model  $M_1$  over

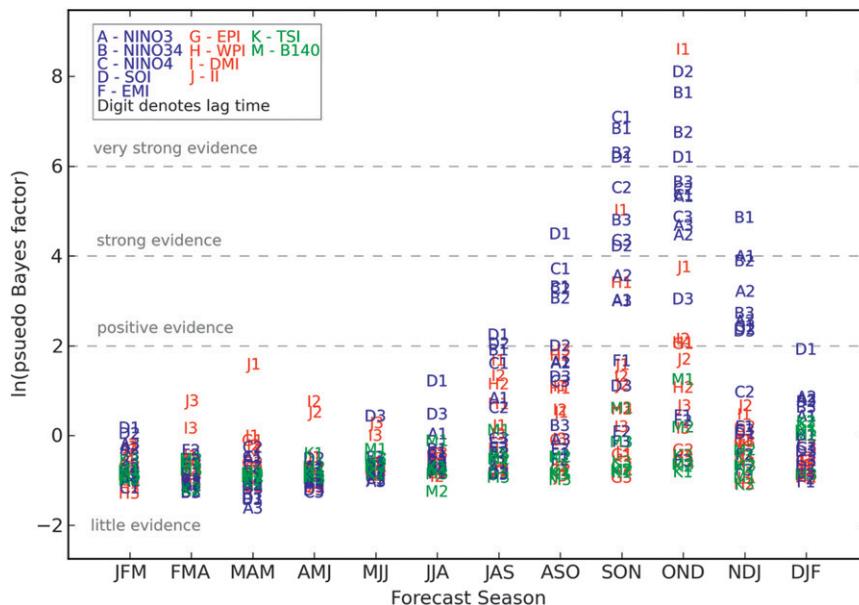


FIG. 1. Plot of  $\ln(\text{PsBF})$  (evidence for forecasting seasonal rainfall) vs season for the grid cell with centroid  $(36.0^\circ\text{S}, 145.0^\circ\text{E})$ .

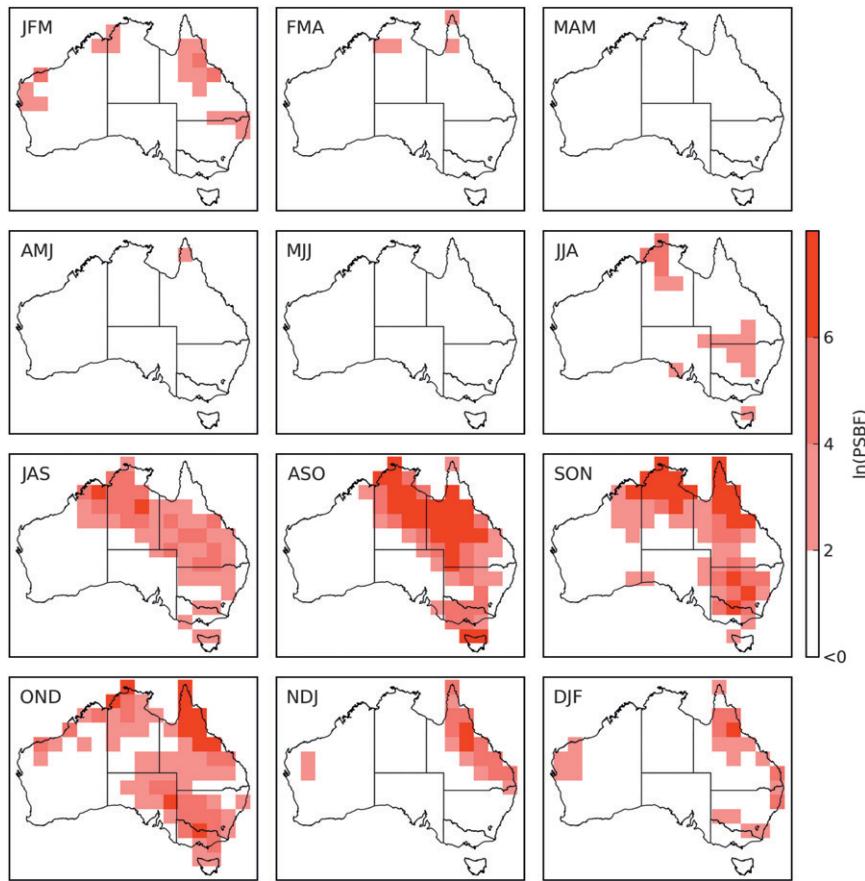


FIG. 2. Map of  $\ln(\text{PsBF})$  showing the evidence for forecasting Australian seasonal rainfall using the 1-month-lagged SOI.

$M_0$ . When a strong relationship between a climate index and seasonal rainfall underlies the model  $M_1$ , the value of PsBF is expected to be high.

By using the same reference model  $M_0$ , PsBFs can be calculated for multiple models  $M_1$ . In our case  $M_1$  represents competing models using different predictors;  $M_0$  is held fixed as the model with no predictor and therefore represents a climatology forecast. A common set of events is used in Eq. (2) so that the derived PsBF values are comparable.

The PsBF is a variant of the classical Bayes factor (Geisser and Eddy 1979; Gelfand and Dey 1994). While Jeffreys established a scale for describing the strength of evidence according to the Bayes factor (Kass and Raftery 1995), such a scale does not exist for the PsBF. In this study, we set up Table 2 to align qualitative descriptions of the PsBF as evidence of forecasting ability. The table is a modified version of the scale for the Bayes factor given in Kass and Raftery (1995) and has been calibrated as described in Robertson and Wang (2012). There is a 95% chance that the relationship between

a climate index and seasonal rainfall is true if the  $\ln(\text{PsBF})$  is greater than 4. Additionally, the lower bound of  $\ln(\text{PsBF})$  values was found to be above -2 (see Fig. 1 in the results). On balance  $\ln(\text{PsBF})$ s between -2 and 2 are considered noise, with values greater than 2 increasingly signifying positive statistical evidence.

In this study, we calculated the PsBFs for forecasts of seasonal rainfall totals from the first day of each month. Forecasts were made separately for each of the 122 grid cells using single predictor models. The predictors evaluated were 13 monthly climate indices lagged by 1–3 months, giving 39 candidate predictor models for each season.

### 3. Results and discussion

#### a. Seasonal evidence at a selected grid point

To assist in the interpretation of the maps in the subsequent sections of these results, we first present a graph of the  $\ln(\text{PsBF})$  versus season for a selected grid cell (Fig. 1).

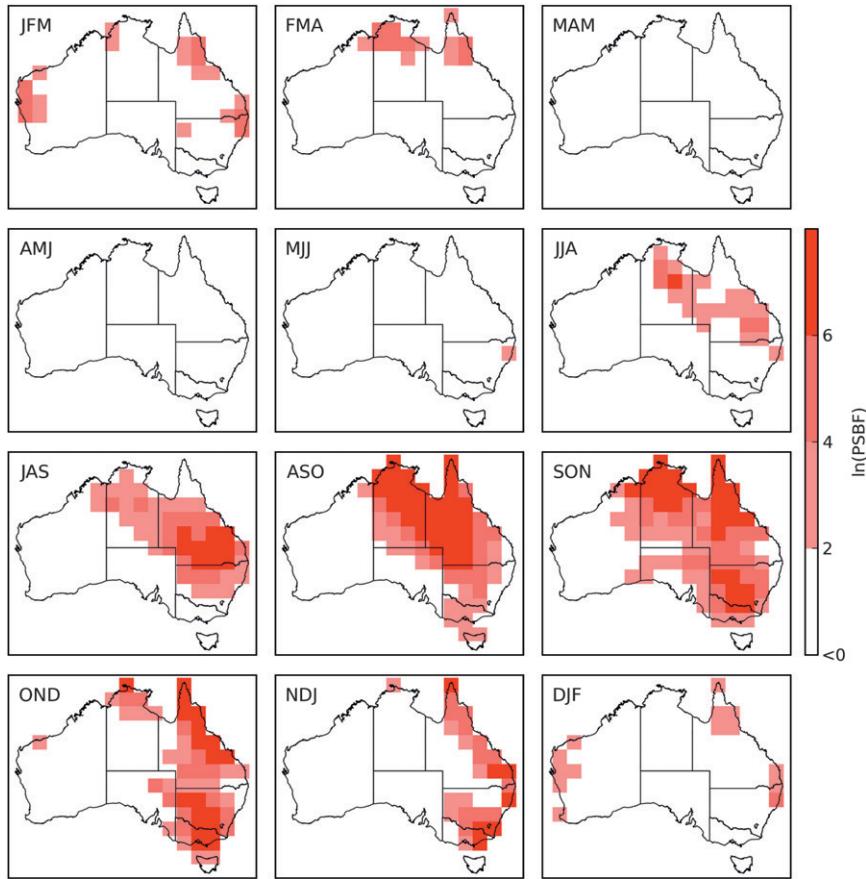


FIG. 3. As in Fig. 2, but using Niño-3.4.

The centroid of this grid cell is in southeastern Australia at (36.0°S, 145.0°E). Classifications of the  $\ln(\text{PsBF})$  as evidence of forecasting ability are drawn on the graph as per Table 2. In this location, there is little evidence supporting the use of lagged climate indices to forecast rainfall in the first half of the year. There is, however, evidence supporting the use of lagged climate indices to forecast July–September (JAS) to November–January (NDJ) rainfall, with predictability peaking for October–December (OND) forecasts. For forecasting September–October (SON) and OND rainfall, there is strong evidence that both Pacific and Indian region indices are useful.

*b. Seasonal evidence for forecasting rainfall using 1-month-lagged SOI, Niño-3.4, DMI, and TSI*

Figures 2–5 show the evidence for forecasting rainfall in each season using the 1-month-lagged SOI, Niño-3.4, DMI, and TSI. This subset was selected to compare indices from each geographical region and, in the case of SOI and Niño-3.4, distinguish between atmospheric and oceanic climate indices. There is strong evidence supporting the use of the 1-month-lagged SOI to forecast JAS

to December–February (DJF) rainfall in eastern and central northern Australia, with peak predictability occurring in August–October (ASO) and SON (Fig. 2). The spatial and seasonal pattern of evidence supporting the use of the 1-month-lagged Niño-3.4 is similar to but generally stronger than SOI (Fig. 3). SOI has previously been shown to have a significant lagged relationship with SON rainfall in the areas described above (Chiew et al. 1998). However, the same study did not find any significant correlation between SST anomalies averaged over the entire Niño-3, Niño-3.4, and Niño-4 region and SON rainfall. In contrast, our result shows that the SST anomaly in the smaller Niño-3.4 region has greater supporting evidence for forecasting eastern Australian rainfall than SOI, suggesting that SST anomalies are in fact more important.

There is strong—very strong evidence supporting the use of the 1-month-lagged DMI to forecast Australian SON and OND rainfall in a small band across southeastern Australia and at the top edges of the continent (Fig. 4). This result is discussed further in the next subsection on SON rainfall.

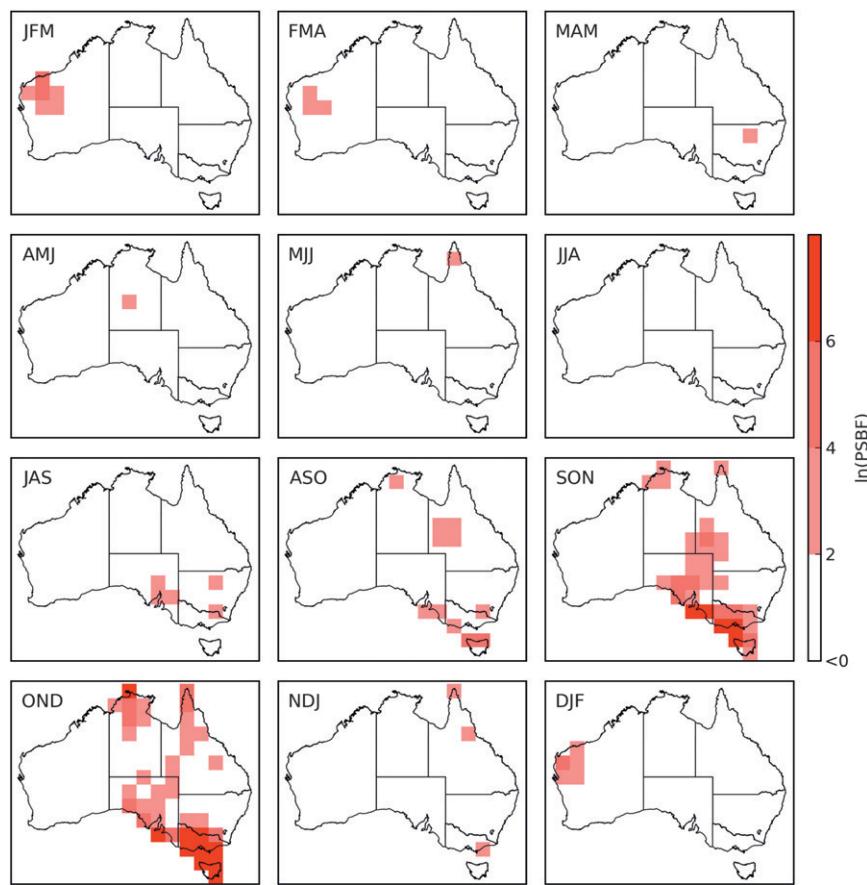


FIG. 4. As in Fig. 2, but using DMI.

There is strong—very strong evidence supporting the use of the 1-month-lagged TSI to forecast OND and NDJ rainfall in the western half of Australia (Fig. 5). The mechanism for this relationship is not understood as the relationship between TSI and seasonal rainfall has not been studied this region. A region where a relationship has been established is in southeastern Australia, where TSI is significantly correlated with concurrent March–May (MAM) rainfall (Murphy and Timbal 2008). Warm SST anomalies in the Tasman Sea are associated with increased rainfall in this region. Our results show that this relationship does not persist and there is no evidence supporting the use of the lagged TSI to forecast seasonal rainfall in southeastern Australia in any season.

#### c. Evidence for forecasting SON rainfall

Concurrent relationships between climate indices and Australian SON rainfall are well established (Risbey et al. 2009). We extend these results to show that there is strong evidence supporting the use of Pacific region indices (Niño-3, Niño-3.4, Niño-4, EMI,

and SOI) lagged out to 3 months to forecast SON rainfall in eastern and northern Australia (Figs. 6–8). As expected, the strength and spatial extent of the evidence decreases as lag time is increased. At a 3-month lag, the evidence supporting the use of Niño-3.4 and Niño-4 for forecasting SON rainfall remains particularly strong across large parts of eastern and northern Australia. The evidence is positive to strong for the 1-month-lagged EMI in central Australia and parts of western Australia. The proximity of an SST anomaly component located to the north of Australia is most likely why the lagged EMI has relationships farther west than the other indices.

There is strong evidence supporting the use of DMI and WPI lagged 1 month to forecast SON rainfall in southeastern Australia (Fig. 5). However, there is little evidence for using DMI lagged more than 1 month to forecast SON rainfall (Figs. 7, 8). This result is partly in contrast to Kirono et al. (2010), who did not find any significant correlation between SON rainfall and the DMI averaged over the previous 2 months. However, their study did identify a significant correlation between

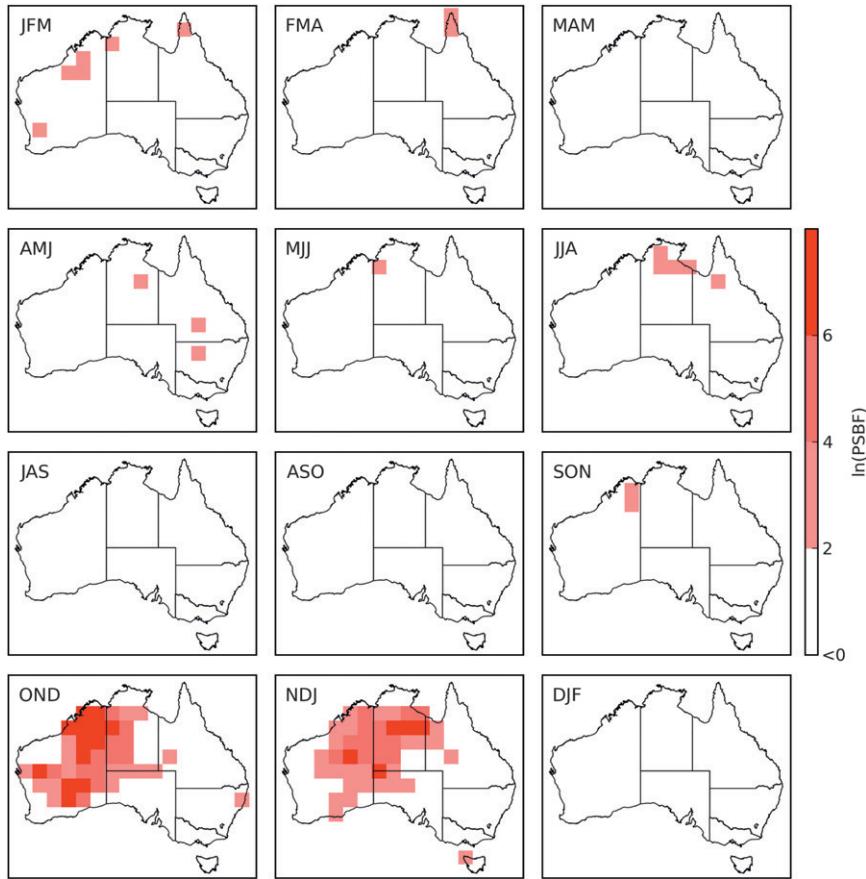


FIG. 5. As in Fig. 2, but using TSI.

SON rainfall in southeastern Australia and WPI averaged over the previous 2 months. We find there is potential predictability arising from sea surface temperatures near Indonesia. There is strong evidence supporting the use of II lagged out to 2 months for forecasting SON rainfall in the central northern tips and in parts of eastern Australia.

There is little evidence supporting the use of extratropical climate indices to forecast SON rainfall anywhere in Australia (Figs. 6–8). While Risbey et al. (2009) concluded that B140 has a significant concurrent relationship with SON rainfall across most of eastern Australia, it is clear that lagged B140 is not useful for forecasting it. One explanation for this is that B140 is associated with higher-frequency weather systems such as cold fronts, rather than the large-scale circulation.

The evidence supporting the use of climate indices to forecast Australian SON rainfall is summarized by Fig. 9. The summary shows the climate index and lag time with the highest supporting evidence in each grid cell. For brevity we just present and discuss summary figures for the three remaining main seasons.

#### *d. Evidence for forecasting DJF rainfall*

Compared to SON, there is less evidence for forecasting DJF rainfall across Australia and the focus has shifted from east to west. There is strong evidence supporting the use of the 1-month-lagged SOI to forecast DJF rainfall in northeastern Australia. Farther south, there is strong evidence supporting the use of the 2- or 3-month-lagged Niño-3 near the coast (Fig. 10). For this region, it is apparent that DJF rainfall predictability stems from SST anomalies in the eastern Pacific at around the peak of an El Niño or La Niña event. There is strong evidence supporting the use of the 3-month-lagged TSI to forecast DJF rainfall to the north of the Great Australian Bight. As was discussed in section 3b, the mechanism for the relationship between TSI and western Australian seasonal rainfall is unclear. However, this result confirms that the September Tasman SST anomalies are related to western Australian rainfall in the next few seasons. A partial explanation of the relationship could be that the warming trend in the September TSI explains rainfall trends over

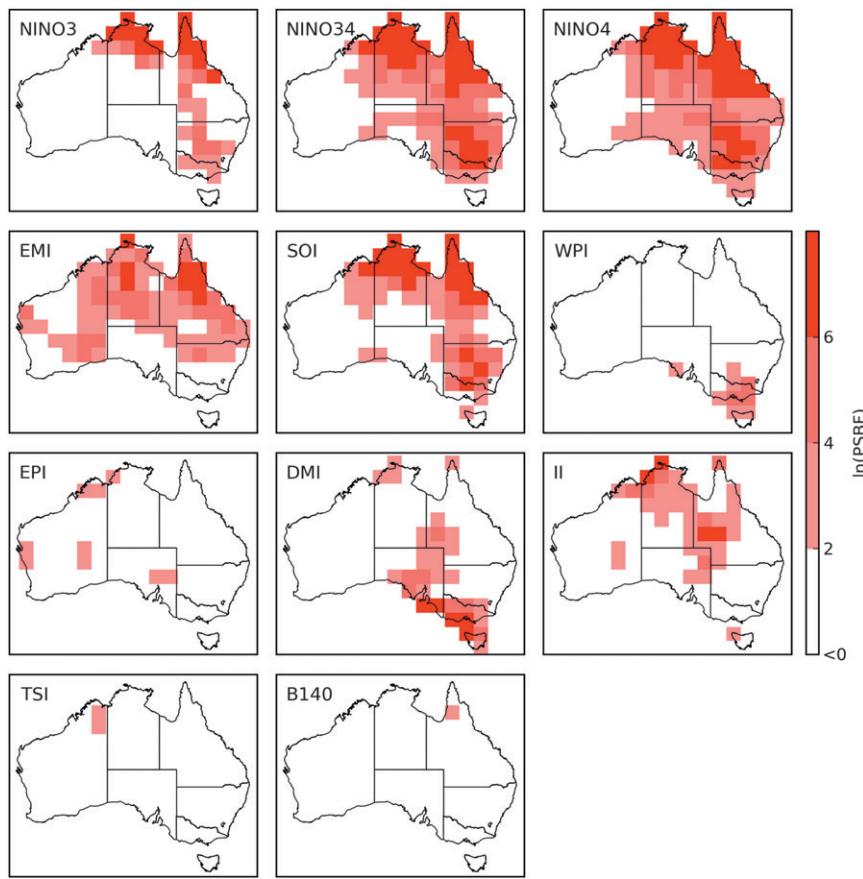


FIG. 6. Map of  $\ln(\text{PsBF})$  showing the evidence for forecasting Australian SON rainfall using 1-month-lagged climate indices.

the western half of the continent late in the year. We further discuss the implications of trends in section 4. There is positive to strong evidence for using 1-month-lagged II and EPI to forecast DJF rainfall in northwestern Australia. It is apparent that the rainfall there is being driven by proximate SST anomalies.

#### e. Evidence for forecasting MAM rainfall

There is little evidence supporting the use of any lagged climate indices to forecast MAM rainfall across most of Australia (Fig. 11). However, there is positive evidence supporting the use of the 2- or 3-month-lagged EMI to forecast MAM rainfall in the northern tropical regions. Exploring this result further, we find there is also evidence supporting the use of the 1-month-lagged EMI to forecast February–April (FMA) rainfall in the same region (not shown). The EMI represents a special case of El Niño, known as El Niño Modoki, which is characterized by a pattern of warm SST anomalies in the central equatorial Pacific, straddled by cool SST anomalies to the west and east (Ashok et al. 2007). A recent

study found that El Niño Modoki events are associated with a shorter duration and more intense monsoon season (the normal season is defined as December–March) across northern Australia (Taschetto et al. 2009). Although we do not seek to provide physical interpretations, this previous finding adds credibility to our result that the December and January El Niño Modoki patterns are indicators of northern Australian rainfall through the latter half of the monsoon season.

#### f. Evidence for forecasting JJA rainfall

After a low point in seasonal rainfall predictability for MAM, there is increased evidence supporting the use of lagged climate indices to forecast June–August (JJA) rainfall (Fig. 12). There is positive evidence supporting the use of 1-month-lagged Niño-4 to forecast JJA rainfall across parts of eastern Australia. In the central north region, the 1-month-lagged Niño-3.4 and SOI indices are best supported. The predictability associated with SOI in the central northern tip of Australia is thought to result from the direct influence of Darwin sea level

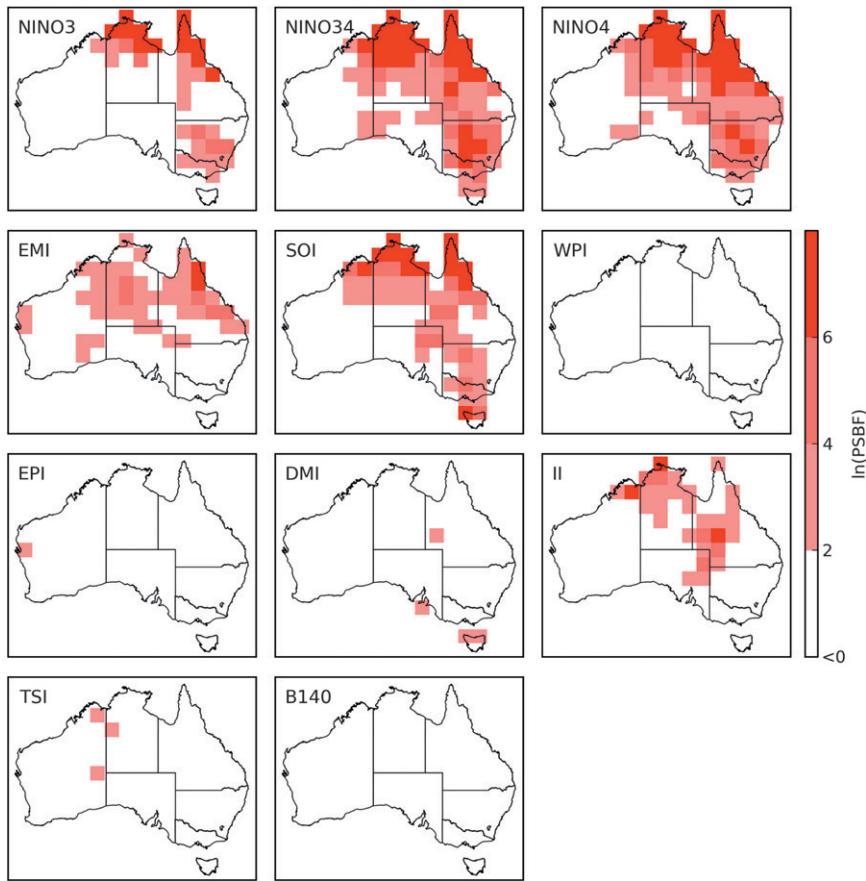


FIG. 7. As in Fig. 6, but using 2-month-lagged climate indices.

pressure. There is positive supporting evidence for using 3-month-lagged B140 to forecast JJA rainfall over southwestern Australia. The positive evidence for 3-month-lagged B140 is somewhat unexpected (anywhere), given that in section 3c we suggested that B140 was associated with higher-frequency weather patterns. One explanation for the result in southwestern Australia is that a reasonably consistent relationship exists whereby blocking conditions over southeastern Australia (at 140°E) in March tend to be followed by blocking conditions over southwestern Australia a few months later. The explanation is based on the observation that blocking over the Great Australia Bight favors rainfall in Western Australia (Risbey et al. 2009), but further investigation is required to confirm this.

#### *g. Evidence for forecasting seasonal rainfall using D20 and AAO*

We comment only briefly on the evidence supporting the use of D20 and AAO for forecasting seasonal rainfall because their records are relatively short. There is

positive evidence supporting the use of D20 to forecast JJA–NDJ rainfall in the regions of eastern and northern Australia. However, there is overlapping and stronger evidence supporting the use of SST indices such as Niño-4, suggesting that little advantage is to be obtained by using subsurface temperatures. There is little evidence supporting the use of AAO to forecast Australian seasonal rainfall. This is not unexpected given that Meneghini et al. (2007) reported generally weak concurrent relationships between a similar AAO and Australian seasonal rainfall.

#### *h. Comparison of evidence for 1950–79 and 1980–2009*

It is not expected that the evidence for using lagged climate indices to forecast seasonal rainfall is consistent over all decades. However, analyzing the evidence in individual decades is problematic because of the increased chance of statistical error due to the small sample size. Also, the precise location of seasonal rainfall affected by the large-scale circulation is variable within a region and therefore a long period is required to

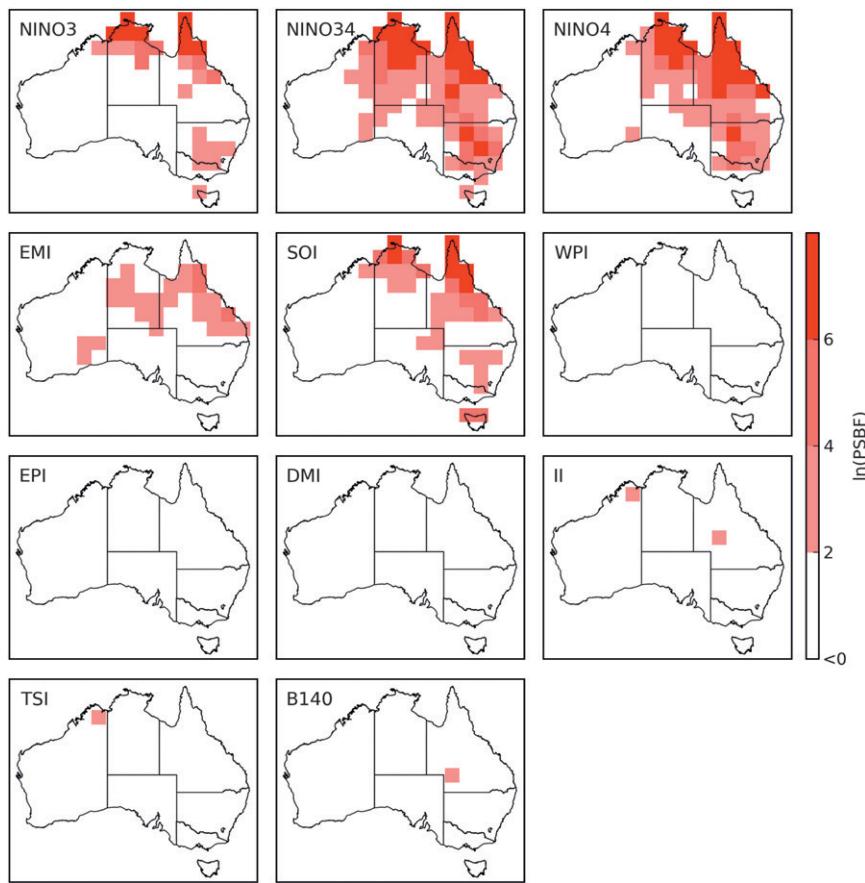


FIG. 8. As in Fig. 6, but using 3-month-lagged climate indices.

build up the spatial pattern of evidence. Taking into account the two issues above, we judge whether the evidence for using lagged climate indices is consistent over time by calculating the PsBFs for the periods 1950–79 and 1980–2009 and by using only the events of interest in Eq. (2) (Figs. 13, 14). We can then make some general inferences about the relative performance in the two periods; we cannot make a direct comparison with the other results, as there is only half the number of events.

The spatial pattern of evidence supporting the use of climate indices to forecast seasonal rainfall is very similar between the two periods. However, the evidence is weaker in the most recent 30-yr period of 1980–2009 than the preceding 30-yr period of 1950–79. It is likely that the difference in evidence for the two periods is (in part) due to variability at decadal or even longer time scales. According to a classification of ENSO events according to the average SOI from June to November 2011 (<http://www.longpaddock.qld.gov.au/products/australiasvariableclimate/ensoyearclassification.html>), the two periods have an equal number of neutral

ENSO events. However, 1980–2009 has a higher proportion of El Niño events and 1950–79 a higher proportion of La Niña events. Therefore, there may be better supporting evidence for using lagged climate indices to forecast seasonal rainfall in periods when La Niña conditions dominate. Capturing these longer-term effects will continue to be a challenge in seasonal forecasting.

#### 4. Further discussion

A clear result from this study is that there is simply little evidence supporting the use of lagged climate indices to forecast seasonal rainfall in many locations. One example is the southwest corner of western Australia (SWWA). In this region, there is little evidence for forecasting SON rainfall and only marginally positive evidence for using 3-month-lagged climate indices to forecast JJA and DJF rainfall, suggesting very weak relationships between the climate variables studied and rainfall in the region. This follows the result of Feng et al. (2010), who found that SWWA rainfall is linked to

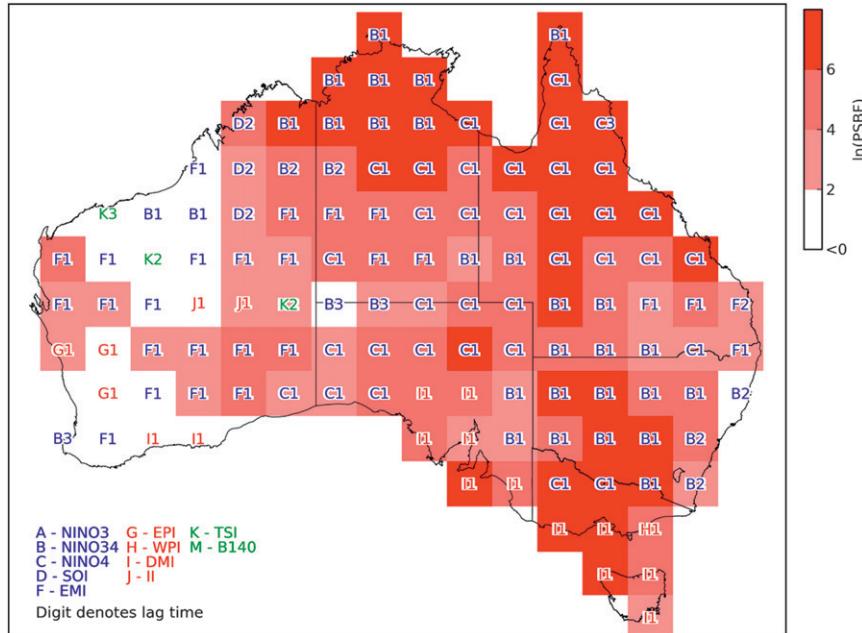


FIG. 9. Map of max  $\ln(\text{PsBF})$  and associated climate index, showing the climate index with the strongest supporting evidence for forecasting Australian SON rainfall in each grid cell: Pacific (blue labels), Indian (red labels), and extratropical (green labels) regions.

a large-scale circulation over southwestern Australia, rather than circulations in the Indian, Pacific, or Antarctic regions.

The results show that there is stronger evidence for using lagged Niño-3.4 and Niño-4 to forecast Australian

seasonal rainfall than lagged Niño-3. A recent study of the zonal distribution of El Niño events and Australian seasonal rainfall found that the concurrent relationship is more sensitive to variation in central and western Pacific Ocean SST anomalies than eastern

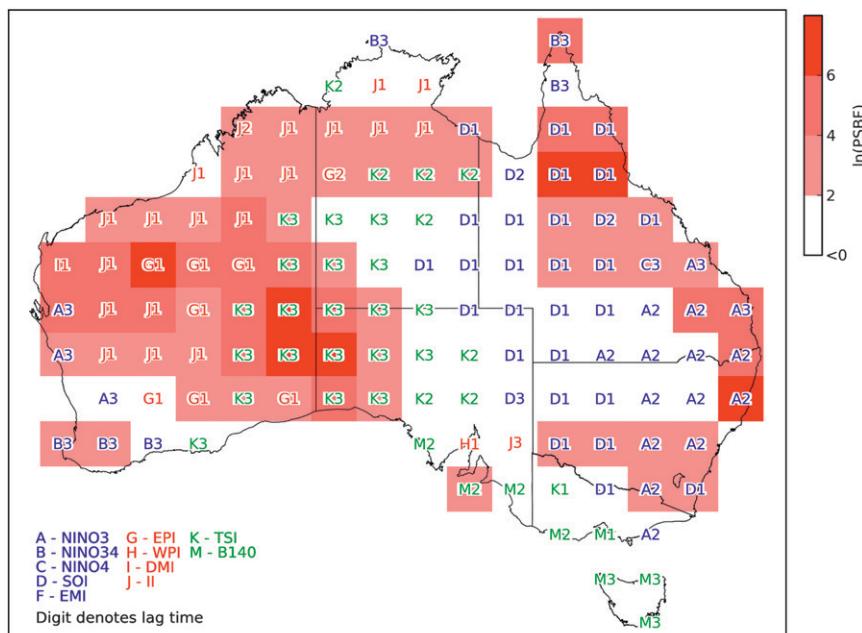


FIG. 10. As in Fig. 9, but for DJF rainfall.

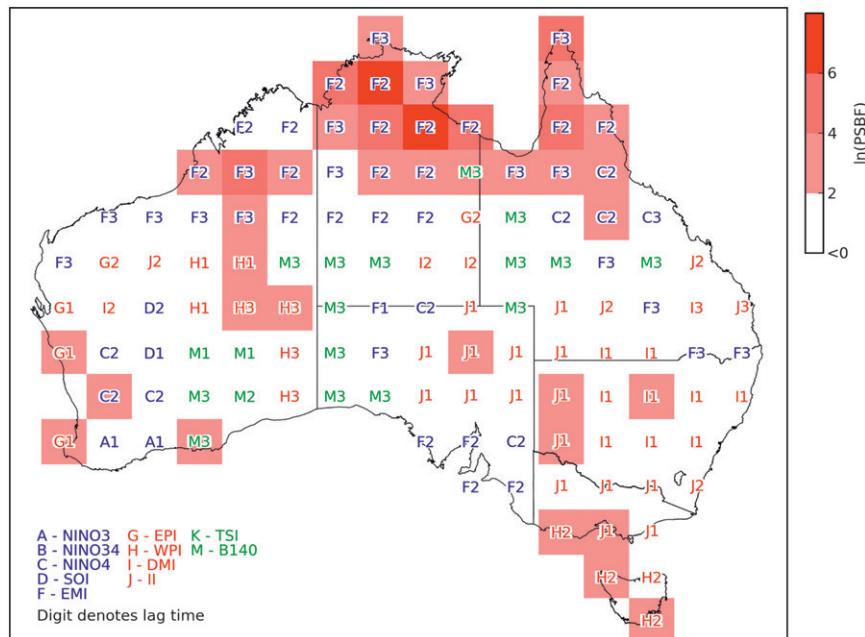


FIG. 11. As in Fig. 9, but for MAM rainfall.

Pacific Ocean SST anomalies (Wang and Hendon 2007). The theory presented in that study is that if an El Niño pattern is shifted eastward then Australian rainfall may not be affected as strongly as it would be if an El Niño pattern was shifted westward. The results of our study add further empirical evidence to this theory.

There is some debate on the degree to which the Indian oceanic-atmospheric circulation is independent of the Pacific oceanic-atmospheric circulation (Allan et al. 2001; Ashok et al. 2003; Meyers et al. 2007). Our results show that the evidence supporting the use of DMI is limited to forecasting SON and OND rainfall in south-east Australia. Although this evidence for DMI is very

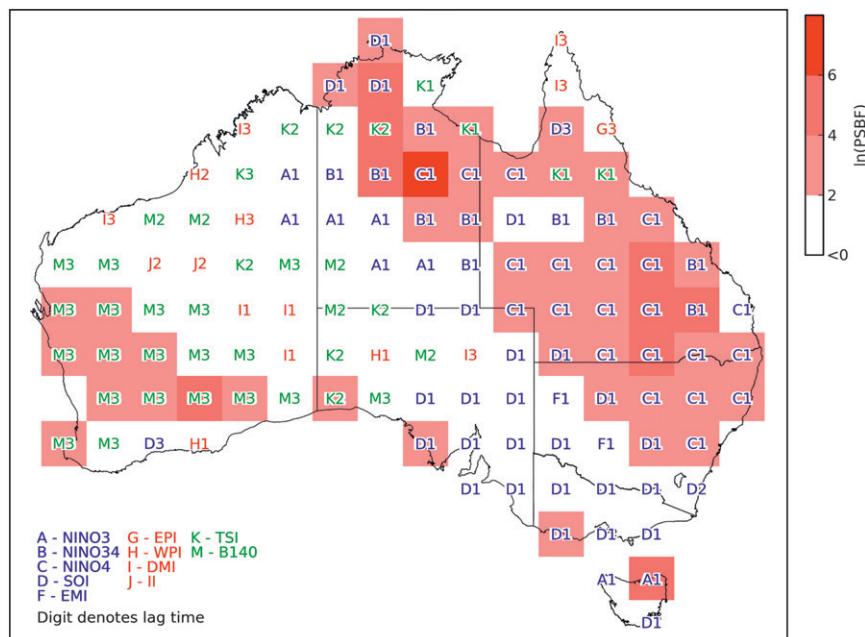


FIG. 12. As in Fig. 9, but for JJA rainfall.

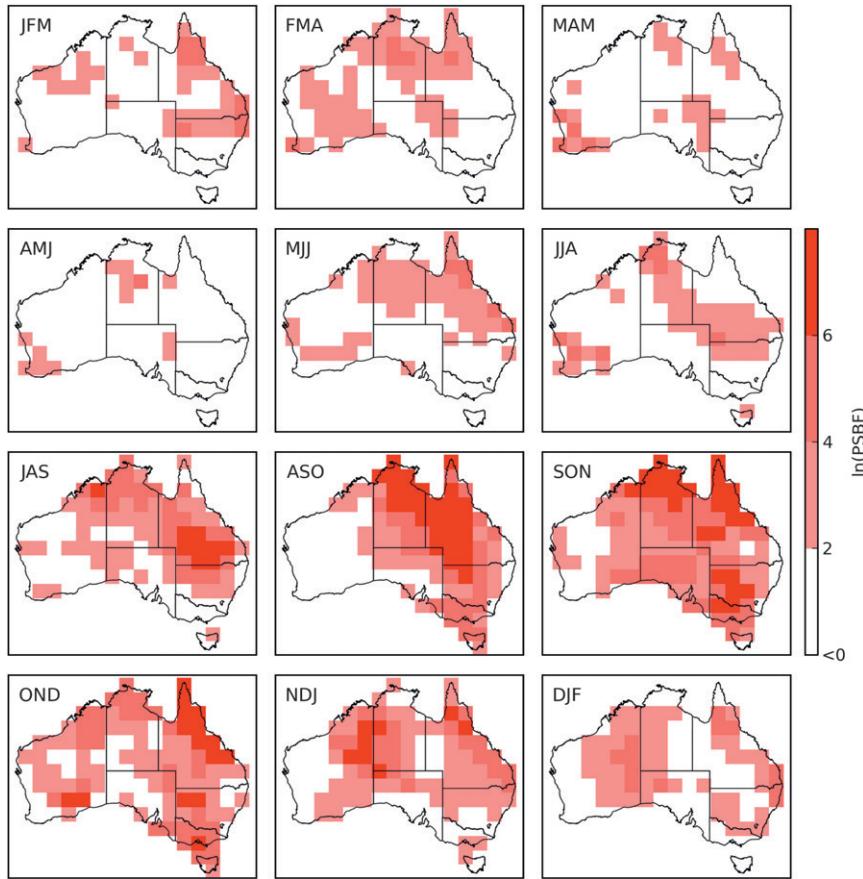


FIG. 13. Map of max  $\ln(\text{PsBF})$  showing the max evidence for forecasting Australian seasonal rainfall in each grid cell considering all lag times, 1950–79.

strong in these seasons, the evidence for the western Pacific region indices, Niño-3.4 and Niño-4, is also very strong in these seasons. Further analysis is required to ascertain whether the DMI and Niño indices provide unique information.

We note that some of the evidence for using lagged climate indices to forecast seasonal rainfall in a particular location may be directly attributable to long-term trends in the lagged climate indices and seasonal rainfalls. Given the large number of relationships investigated in this study, we leave detailed trend analysis to further study and discuss some cases where predictability may arise from trends. As discussed briefly in the results, there is strong evidence for using the lagged TSI to forecast SON, OND, and DJF rainfall in western Australia. As this relationship is poorly understood, we performed a brief analysis of the effect of trends by applying linear regression to remove trends in the September TSI and the seasonal rainfalls. While there was a reduction in the evidence, the index was still supported for forecasting rainfall in the next few seasons. Furthermore, a visual inspection of the September TSI

suggests that there was a step change in 1996–97 and linear trend removal may not be appropriate. Smith and Timbal (2011) suggested that a first-order difference method is appropriate for removing linear and nonlinear trends. In their study, they found that the concurrent relationship between DMI and southern Australian SON rainfall was reduced if long-term trends were removed using first-order differences. This has implications for our results, as we may see erosion of the evidence for using the 1-month-lagged DMI to forecast SON and OND rainfall in southeastern Australia after a removal of trends.

There is little evidence supporting the use of AAO and B140 for forecasting seasonal rainfall because the lagged relationships are weak or nonexistent. It may be possible to circumvent the lack of lagged relationships through the use of dynamical prediction systems. If climate indices are derived from dynamical forecasts of oceanic and atmospheric variables, then it may be possible to exploit concurrent relationships with rainfall by including the indices as predictor variables in statistical prediction systems. This will be the subject of

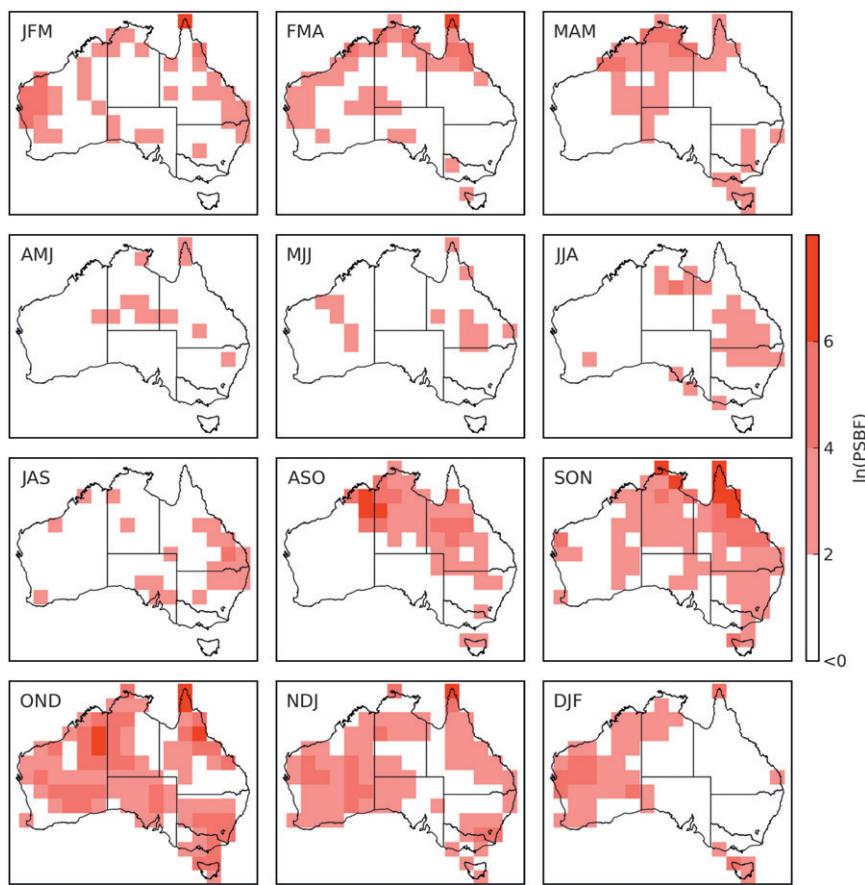


FIG. 14. As in Fig. 13, but for 1980–2009.

a future study in hybrid statistical–dynamical prediction systems.

In a given location, the evidence supporting the use of different climate indices may be similar. This is expected because many of the climate indices are correlated with one another. Another reason that climate indices may show similar evidence is that seasonal rainfall at a given location may be influenced by different components of the atmospheric circulation. From a purely statistical viewpoint, this notion leads to the issue of model uncertainty in the forecasting problem. In the future, we will report on the application of Bayesian model averaging (Hoeting et al. 1999) to the forecasting problem. In this study, we have focused on single predictor models. In future research, we will investigate the evidence for using multiple predictor models that account for interactions between variables to forecast Australian seasonal rainfall.

## 5. Conclusions

Statistical prediction systems will continue to play an important role in seasonal rainfall forecasting for the

foreseeable future. Lagged climate indices are potentially useful predictors of seasonal rainfall, particularly those derived from low-frequency oceanic–atmospheric processes. A Bayesian joint probability modeling approach was applied to find cross-validation predictive densities of the seasonal rainfalls in grid cells across all of Australia, using lagged climate indices as predictors. For each climate index, the evidence supporting its use as a predictor of seasonal rainfall, relative to a naïve no-predictor model, was quantified by the pseudo-Bayes factor.

The lagged climate indices with the strongest and most spatially extensive evidence for forecasting Australian seasonal rainfall are derived from oceanic and atmospheric variables in the Pacific region. There is evidence supporting the use of climate indices that are derived from oceanic variables in the Indian and extra-tropical regions for some regions and seasons. The spatial structure of evidence for forecasting seasonal rainfall varies with season. The strongest and most spatially extensive evidence is found from ASO to NDJ. The weakest and least spatially extensive evidence is found from MAM to May–July (MJ).

Climate indices that are lagged 1 month show the strongest evidence for forecasting seasonal rainfall and the strength of evidence decreases as lag time is increased. Evidence for climate indices derived from Pacific region sea surface temperatures shows the strongest persistence in the relationships with seasonal rainfall as lag time is increased. Climate indices from the Indian and extratropical region or derived from the higher-frequency atmospheric indices tended to show less persistence in the relationship.

The evidence supporting the use of climate indices to forecast Australian seasonal rainfall is not consistent across decades. Sub-period analysis reveals that the evidence is stronger for the period 1950–79 than 1980–2009. Therefore, climate variability at decadal and longer time scales can impact the usefulness of climate indices for forecasting Australian seasonal rainfall.

This study leads to opportunities for further research. We plan to assess the evidence for combining climate indices in multiple predictor models and report on the application of Bayesian model averaging to the forecasting problem. We also plan to report on the use, in the statistical prediction system, of predictions from dynamic climate models in combination with the lagged climate indices from observed data.

**Acknowledgments.** This research has been supported by the Water Information Research and Development Alliance between the Australian Bureau of Meteorology and CSIRO Water for a Healthy Country Flagship. We thank Neil Plummer, Dr. David Jones, and Dr. Andrew Watkins from the Australian Bureau of Meteorology, and Dr. Harry Hendon from the Centre for Australian Weather and Climate Research for valuable discussions. We would also like to thank Dr. Ian Smith, Dr. Enli Wang, and three anonymous reviewers for their reviews of early versions of this paper.

## REFERENCES

- Allan, R., and Coauthors, 2001: Is there an Indian Ocean dipole and is it independent of the El Niño-Southern Oscillation? *CLIVAR Exchanges*, No. 6, International CLIVAR Project Office, Southampton, United Kingdom, 18–22.
- Anderson, J., H. Van Den Dool, A. Barnston, W. Chen, W. Stern, and J. Poshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Amer. Meteor. Soc.*, **80**, 1349–1361.
- Ashok, K., Z. Guan, and T. Yamagata, 2003: A look at the relationship between the ENSO and the Indian Ocean dipole. *J. Meteor. Soc. Japan*, **81**, 41–56.
- , S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007, doi:10.1029/2006JC003798.
- Barnston, A. G., Y. He, and M. H. Glantz, 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Niño episode and the 1998 La Niña onset. *Bull. Amer. Meteor. Soc.*, **80**, 217–243.
- Chiew, F. H. S., T. C. Piechota, J. A. Dracup, and T. A. McMahon, 1998: El Niño/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. *J. Hydrol.*, **204**, 138–149.
- Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **6**, 249–266.
- Feng, J., J. Li, and Y. Li, 2010: A monsoonlike southwest Australian circulation and its relation with rainfall in southwest Western Australia. *J. Climate*, **23**, 1334–1353.
- Geisser, S., and W. Eddy, 1979: A predictive approach to model selection. *J. Amer. Stat. Assoc.*, **74**, 153–160.
- Gelfand, A. E., and D. K. Dey, 1994: Bayesian model choice: Asymptotics and exact calculations. *J. Roy. Stat. Soc.*, **56B**, 501–514.
- Goddard, L., S. Mason, S. Zebiak, C. Ropelewski, R. Basher, and M. Cane, 2001: Current approaches to seasonal to interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152.
- Halide, H., and P. Ridd, 2008: Complicated ENSO models do not significantly outperform very simple ENSO models. *Int. J. Climatol.*, **28**, 219–233.
- Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky, 1999: Bayesian model averaging: A tutorial. *Stat. Sci.*, **14**, 382–401.
- Jones, D. A., W. Wang, and R. Fawcett, 2009: High-quality spatial climate data-sets for Australia. *Aust. Meteor. Oceanogr. J.*, **58**, 233–248.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kass, R. E., and A. E. Raftery, 1995: Bayes factors. *J. Amer. Stat. Assoc.*, **90**, 773–795.
- Kirono, D. G. C., F. H. S. Chiew, and D. M. Kent, 2010: Identification of best predictors for forecasting seasonal rainfall and runoff in Australia. *Hydrolog. Processes*, **24**, 1237–1247.
- McBride, J. L., and N. Nicholls, 1983: Seasonal relationships between Australian rainfall and the Southern Oscillation. *Mon. Wea. Rev.*, **111**, 1998–2004.
- Meneghini, B., I. Simmonds, and I. N. Smith, 2007: Association between Australian rainfall and the Southern Annular Mode. *Int. J. Climatol.*, **27**, 109–121.
- Meyers, G., P. McIntosh, L. Pigot, and M. Pook, 2007: The years of El Niño, La Niña, and interactions with the tropical Indian Ocean. *J. Climate*, **20**, 2872–2880.
- Michaels, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.*, **26**, 1589–1600.
- Mo, K. C., 2000: Relationships between low-frequency variability in the Southern Hemisphere and sea surface temperature anomalies. *J. Climate*, **13**, 3599–3610.
- Murphy, B. F., and B. Timbal, 2008: A review of recent climate variability and climate change in southeastern Australia. *Int. J. Climatol.*, **28**, 859–879.
- Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. *J. Climate*, **19**, 3279–3293.
- Rajeevan, M., D. Pai, R. Anil Kumar, and B. Lal, 2007: New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dyn.*, **28**, 813–828.
- Risbey, J. S., M. J. Pook, P. C. McIntosh, M. C. Wheeler, and H. H. Hendon, 2009: On the remote drivers of rainfall variability in Australia. *Mon. Wea. Rev.*, **137**, 3233–3253.

- Robertson, D. E., and Q. J. Wang, 2012: A Bayesian approach to predictor selection for seasonal streamflow forecasting. *J. Hydrometeor.*, **13**, 155–171.
- Ruiz, J. E., I. Cordery, and A. Sharma, 2006: Impact of mid-Pacific Ocean thermocline on the prediction of Australian rainfall. *J. Hydrol.*, **317**, 104–122.
- Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360–363.
- Smith, I. N., and B. Timbal, 2011: Links between tropical indices and southern Australian rainfall. *Int. J. Climatol.*, **32**, 33–40. doi:10.1002/joc.2251.
- Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296.
- Stone, R. C., G. L. Hammer, and T. Marcussen, 1996: Prediction of global rainfall probabilities using phases of the Southern Oscillation Index. *Nature*, **284**, 252–255.
- Taschetto, A. S., C. C. Ummenhofer, A. Sen Gupta, and M. H. England, 2009: Effect of anomalous warming in the central Pacific on the Australian monsoon. *Geophys. Res. Lett.*, **36**, L12704, doi:10.1029/2009GL038416.
- Troup, A. J., 1965: Southern oscillation. *Quart. J. Roy. Meteor. Soc.*, **91**, 490–506.
- Verdon, D. C., and S. W. Franks, 2005: Indian Ocean sea surface temperature variability and winter rainfall: Eastern Australia. *Water Resour. Res.*, **41**, W09413, doi:10.1029/2004WR003845.
- Wang, G., and H. H. Hendon, 2007: Sensitivity of Australian rainfall to inter-El Niño variations. *J. Climate*, **20**, 4211–4226.
- Wang, Q. J., and D. E. Robertson, 2011: Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences. *Water Resour. Res.*, **47**, W02546, doi:10.1029/2010WR009333.
- , —, and F. H. S. Chiew, 2009: A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. *Water Resour. Res.*, **45**, W05407, doi:10.1029/2008WR007355.
- Wright, W. J., 1993: Seasonal climate summary southern hemisphere (autumn 1992): Signs of a weakening ENSO event. *Aust. Meteor. Mag.*, **42**, 191–198.