Dynamical, Statistical–Dynamical, and Multimodel Ensemble Forecasts of Australian Spring Season Rainfall

EUN-PA LIM, HARRY H. HENDON, DAVID L. T. ANDERSON, ANDREW CHARLES, AND OSCAR ALVES

Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, Australia

(Manuscript received 23 February 2010, in final form 30 September 2010)

ABSTRACT

The prediction skill of the Australian Bureau of Meteorology dynamical seasonal forecast model Predictive Ocean Atmosphere Model for Australia (POAMA) is assessed for probabilistic forecasts of spring season rainfall in Australia and the feasibility of increasing forecast skill through statistical postprocessing is examined. Two statistical postprocessing techniques are explored: calibrating POAMA prediction of rainfall anomaly against observations and using dynamically predicted mean sea level pressure to infer regional rainfall anomaly over Australia (referred to as ''bridging''). A ''homogeneous'' multimodel ensemble prediction method (HMME) is also introduced that consists of the combination of POAMA's direct prediction of rainfall anomaly together with the two statistically postprocessed predictions.

Using hindcasts for the period 1981–2006, the direct forecasts from POAMA exhibit skill relative to a climatological forecast over broad areas of eastern and southern Australia, where El Niño and the Indian Ocean dipole (whose behavior POAMA can skillfully predict at short lead times) are known to exert a strong influence in austral spring. The calibrated and bridged forecasts, while potentially offering improvement over the direct forecasts because of POAMA's ability to predict the main drivers of springtime rainfall (e.g., El Niño and the Southern Oscillation), show only limited areas of improvement, mainly because strict cross-validation limits the ability to capitalize on relatively modest predictive signals with short record lengths. However, when POAMA and the two statistical–dynamical rainfall forecasts are combined in the HMME, higher deterministic and probabilistic skill is achieved over any of the single models, which suggests the HMME is another useful method to calibrate dynamical model forecasts.

1. Introduction

Over the last two decades significant effort has been focused on developing seasonal forecast systems based on coupled general circulation models (e.g., Stockdale et al. 1998; Palmer et al. 2004; Troccoli et al. 2008; Stockdale et al. 2010). The strength of dynamical coupled models stems from their ability to simulate and predict the main modes of atmosphere–ocean interactions that are important for seasonal climate variability [e.g., El Niño–Southern Oscillation (ENSO)]. Also, they explicitly incorporate nonlinear processes in the atmosphere and ocean, so that climate evolution, which is chaotic in nature, can be captured in the probabilistic domain through multimember or multimodel ensemble forecasts. The merit of dynamical models is especially important for climate prediction over subtropical and extratropical regions where climate is governed by internal dynamics as much as by the lower boundary forcing (i.e., sea surface temperature), which makes skillful prediction with a statistical forecast system more difficult in these regions compared to the tropics.

The Australian Bureau of Meteorology (BOM) together with the Commonwealth Scientific and Industrial Research Organization (CSIRO) has been developing a coupled atmosphere–ocean climate prediction model, Predictive Ocean Atmosphere Model for Australia (POAMA) that is based on atmosphere and ocean general circulation models. The current operational version of POAMA (v1.5b) demonstrates internationally competitive skill in predicting ENSO and the Indian Ocean Dipole mode (IOD; Wang et al. 2008; Hendon et al. 2009; Zhao and Hendon 2009), which are two of the most important influences on Australian climate in winter and spring (e.g., McBride and Nicholls 1983; Nicholls 1989; Drosdowsky

Corresponding author address: Eun-Pa Lim, CAWCR, Bureau of Meteorology, GPO Box 1289K, Melbourne, VIC 3001, Australia.
E-mail: e.lim@bom.gov.au

DOI: 10.1175/2010MWR3399.1

and Chambers 2001; Meyers et al. 2007). ENSO is known to affect northeastern Australian rainfall directly through changes in the Southern Oscillation (e.g., northeastern Australia experiences anomalously high surface pressure during El Niño; McBride and Nicholls 1983). On the other hand, IOD affects rainfall across the southwest and the southeast of Australia through excitation of equivalent–barotropic Rossby wave trains that are forced by the variations of convective heating in the equatorial Indian Ocean. For instance during a positive IOD event, a Rossby wave train emanates from the suppressed convection in the eastern Indian Ocean, which results in an anticyclonic pressure–geopotential height anomaly across southern portions of Australia (e.g., Saji et al. 2005). This anticyclonic anomaly acts to deflect and weaken rain-bearing weather systems across the southern portions of the continent (e.g., Ummenhofer et al. 2009). The winter rainfall variance in the southern part of Australia that is explained by IOD is mostly independent of ENSO (e.g., Risbey et al. 2009); whereas, the spring rainfall variance explained by IOD is strongly associated with ENSO as well because IOD and ENSO highly covaries in the spring season (Lim et al. 2009).

The fact that Australian rainfall variation is strongly associated with some important modes of tropical SST variability such as ENSO and IOD and POAMA can predict these modes with good skill implies some useful skill for prediction of Australian springtime climate. The first aim of this study is, therefore, to assess the skill of POAMA probabilistic forecasts of springtime rainfall variability, using a retrospective forecast set for the period 1981–2006.

An extension to purely dynamical prediction is to exploit dynamically predicted climate components in a statistical model (e.g., Feddersen et al. 1999). Two approaches are possible. The first, statistical calibration, attempts to overcome systematic bias of predicted patterns of climate variability, for instance the pattern of rainfall anomaly expected during El Niño, by adjusting the predicted spatial patterns of variability toward its observed counterpart (Ward and Navarra 1997; Feddersen et al. 1999; Mo and Straus 2002; Kang et al. 2004). The second, known as statistical bridging, is the prediction of a variable such as rainfall through its statistical relationship with other model-predicted large-scale climate components such as temperature, pressure, or wind fields (Voldoire et al. 2002; Tippett et al. 2005; Lin and Derome 2005; Kang et al. 2007). This approach attempts to take advantage of the model's ability to predict large-scale patterns of variability, such as the Southern Oscillation, in order to overcome the model's inability to faithfully simulate the associated regional rainfall variability. In this study we test the feasibility of improving regional forecast skill of Australian rainfall through these two distinct methods of statistical postprocessing—calibrating the patterns of direct rainfall predictions from POAMA against observations and using predicted mean sea level pressure (MSLP) as an atmospheric bridge of the lower boundary forcing to rainfall.

Finally, we explore a "homogeneous" multimodel ensemble (HMME) prediction method[1] by combining the direct forecasts from POAMA and the forecasts from the two statistical–dynamical prediction schemes in order to seek improved rainfall forecast skill over Australia. The motivation for exploring the HMME is because the overall performance of a multimodel ensemble system is, in general, known to be better than the individual component models as a result of offsetting errors and increasing ensemble spread (Doblas-Reyes et al. 2000; Zebiak 2003; Hagedorn et al. 2005; Weigel et al. 2008). We indeed show that the HMME is more skilful than the individual component models in regard to forecast reliability, correlation, and error assessments of probabilistic and deterministic forecasts because there is independent information in each of the component models. Hence, the HMME approach can improve the prediction skill in a fashion similar to that achieved by a conventional multimodel ensemble based on different dynamical models, but in a more cost-effective way.

Details of the POAMA system and verification methods are described in section 2, and the skill of probabilistic forecasts for above median rainfall from POAMA is analyzed in section 3. In section 4 statistical–dynamical models are discussed and the skill from these models is compared to that of the direct predictions from POAMA. This will be followed by the skill assessment of the HMME in section 5. Concluding remarks are given in section 6.

2. Description of the POAMA system and verification methods

The POAMA version 1.5b system consists of the BOM Atmospheric Model version 3 (BAM3; Colman et al. 2005) and the Australian Community Ocean Model version 2 (ACOM2; Schiller et al. 2002; Oke et al. 2005). The horizontal structure of BAM3 is represented by spherical harmonics with a triangular truncation at wavenumber 47 (denoted T47, which has approximately 250-km resolution), and the vertical variation is represented by 17 sigma levels. ACOM2 has a zonal resolution of 2° longitude and a telescoping meridional resolution of 0.5° latitude within 8° latitude of the equator, gradually changing to

[1] This terminology was given as the individual models participating in this system utilize the outputs of one model, POAMA.

1.5° latitude near the poles. Vertically, ACOM2 has 25 levels, with 12 levels in the top 185 m. The atmosphere and ocean models are coupled every 3 h by the Ocean Atmosphere Sea Ice Soil (OASIS) coupling software (Valke et al. 2000).

POAMA forecasts are initialized with observed atmospheric and oceanic conditions. The atmospheric initial conditions are provided by a new Atmosphere and Land Initialization (ALI) scheme (Hudson et al. 2010). ALI nudges zonal and meridional winds, temperature, and humidity in BAM3 to those of the reanalyses from 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; Uppala et al. 2005) during 1980–2001 and to the global analyses from the BOM's numerical weather prediction system (GASP) during 2002–06 for hindcasts. Likewise, ALI nudges to GASP for the real-time forecasts. The initial conditions produced from ALI are similar to the analyses of ERA-40/GASP but result in less initial forecast shock than if the ERA-40/GASP analyses were directly used as initial conditions. Land-surface conditions are initialized indirectly in response to the nudged atmospheric conditions.

The ocean data assimilation system provides an estimate of the state of the upper ocean based on the optimum interpolation (OI) of available subsurface temperature observations (Smith et al. 1991), together with a strong relaxation of the SST to observed analyses. Further details of the POAMA system can be found in Hudson et al. (2010) and at http://poama.bom.gov.au/documentation/index.htm.

For this study, a 10-member ensemble of hindcasts for the period of 1981–2006 from POAMA v1.5b is used. The 10 ensemble members were generated by perturbing the atmospheric initial conditions by using 6-h consecutively earlier analyses from the first day of each month. Identical ocean initial conditions were used for all 10 members. We assess the skill of probabilistic forecasts (based on the 10 ensemble members) of above median rainfall, which follows the format of the current operational seasonal forecasts issued at BOM. Forecast probability is simply the fraction of the total ensemble members that predicts above median rainfall. We also examine the skill of deterministic forecasts for predicting rainfall anomaly based on the ensemble mean. We limit our interest to rainfall forecasts in austral spring [September–November (SON)] when skillful rainfall prediction would be greatly valued for water management and agriculture in Australia. We verify against observed behavior based on the monthly gridded analyses of Australian rainfall that are produced by the National Climate Centre (0.25° latitude × 0.25° longitude; Jones and Weymouth 1997).

Rainfall anomalies from the monthly climatologies of both the dynamical forecasts and verification data were seasonally averaged and then normalized by their respective standard deviations at each grid point over Australia. The monthly climatology of the dynamical forecasts was a function of forecast start month and lead time. By forming anomalies relative to the model's climatology and standardizing the anomalies relative to the model's variability, some aspects of the mean model bias are removed and the hindcast rainfall is ''calibrated'' to have the same standard deviation as observed (Stockdale et al. 1998; Weigel et al. 2008). Hereafter, POAMA rainfall anomalies calculated in this manner are referred to as direct rainfall prediction from POAMA.

To reduce as much as possible the introduction of artificial skill, especially in consideration of the relatively short record length of available hindcasts (i.e., 26 yr), all calculations in this study were cross validated. Leave-one-out cross validation was performed as follows: 1) One year of rainfall data was left out; 2) the climatology and standard deviation were obtained with the remaining data (training data) for both model and observation, and among the resultant rainfall anomalies the forecast and observed medians were found; 3) the standardized rainfall anomalies (predicted and observed) in the year left out were calculated based on the climatology and standard deviation obtained in the training period; and 4) the probabilistic forecast of above median rainfall with 10 ensemble members was verified. This process was repeated throughout the entire period. Cross validation of the statistical calibration and bridging methods are described in section 4.

3. POAMA forecasts

The attributes diagram (Wilks 2006; Palmer et al. 2008) is widely used as a compact way of displaying many of the features of probabilistic forecasts. Figure 1a displays the attributes diagram of POAMA forecasts of rainfall being above the median for all grid points over Australia in SON at lead time[2] (LT) of zero month based on the entire 26-yr record of hindcasts. In Fig. 1a, the forecasts falling in the gray areas are considered to be reliable forecasts as they correctly indicate the occurrence/nonoccurrence of the exceedance of median rainfall and have smaller magnitudes of error than a climatological forecast. The overall distribution of POAMA probabilistic forecasts

[2] The period of time between the issue time of the forecast and the beginning of the forecast validity period (WMO users guide, http://www.bom.gov.au/wmo/lrfvs/users.shtml). In this study, lead time zero month means that SON forecasts were initialized on 1 September.

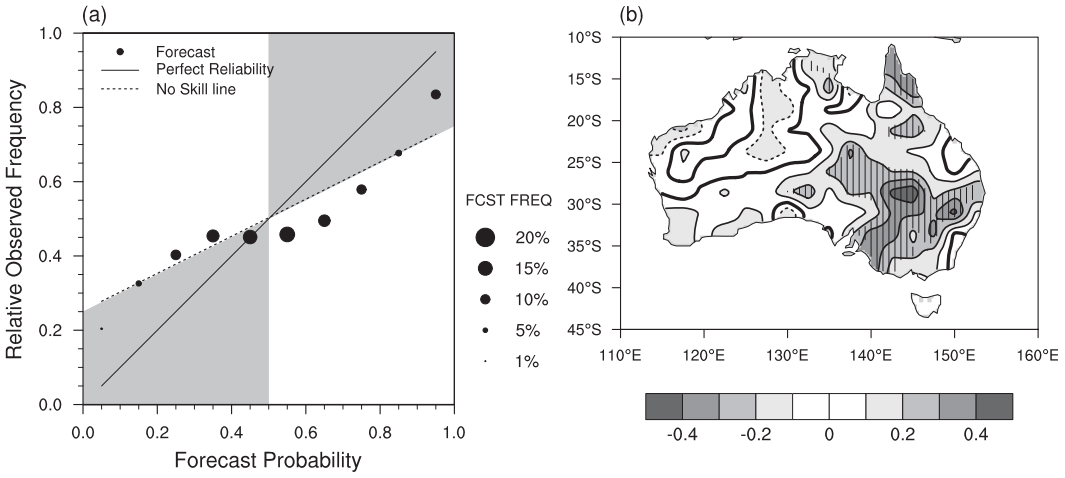

FIG. 1. (a) Attributes diagram of POAMA prediction of above median rainfall in austral spring (SON) at lead time 0 (LT 0), using 26 yr of hindcasts (1981–2006) over all grid points of Australia. The forecasts in the gray areas are considered to be reliable forecasts. (b) BSS of POAMA forecasts of exceeding median rainfall in the hindcast period, taking a climatological forecast as a reference forecast system. The contour interval is 0.1. The solid (dash) contours indicate positive (negative) changes made by POAMA compared to the climatological forecast. Gray shadings denote magnitudes of change. Hatching indicates the areas of statistically significant improvements by POAMA in the BSS, tested by bootstrapping 2000 samples and rejecting a null hypothesis (H_0: BSS $\leq$ 0) at the 10% level.[3]

suggests that the forecasts tend to be overconfident, and the reliability of most forecasts is far from being perfect. For predicted median rainfall with 50%–80% probabilities, the magnitude of error is larger than that of a climatological forecast. On the other hand, forecasts for greater than 90% chance of exceeding the median are highly reliable, and notably, such emphatic forecasts are not rare in POAMA (9% of the total forecast population in SON). In the range of 0%–40% probabilities, the magnitude of error is similar to a climatological forecast. In contrast to the poor-to-moderate forecast reliability, forecast resolution is good as indicated by different probability forecasts being followed by different observed outcomes. For instance, in Fig. 1a the relative observed rainfall frequency is very different when 10% chance of rainfall exceeding the median is predicted, compared to when 90% chance of rainfall exceeding the median is predicted.

Skill of probabilistic forecasts is also assessed using the Brier Score (BS), which is the mean squared error of probability forecasts against the occurrence of an event [Wilks 2006, his Eq. (7.34)]:

$$\mathrm{BS} = \frac{1}{n}\sum_{k=1}^{n}(F_k - O_k)^2, \tag{1}$$

where n is the total number of the forecasts (in our case 26 yr), F_k is the forecast probability for rainfall being above the median, and O_k is the observed rainfall occurrence in the kth year. Here O_k is 1 if the observed rainfall is above the median, and 0 if it is not. The Brier Skill Score (BSS) relative to a climatological forecast is then computed, which is the fractional change of the BS of the POAMA forecasts compared to the BS of the climatological forecast:

$$\mathrm{BSS}_{\mathrm{poama}} = 1 - \frac{\mathrm{BS}_{\mathrm{poama}}}{\mathrm{BS}_{\mathrm{clim}}}. \tag{2}$$

The BS of the climatological forecast ($\mathrm{BS}_{\mathrm{clim}}$) was computed with a sampling error correction term that adds sampling uncertainty to $\mathrm{BS}_{\mathrm{clim}}$ to make a fair comparison with the BS of POAMA whose probabilities are imperfectly estimated with the limited number of samples (i.e., 10 ensemble members; Weigel et al. 2007; Mason and Stephenson 2008). This correction term was simply $\mathrm{BS}_{\mathrm{clim}}$ divided by the ensemble size of POAMA in this case.

The BSS for springtime rainfall at LT 0 (Fig. 1b) demonstrates up to 20%–40% improvement of the dynamical forecasts over the climatological forecast across a broad region of eastern and southern Australia, where rainfall is strongly influenced by ENSO and IOD. On the

[3] Bootstrapping was performed by randomly resampling the data with replacement. The BSS at each grid point was calculated with the resampled data and this process was repeated 2000 times. Our null hypothesis BSS $\leq$ 0 (i.e., no skill improvement made by a forecast system compared with a reference forecast system) was rejected if equal to or less than 10% of the recalculated 2000 BSSs satisfied the hypothesis. This test is done on each grid point.

other hand, the performance of POAMA is no better than a climatological forecast in the west.

4. Statistical calibration and bridging

We now explore the benefits of two statistical techniques to improve forecast skill: calibrating the direct rainfall predictions at each grid point and bridging POAMA-predicted MSLP over the Southern Hemisphere (SH) extratropical region (25°–75°S) to rainfall to make predictions. The rationale for the calibration scheme is that direct forecasts of rainfall from POAMA may suffer from systematic bias as a result of, for instance, a systematic bias in the rainfall teleconnections to Australia associated with ENSO (e.g., misplacement of rainfall component related to ENSO; Lim et al. 2009). Such a bias can be corrected, in principal, a posteriori. The rationale for the bridging scheme is that POAMA may have some skill in predicting the large-scale variations of circulation that exert a strong control over local rainfall.

It is an important premise in statistical–dynamical modeling that a dynamical model should have skill in predicting the variable that is used as a predictor because statistical postprocessing cannot generate skill. As demonstrated in the previous section, POAMA can directly predict Australian rainfall with some skill, which qualifies POAMA rainfall forecasts to be used in the calibration model. POAMA also has skill in predicting SH extratropical MSLP, which will be discussed later.

a. Common techniques

The basic algorithm is the same in both calibration and bridging models. First, the components of the predictors that covary with observed Australian rainfall are identified by using the singular value decomposition analysis (SVDA) technique (also known as maximum covariance analysis; Bretherton et al. 1992; Ward and Navarra 1997). This technique expands a predictor field $\mathbf{X}(i, t)$ and a predictand field $\mathbf{Y}(j, t)$ in terms of spatial patterns $\mathbf{G}(i, m)$ (right singular vector) and $\mathbf{H}(j, m)$ (left singular vector), respectively, which maximize the covariance between the two fields. The expansion coefficients, $u(m, t)$ and $v(m, t)$, of those patterns are derived as

$$\mathbf{X}(i, t) \approx \sum_{m=1}^{M} u(m, t)\mathbf{G}(i, m) \tag{3}$$

$$\mathbf{Y}(j, t) \approx \sum_{m=1}^{M} v(m, t)\mathbf{H}(j, m), \tag{4}$$

where i and j indicate the spatial coordinates of $\mathbf{X}$ and $\mathbf{Y}$, respectively, t is a time coordinate, and m and M represent SVD modes. For the calibration scheme, $\mathbf{X}$ is the predicted standardized rainfall anomaly from POAMA. For the bridging scheme, $\mathbf{X}$ is the MSLP anomaly predicted by POAMA. In both cases, $\mathbf{Y}$ is observed standardized rainfall anomaly.

Second, we linearly regress the predictand field $\mathbf{Y}(j, t)$ onto the time series of expansion coefficients of a selected number of SVD modes of the predictor $[u(m, t)]$, using multiple linear regression; that is,

$$\hat{\mathbf{Y}}(j, t) \approx \sum_{m=1}^{M} \mathbf{A}(j, m)u(m, t), \tag{5}$$

where $\hat{\mathbf{Y}}(j, t)$ is the predicted $\mathbf{Y}(j, t)$ in the regression model, and $\mathbf{A}(j, m)$ is a set of regression coefficients that minimize the expected root-mean-squared difference between $\hat{\mathbf{Y}}(j, t)$ and $\mathbf{Y}(j, t)$. In this step, we need to set M, the number of retained SVD modes to include as predictors in this regression model. To avoid overfitting, we require no more than 10 degrees of freedom ($t - M - 1 \geq 10$), with at least five data points per predictor ($M \leq t/5$) (Mo and Straus 2002).

In constructing the statistical–dynamical schemes, we used all 10 ensemble members of forecasts by concatenating them and forming one long time series of $\mathbf{X}$ while repeating the predictand time series 10 times to form a compatible long time series of $\mathbf{Y}$ (Ward and Navarra 1997).

b. Calibration

The leading five SVD modes, which explain 96% of the covariance between dynamically predicted and observed Australian rainfall anomalies, were retained in our statistical calibration model. The first two SVD patterns of dynamically predicted rainfall are displayed in Figs. 2a,c, and the correlations of their expansion coefficients with observed rainfall are displayed in Figs. 2b,d. POAMA's rainfall anomaly pattern that dominates the covariance with the observed rainfall (83%) has a large loading over eastern Australia (Fig. 2a). This forecast rainfall pattern is related to the forecast SST pattern that has a maximum anomaly over the central tropical Pacific flanked by the opposite signed anomalies in the eastern Pacific and the western Pacific–eastern Indian oceans. According to the observation, in the spring this tripole SST structure over the tropical Pacific (referred to as ENSO Modoki, central Pacific ENSO or warm-pool ENSO; e.g., Hendon et al. 2009) and associated variations of convection near the date line significantly impact the rainfall variation in the northwest and northeast of Australia (Wang and Hendon 2007; Lim et al. 2009). Also, in spring, the variability of the eastern Indian Ocean SST (which covaries but with opposite sign with

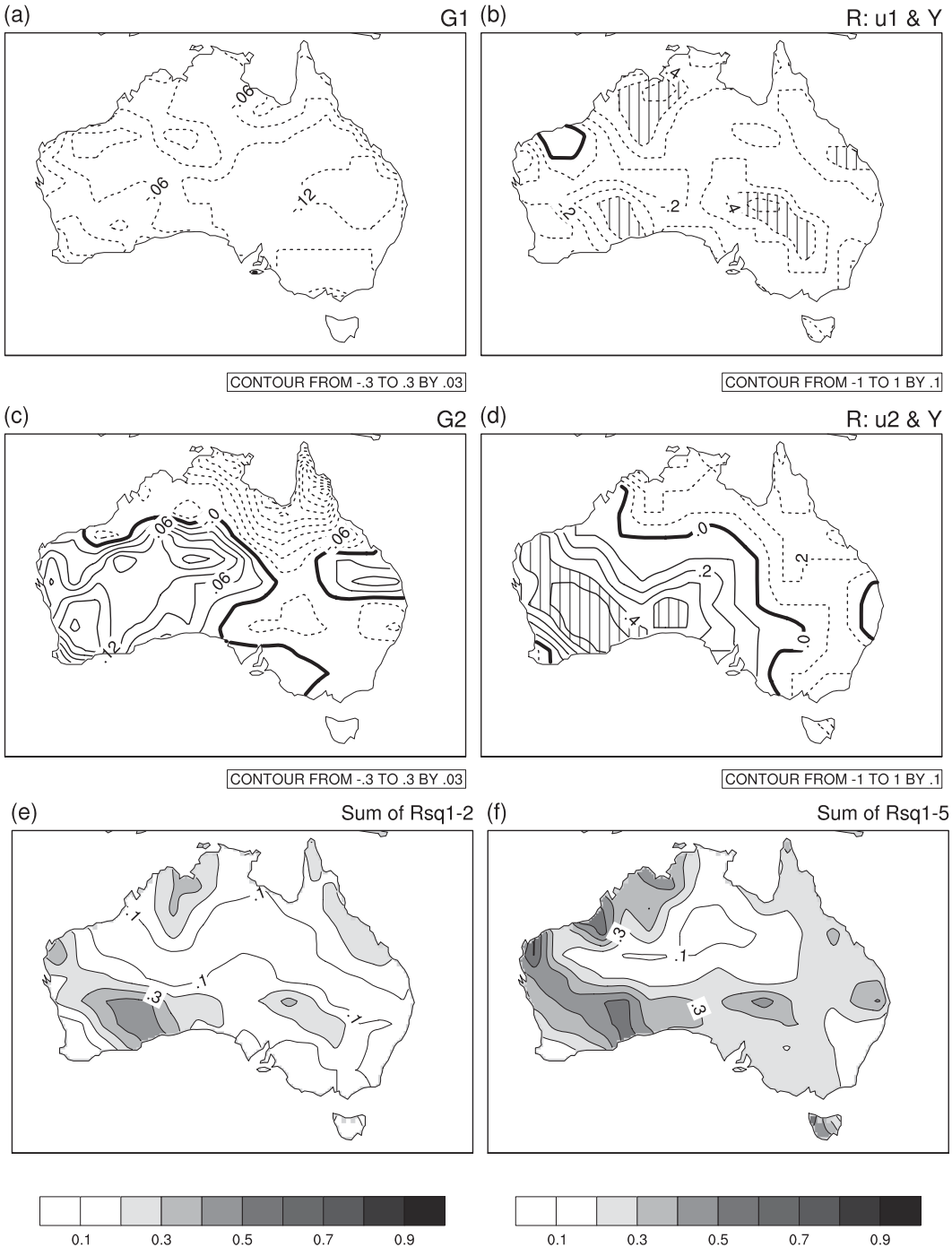

FIG. 2. (a),(c) The leading two SVD right vectors of the calibration scheme (POAMA predicted rainfall components that most covaries with observed rainfall). The contour interval is 0.03. (b),(d) Correlation of observed rainfall with the expansion coefficient of the SVD right vector shown in (a) and (c), respectively. The contour interval is 0.1, and the areas of the statistically significant correlation at the 5% level (two-tailed Student's *t* test with 26 samples) are hatched. (e),(f) Explained variance of observed rainfall by the expansion coefficients of the first two and the first five SVD right vectors, respectively. The contour interval is 0.1, and gray shading starts from 0.2.

anomalies in the western Indian Ocean and in the equatorial central to eastern Pacific Ocean) excites an equivalent-barotropic Rossby wave train that produces a pressure/geopotential height anomaly over the Great Australian Bight, thereby affecting rainfall across southern portions of the continent (Saji et al. 2005; Lim et al. 2009; Ummenhofer et al. 2009). Therefore, it appears that our calibration technique can capture the rainfall signal associated with the tripole SST pattern across the eastern Pacific to the eastern Indian Ocean in the dynamical model and transfer the information to the correct locations as shown in Fig. 2b.

The second SVD pattern of POAMA rainfall (Fig. 2c) that accounts for 7% of the covariability with the observed rainfall is associated with the model's simulations of traditional ENSO from which the maximum SST anomaly is found in the far eastern Pacific together with the opposite-signed anomaly in the far western Pacific. The overall pattern of the model's ENSO-related rainfall component is similar to the observed rainfall pattern associated with ENSO in spring, but the model's rainfall responses are shifted northeast compared to its observed counterpart (Figs. 2c,d).

The first two SVD time series of dynamically predicted rainfall account for up to 40% of observed rainfall variance in the southern part of Western Australia (Fig. 2e) but not much elsewhere. Additional three higher SVD modes account for 20%–30% of the rainfall variance of eastern Australia (Fig. 2f). The larger amount of explained variance of observed rainfall in the west implies higher predictability in the west than in the east by this calibration scheme.

c. *Bridging*

For statistical bridging, we have chosen extratropical MSLP as a predictor variable because MSLP is directly associated with rainfall (more than, e.g., mid- or upper-tropospheric geopotential heights with Australian rainfall). Also, MSLP carries information of both tropical and extratropical climate modes such as ENSO, IOD, and Southern Annular Mode (SAM) that are important drivers of Australian spring rainfall (e.g., McBride and Nicholls 1983; Hendon et al. 2007; Lim et al. 2009; Risbey et al. 2009).

Our predictor field $\mathbf{X}$ in the statistical bridging scheme is a dynamically predicted MSLP anomaly (hereafter referred to as MSLP) prefiltered by projecting the POAMA forecasts onto the leading 10 empirical orthogonal function modes of observed MSLP. Prefiltering of MSLP data with EOF analysis reduces the number of degrees of freedom in the MSLP field, and therefore, can result in more stable SVD modes (Barnett and Preisendorfer 1987; Bretherton et al. 1992).

The 10 leading EOFs of observed MSLP (weighted by square root of cosine latitude) explain 90% of the total MSLP variability in the SON season in the SH extratropics (25°–75°S) according to our analysis on the National Centers for Environmental Prediction (NCEP)–Department of Energy (DOE) reanalysis II data (Kanamitsu et al. 2002). The first four observed EOF modes of MSLP explain about 70% of the total variance and are displayed in Fig. 3. The signs of the EOF patterns were determined to be consistent with positive rainfall anomaly over Australia. The first observed EOF mode (Fig. 3a) represents SAM (Thompson and Wallace 2000; Marshall 2003). We note that the positive phase of SAM is weakly correlated with La Niña during austral spring (Table 1; also see, e.g., Silvestri and Vera 2003).

The second observed EOF mode of MSLP shows the wave train pattern that is commonly referred to as the Pacific–South American pattern (PSA). The PSA has been suggested to result from energy propagation from the tropics especially in response to equatorial Pacific SST variability during ENSO (Karoly 1989; Mo 2000; Kiladis and Mo 1998). In the austral spring, the western Pacific Ocean and the eastern Indian Ocean SST anomalies highly covary with the same sign (but oppositely signed to the eastern Pacific SST anomaly) during ENSO and serve as the direct source for the excitation of an equivalent-barotropic Rossby wave propagating from the Maritime Continent–eastern Indian Ocean region to the southeastern end of the Pacific Ocean as displayed in Fig. 3b. Such association between PSA and the western Pacific–eastern Indian Ocean SSTs during ENSO is confirmed by the strong negative correlation of the time series of the second EOF (PC2) with the Niño-3 index and with the IOD index (DMI; Saji et al. 1999) in Table 1. The high correlation of PC2 with DMI mostly stems from its correlation with the eastern pole of DMI (10°S–0°, 90°–110°E; $r \sim 0.65$).

The positive loading of the third observed EOF mode of MSLP shows some association with the ENSO events whose maximum cold SST anomaly occurs over the central Pacific as captured by the Niño-4 index or the El Niño Modoki Index (EMI; Ashok et al. 2007, 2009). The correlation of PC3 with the Niño-4 index and EMI is approximately -0.5. The fourth EOF mode of MSLP also has a moderate relation with the Niño-3 index, EMI, and DMI (Table 1).

The first 10 leading EOF modes of observed MSLP account for more than 40% of the observed rainfall variance over most locations across Australia and up to 70% in the northwest (Fig. 4a). Figure 4b shows that the first four EOFs of MSLP account for most of the rainfall variance in the northern and western parts of the country, whereas Fig. 4c shows that the rainfall variance in the

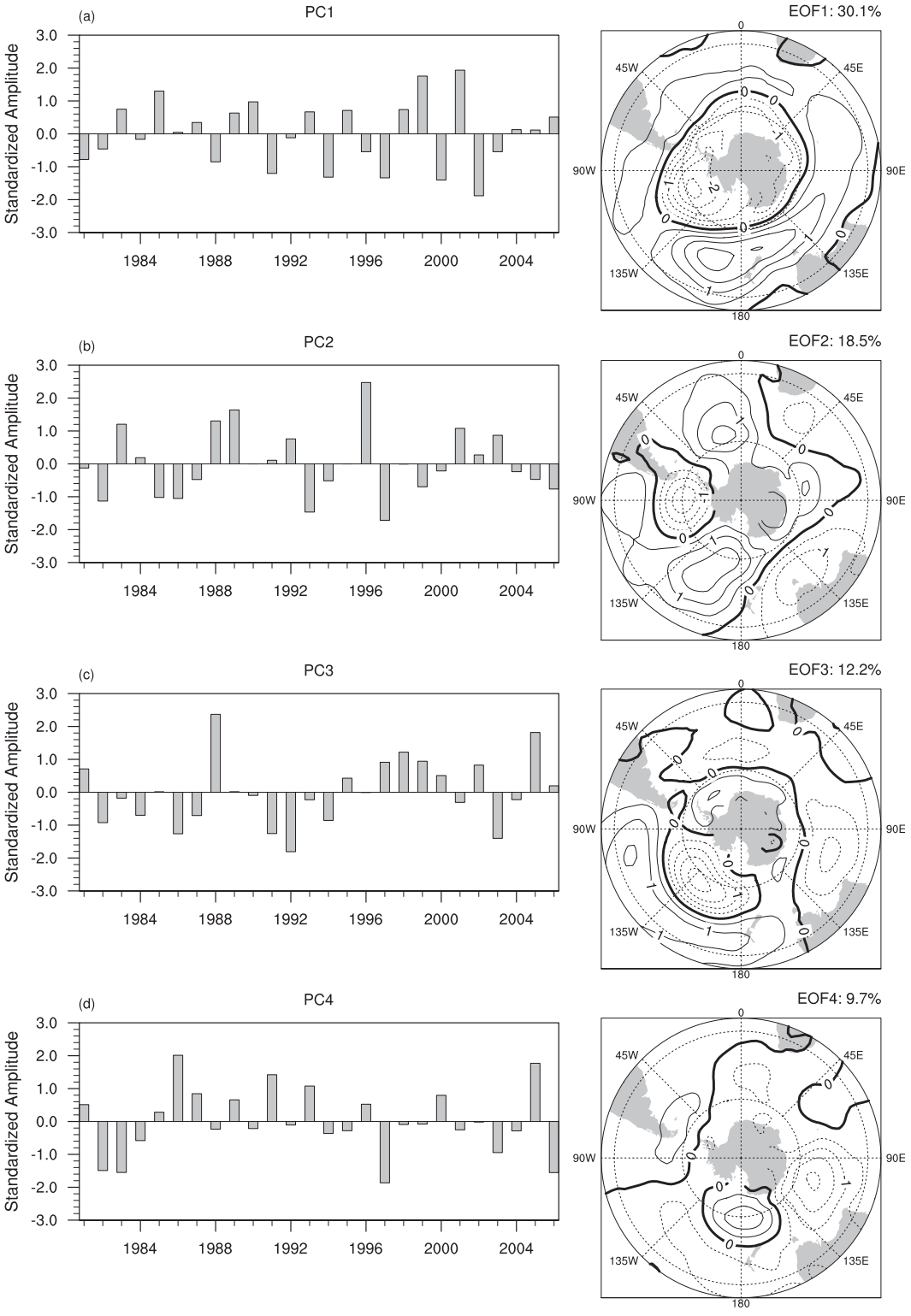

FIG. 3. (left) First four standardized principal component (PC) time series of observed MSLP and (right) the corresponding EOF patterns obtained by regression of observed MSLP field onto the standardized PCs in the period of 1981–2006 from the NCEP–DOE reanalysis II dataset. The domain is 25°–75°S. The contour interval is 0.5 hPa. The solid (dashed) contours indicate positive (negative) loadings. The percentage at the top of each map is MSLP variance accounted for by each mode.

TABLE 1. Correlation of the first four principal component (PC) time series of observed MSLP and Indo-Pacific SST indices depicting ENSO (NIÑO-3, NIÑO-4, EMI) and the DMI in the period 1981–2006. Correlation coefficient r greater than 0.38 is statistically significant at the 5% level (boldfaced) as estimated by a two-tailed Student's t test, assuming 26 degrees of freedom.

	NIÑO-3[a]	NIÑO-4[b]	EMI[c]	DMI[d]
PC1	**−0.40**	−0.27	−0.06	−0.33
PC2	**−0.50**	**−0.41**	−0.19	**−0.62**
PC3	−0.30	**−0.49**	**−0.49**	−0.13
PC4	−0.31	−0.01	0.37	−0.36

[a] NIÑO-3 index = $\overline{\text{SST}}_{(5°\text{S–}5°\text{N},90°\text{–}150°\text{W})}$
[b] NIÑO-4 index = $\overline{\text{SST}}_{(5°\text{S–}5°\text{N},160°\text{E–}150°\text{W})}$
[c] EMI = $\overline{\text{SST}}_{(10°\text{S–}10°\text{N},165°\text{E–}140°\text{W})} - 0.5 \times \overline{\text{SST}}_{(15°\text{S–}5°\text{N},70°\text{–}110°\text{W})} - 0.5 \times \overline{\text{SST}}_{(10°\text{S–}20°\text{N},125°\text{–}145°\text{E})}$
[d] DMI = $\overline{\text{SST}}_{(10°\text{S–}10°\text{N},50°\text{–}70°\text{E})} - \overline{\text{SST}}_{(10°\text{S–}0°,90°\text{–}110°\text{E})}$

southern and eastern parts is largely explained by the higher-order EOFs (5–10). To estimate the upper limit of predictability of rainfall by using EOFs of MSLP as predictors, we recalculate the rainfall variance explained by the observed MSLP EOFs in a cross-validated fashion. That is, we leave out one year of data, compute the EOFs of MSLP, and develop the regression of the PCs with observed rainfall on the remaining years, then apply this regression model to predict the rainfall in the left out year. We then repeat this over all years. The rainfall variance explained by MSLP PCs with this assumption of perfect forecasts of MSLP PCs (Figs. 4d–f) reduces by 10%–20% in the central eastern part of Australia because of the large year-to-year variations of the higher EOF modes of MSLP variability as indicated by the comparison between Figs. 4c and 4f. In contrast, predictability of the rainfall in western Australia remains high (40%–70%) if MSLP is perfectly predicted.

Because this bridging technique uses prefiltered MSLP predicted by POAMA as a predictor field, it is a natural first step to examine if POAMA has skill to predict the prefiltered MSLP. Prediction of prefiltered MSLP is carried out by filtering POAMA MSLP forecasts through the observed 10 EOF patterns of MSLP and obtaining the resultant expansion coefficients (viz., predicted PCs). The correlation of the observed PCs with the predicted PCs of MSLP is provided in Table 2. It is encouraging to see that in the austral spring season, SAM (PC1) is skillfully predicted by POAMA ($r \sim 0.7$) and

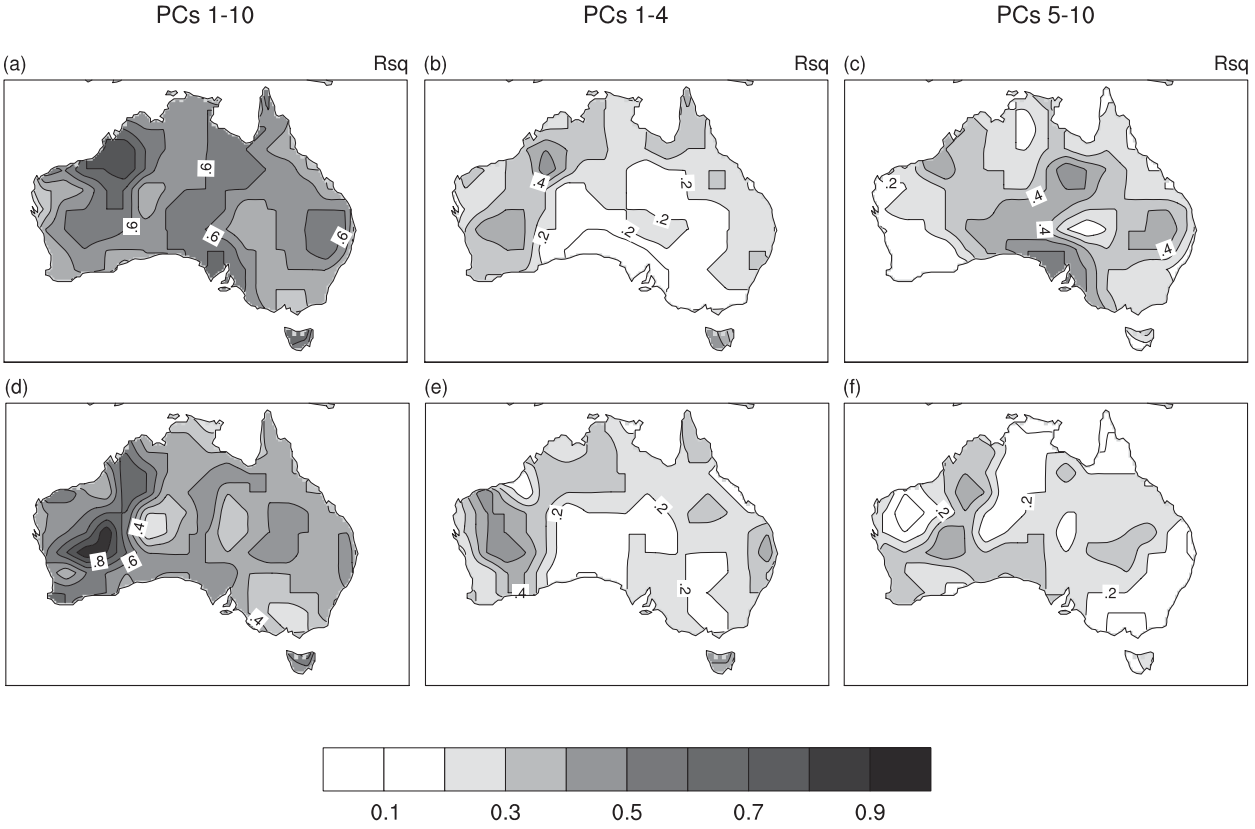

FIG. 4. Explained rainfall variance by (a) the first 10 PCs, (b) the first 4 PCs, and (c) PCs 5–10 of observed MSLP in SON 1981–2006. (d)–(f) As in (a)–(c) but computed in a cross-validated fashion. The contour interval is 0.1. Gray shading starts from 0.2.

TABLE 2. Correlation of the observed and the predicted PCs of MSLP in 1981–2006; r greater than 0.3 is statistically significant at the 5% level (boldfaced) as estimated by a one-sided Student's t test, assuming 26 degrees of freedom.

	r
PC1	**0.69**
PC2	**0.46**
PC3	**0.39**
PC4	0.23
PC5	0.11
PC6	0.22
PC7	0.25
PC8	0.16
PC9	0.06
PC10	0.08

the MSLP variabilities associated with different types of ENSO (PC2 and PC3) are also predictable with some skill ($r > 0.4$) at least at the shortest lead time.

The predicted PCs were used as a predictor field, $\mathbf{X}(k, t)$ where k runs from 1 to 10, in the SVDA algorithm together with observed rainfall [Eqs. (3) and (4)]. Figure 5a shows the leading four SVD modes of predicted MSLP [$\mathbf{G}(k, m)$ where $k = 1$–10, $m = 1$–4] with observed Australian rainfall. Figure 5a indicates that predicted SAM(PC1) and PSA(PC2) have the highest covariance with observed spring rainfall. The four SVD modes explain 97% of the total covariance between observed rainfall and predicted (with prefiltering) MSLP. Figure 5b shows that the temporal variation of the amplitude of the first SVD right vector is significantly correlated with that of the observed rainfall over the north, the east, and the southwest corner where spring rainfall is strongly influenced by traditional ENSO and IOD (cf. Lim et al. 2009, their Figs. 4a and 5a). In comparison, the other three SVD modes show only moderate relationships with the observed rainfall (Figs. 5c–e). The temporal coefficients of the four SVD right vectors account for 20%–50% of observed rainfall variance in most of northern and eastern Australia in spring (Fig. 5f).

A 10-member ensemble hindcast set was produced from each of these models by projecting 10 members of POAMA rainfall and MSLP predictions to the calibration and bridging schemes, respectively. The two statistical–dynamical models were developed and verified with complete cross validation (including the development of the EOFs and SVDA) by leaving out the verification year while the models were developed with the data in the remaining years. This process was repeated throughout the entire hindcast period.

d. Skill analysis

Figure 6 displays the BSS of statistically calibrated and bridged forecasts in comparison to the direct rainfall forecasts from POAMA. Some local improvements are made by these statistical–dynamical approaches in different parts of Australia where significant relationships exist between the observed rainfall and the predictors (Figs. 2f and 5f). However, forecast skill of both statistical–dynamical models is overall not as good as that of the direct rainfall prediction from POAMA over most of Australia. The reason for the overall poorer performance of the statistical–dynamical models is likely to be related to the strict cross-validation process: we are attempting to correct a relatively modest signal using a short record of hindcasts. Also, in the case of constructing a statistical bridging model, it is difficult to find a dominant predictor variable to explain the variability of Australian rainfall over the entire country and in all seasons (e.g., Murphy and Timbal 2008; Risbey et al. 2009). Tropical SST might be a better candidate as a predictor field because there is a reasonably strong empirical relationship between Australian rainfall and ENSO (e.g., Lim et al. 2009), but POAMA ensemble forecasts for tropical SST tend to have narrow spread (not shown), which increases the frequency of very emphatic forecasts and therefore results in an increase in the magnitude of error.

5. Homogeneous multimodel ensemble prediction

Although the statistical–dynamical models by themselves do not outperform the direct predictions from POAMA, they could contribute to skill improvement in the context of a multimodel ensemble, provided that there is some independent information in each component model (Doblas-Reyes et al. 2000; Zebiak 2003, Coelho et al. 2004; Hagedorn et al. 2005). Some indication of the independence of the three sets of forecasts is shown in Fig. 7, which displays the correlation of the ensemble mean rainfall anomaly from the two statistical–dynamical models with the direct prediction from POAMA. While the bridging forecasts are clearly largely independent (low correlation) from the direct predictions, the calibrated forecasts are strongly related with the direct predictions in the southeast where POAMA demonstrates its best skill in predicting spring rainfall. Nonetheless, about 20% of the variance of the calibrated forecasts is unexplained by the direct predictions over the southeast and the unexplained variance increases away from the southeast. We therefore construct a 30-member homogeneous multimodel ensemble by simply pooling together the rainfall predictions of these two statistical–dynamical models and POAMA (10 members from each single model).

The attributes diagram of the HMME predictions is shown in Fig. 8a. Forecast reliability from the HMME is much improved compared to the reliability of the direct

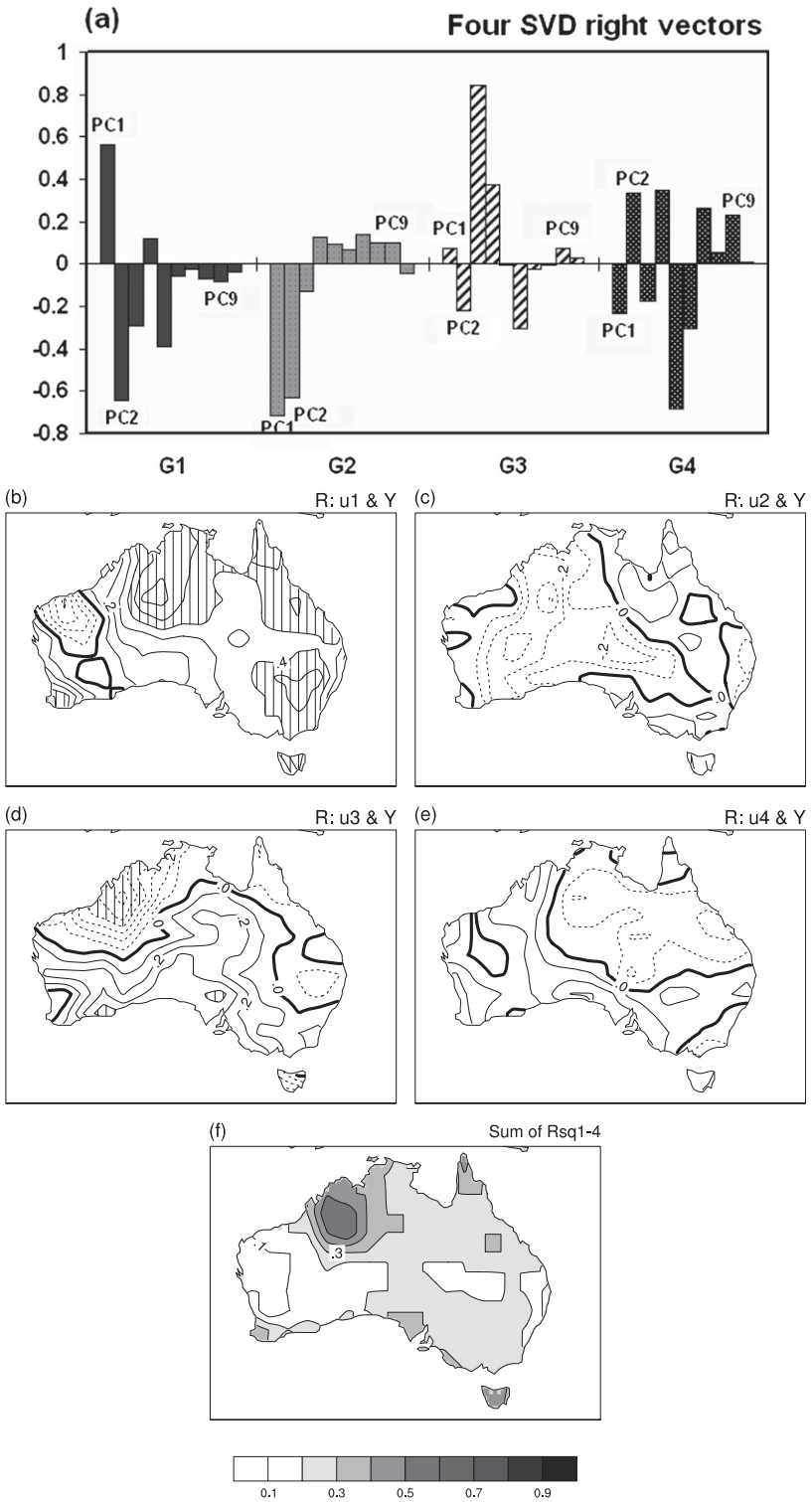

FIG. 5. (a) The first four SVD right vectors of the bridging scheme [POAMA-predicted MSLP (prefiltered) components that most covaries with observed rainfall]. (b)–(e) Correlations of observed rainfall with the expansion coefficients of the four SVD right vectors shown in (a). The contour interval is 0.1, and the areas of the statistically significant correlation at the 5% level (two-tailed Student's *t* test with 26 samples) are hatched. (f) Explained variance of the observed rainfall by the four SVD modes in the period of 1981–2006. The contour interval is 0.1. Gray shading starts from 0.2.

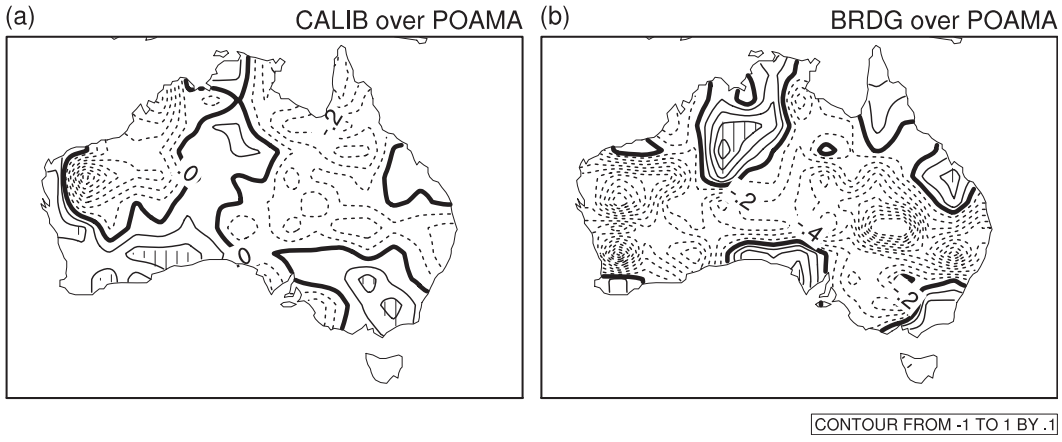

FIG. 6. BSS of the statistical (a) calibration and (b) bridging schemes of predicting above median rainfall at LT 0, taking direct POAMA rainfall forecasts as reference forecasts. The contour interval is 0.1. The solid (dashed) contours indicate positive (negative) changes made by the statistical–dynamical models compared with the direct rainfall predictions from POAMA. Hatching indicates the areas of statistically significant improvements by the calibration and the bridging schemes in the BSS, tested by bootstrapping 2000 samples and rejecting a null hypothesis (H_0: BSS $\leq$ 0) at the 10% level.

prediction of rainfall from POAMA (Fig. 1a) as the forecasts in each probability bin are more closely aligned with the perfect reliability line. Also, the HMME forecasts have smaller magnitudes of error than a climatological forecast as indicated by forecasts found in the gray areas. This improvement of the HMME forecasts over a climatological forecast is also confirmed by the BSS in Fig. 8b—the HMME provides overall improved forecasts compared to the climatological forecast over most of the country except for central Western Australia.

In comparing the HMME forecasts to direct rainfall forecasts from POAMA, we first computed mean reliability error (REL) and resolution (RES) as follows [Wilks 2006, his Eq. (7.40)]:

$$\mathrm{REL} = \frac{1}{n}\sum_{i=1}^{I} N_i(F_i - \overline{O}_i)^2 \tag{6}$$

$$\mathrm{RES} = \frac{1}{n}\sum_{i=1}^{I} N_i(\overline{O}_i - \overline{O})^2, \tag{7}$$

where n is the total number of the forecasts, i indicates forecast categories and runs from 1 to I (I = 10 with 10%

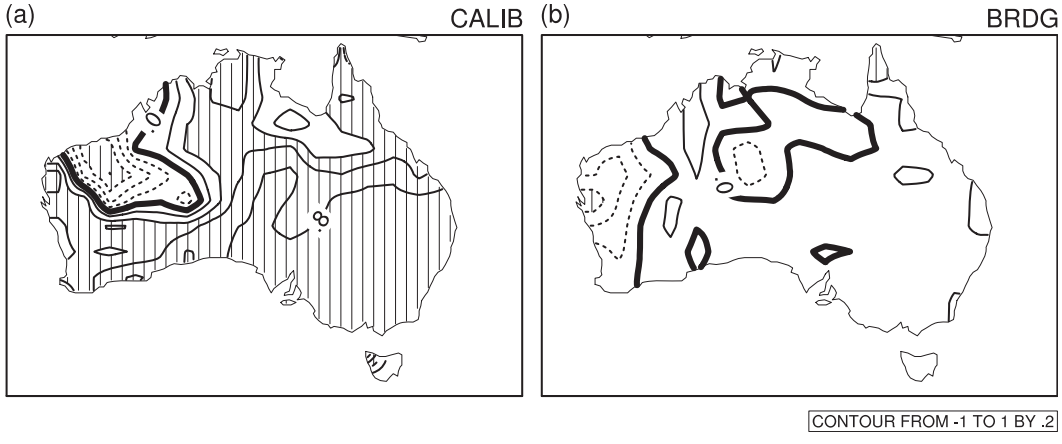

FIG. 7. Correlation of the ensemble mean rainfall forecasts from (a) the statistical calibration scheme and (b) the statistical bridging scheme with the ensemble mean rainfall forecasts from POAMA. The contour interval is 0.2. The solid (dashed) contours indicate positive (negative) correlation. The areas of the statistically significant correlation at the 5% level (two-tailed Student's t test with 26 samples) are hatched.

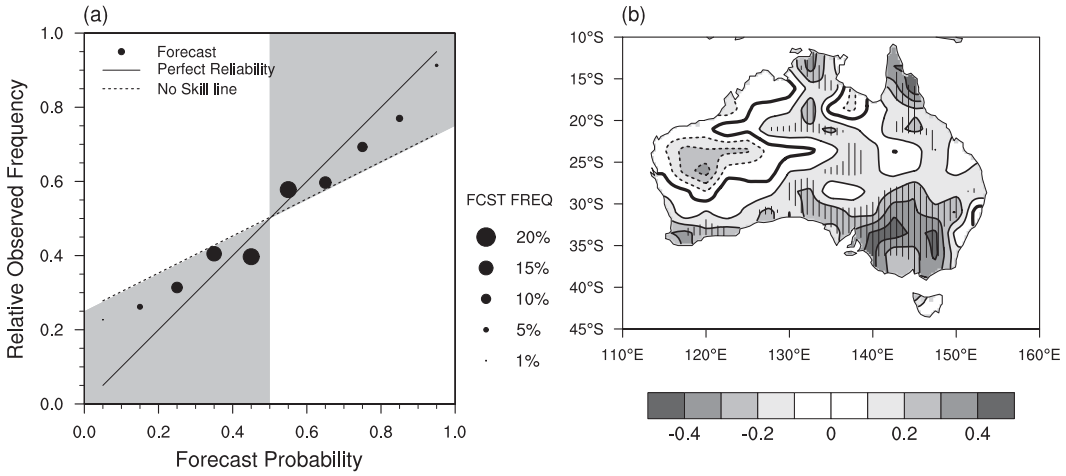

FIG. 8. (a) Attributes diagram of HMME prediction of above median rainfall in SON at LT 0, using 26 yr of hindcasts (1981–2006) over all grid points over Australia. The forecasts in the gray areas are considered to be reliable forecasts. (b) BSS of HMME forecasts of exceeding median rainfall in the hindcast period, taking a climatological forecast as a reference forecast system. The contour interval is 0.1. The solid (dashed) contours indicate positive (negative) changes made by HMME compared with the climatological forecast. Gray shadings denote magnitudes of change. Hatching indicates the areas of statistically significant improvements by HMME in the BSS, tested by bootstrapping 2000 samples and rejecting a null hypothesis (H_0: BSS $\leq$ 0) at the 10% level.

interval in this study), F_i is the forecast probability of rainfall being above the median that falls into the ith bin, $\overline{O}_i$ is the relative frequency of the observed rainfall being above the median given each forecast F_i, and $\overline{O}$ is the climatology of above median rainfall in the observation. The smaller REL and the larger RES is, the more skillful forecasts are. Figures 9a,b display the mean reliability error and resolution of HMME and direct POAMA forecasts at every grid point over Australia. The scatterplots demonstrate that the HMME produces better forecasts with reduced reliability error and increased resolution than POAMA—62% of the grid points demonstrate smaller reliability error, and 49% of the grid points show increased resolution with the HMME forecasts. Although the improvement in resolution is moderate (partly because the resolution of POAMA forecasts is already good), it is a meaningful achievement by the HMME as it is known to be difficult to increase forecast resolution without adding independent information to the prediction system (Stephenson et al. 2005; Doblas-Reyes et al. 2006).

According to Fig. 9c, there is up to 20%–30% improvement of the BSS for the HMME compared to that of the direct predictions from POAMA over the northern and southern parts of the country in SON at LT 0. Such improvements are consistent with the areas where reliability and resolution are both improved, or one improves while the other stays the same (Fig. 9d). In contrast, near the border of New South Wales and Queensland, where direct prediction from POAMA exhibits its best performance (Fig. 1b), skill is slightly reduced by the HMME.

In addition, Doblas-Reyes et al. (2000) pointed out that it is necessary to assess the deterministic skill of a multimodel ensemble system because uncertainty in the probabilistic forecasts can be just an expression of lack of predictable signal if a system produces only noise. Therefore, we formed a grand ensemble mean of the HMME consisting of 30 ensemble members, and evaluated the skill of predicting rainfall anomalies with correlation coefficient and normalized root-mean-squared-error [RMSE normalized by the standard deviation of the observed rainfall (NRMSE)] averaged over all grid points over Australia (Table 3). As indicated in Fig. 1b and Fig. 9c, prediction skill varies considerably region to region, so the average of the correlation coefficients over all grid points is small in Table 3. Nevertheless, the HMME demonstrates an improvement over the single models in terms of correlation, and its NRMSE is not worse than those of the single models, which confirms that the HMME improves skill in predicting Australian spring rainfall.

The improvement by the HMME over the direct output from POAMA (which is overall the best single model in this study) stems from cancellation of forecast errors and broadening of the spread of the ensemble, which is the basis for conventional multimodel ensemble approach with dynamical models (Doblas-Reyes et al. 2000; Hagedorn et al. 2005; Weigel et al. 2008). Hagedorn et al. (2005) argued that even if less skillful component models

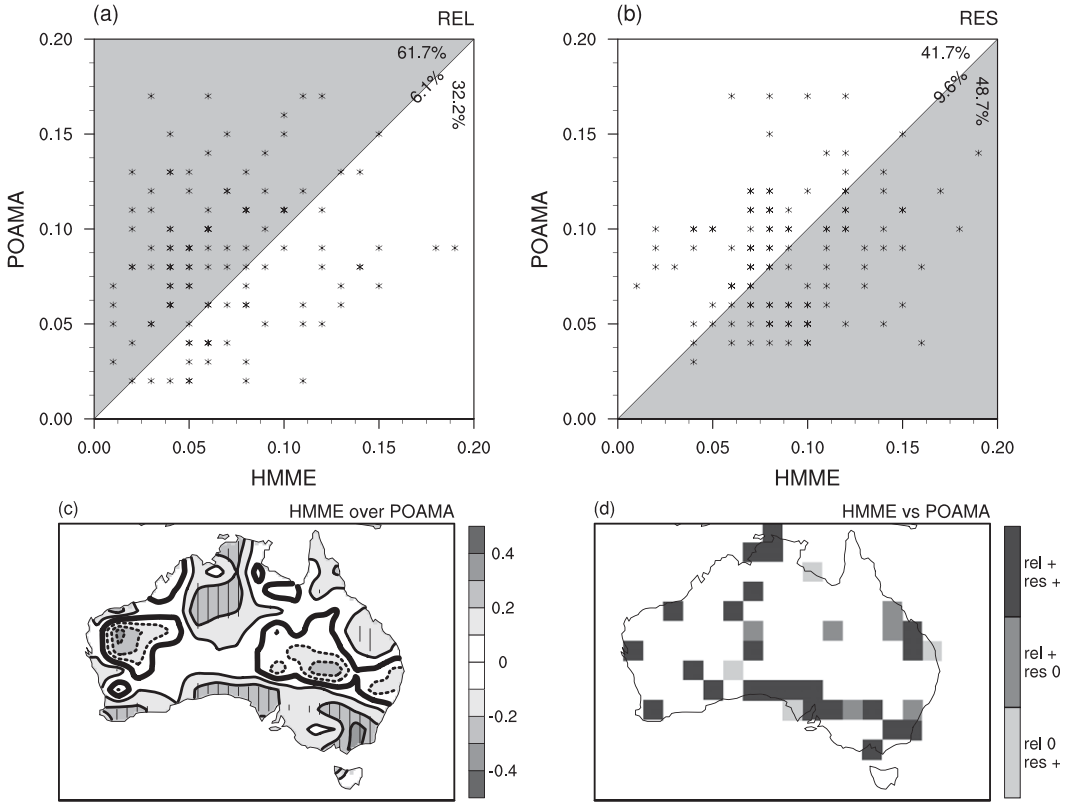

FIG. 9. (a),(b) REL and RES of POAMA and HMME forecasts for above median rainfall at each grid point over Australia. Asterisks in gray areas in (a) and (b) indicate the grid points where HMME has reduced reliability error and increased resolution compared to POAMA, respectively. (c) BSS of HMME forecasts for above median rainfall, taking POAMA forecasts as reference forecasts, in the hindcast period. The contour interval is 0.1. The solid (dashed) contours indicate positive (negative) changes made by HMME compared to POAMA forecast. Gray shadings denote magnitudes of change. Hatching indicates the areas of statistically significant improvements by HMME in the BSS, tested by bootstrapping 2000 samples and rejecting a null hypothesis (H_0: BSS ≤ 0) at the 10% level. (d) Locations where reliability (denoted as rel) and resolution (denoted as res) are both improved by HMME or one is improved by HMME while the other is unchanged, compared to POAMA.

are included in a multimodel system, the skill of the multimodel ensemble system can outperform the ''best'' single model as far as the component models have their own strengths and weaknesses. We demonstrate this with two examples. Figure 10 shows observed rainfall anomalies in 1987 and 1999, which were an El Niño and a La Niña year, respectively. Spring 1987 was anomalously dry whereas spring 1999 was anomalously wet over most of the country in the observations. Both rainfall anomalies are poorly predicted by POAMA over broad areas of the country (Figs. 11a,e). But in 1987 the bridging scheme predicts the dry condition (Fig. 11c) and in 1999 the calibration scheme predicts the wet condition (Fig. 11f) correctly. By combining these predictions, the dry condition in 1987 and wet condition in 1999 both are better predicted than using only direct predictions from POAMA (Figs. 11d,h).

Finally, we have assessed the performance of POAMA, the statistical calibration and bridging schemes and the HMME scheme for the available real-time forecasts. The real-time POAMA v1.5b system, which has been operational since 2007, runs once per day, making a nine-month forecast. For this preliminary test, a 30-member lagged ensemble of the real-time forecasts for rainfall and MSLP are available for the period September 2007 to May 2009. In predicting above median rainfall, the HMME results in higher skill than any of the three

TABLE 3. Average skill of POAMA, the statistical–dynamical schemes, and the HMME scheme in predicting SON rainfall anomaly over Australia for the period 1981–2006. Skill is measured with correlation and NRMSE.

	Correlation	NRMSE
POAMA	0.27	1.39
Calibration	0.10	0.97
Bridging	0.20	0.96
HMME	0.29	0.96

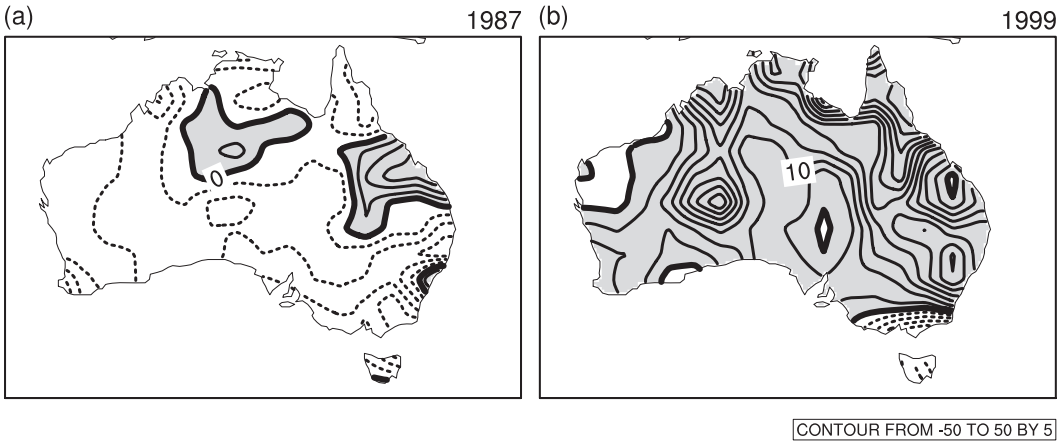

FIG. 10. Observed rainfall anomaly in the spring of (a) 1987 and (b) 1999. The contour interval is 5 mm month^{-1}. The gray shading with solid contours (no shading with dashed contours) indicates positive (negative) rainfall anomaly.

individual models over most of the country during the 21-month period as measured by the BSS. Figure 12 shows the BSS of the HMME, taking POAMA as a reference forecast system. With the HMME, broad areas of Australia except for the southeast demonstrate positive BSS compared to the direct rainfall predictions from POAMA, and 60% of those grid points with positive BSS have greater than 10% improvement.

6. Concluding remarks

We have analyzed the characteristics and skill of probabilistic forecasts of Australian spring rainfall from the POAMA dynamical coupled model forecasting system and have examined the feasibility of forecast skill improvement through statistical–dynamical prediction methods. A 10-member ensemble of hindcasts for 1981–2006 generated from the current operational version of POAMA was employed for this study.

a. Three individual models

The direct probabilistic forecasts from POAMA for rainfall exceeding its median exhibit poor to moderate reliability at LT 0 although the direct forecasts from POAMA outperform a climatological forecast over the eastern part of the country. To examine if additional skill improvement is obtainable through statistical postprocessing, we have developed a statistical calibration scheme and a statistical bridging scheme that use POAMA's predictions of Australian rainfall and extratropical MSLP in the SH as predictors of Australian rainfall, respectively. Forecast skill, as measured by the BSS, from these statistical–dynamical prediction schemes is lower than that from the direct rainfall prediction from POAMA in most of Australia: the poorer performance of the statistical bridging scheme is due to the modest strength of the relationship between Australian spring rainfall and the dominant modes of large-scale climate components (e.g., SAM, ENSO) together with POAMA's moderate skill in predicting MSLP. The poorer performance of the statistical calibration model suggests that major source of errors in POAMA rainfall forecasts is unlikely to be the model's systematic linear bias, which should in principle be corrected statistically. These results suggest then that our hindcast record is probably too short in order to develop stable statistical corrections. For instance, during the period 1981–2006 we sample only a handful of El Niño and La Niña events, which are the primary drivers of climate variability in Australia. Extending the hindcast set back to the beginning of the atmospheric reanalyses (~1958) would seem to be warranted and is being considered with POAMA although ocean initial conditions will be degraded due to a lack of in situ and satellite observations.

b. Homogeneous multimodel ensemble prediction

As an experiment, we combined the direct rainfall predictions from POAMA together with both the statistically calibrated and bridged predictions to make a multimodel ensemble prediction. In SON at LT 0 our homogeneous multimodel ensemble scheme reduces probabilistic forecast error represented by the improvement in the BSS. This skill improvement is gained by large improvement in reliability and slight improvement in resolution. However, the improvement over broad areas of Australia is made at the cost of skill across central western and central eastern Australia.

The POAMA seasonal forecast system is continually evolving and improving – subsequent versions of POAMA will address issues such as correcting model bias and drift,

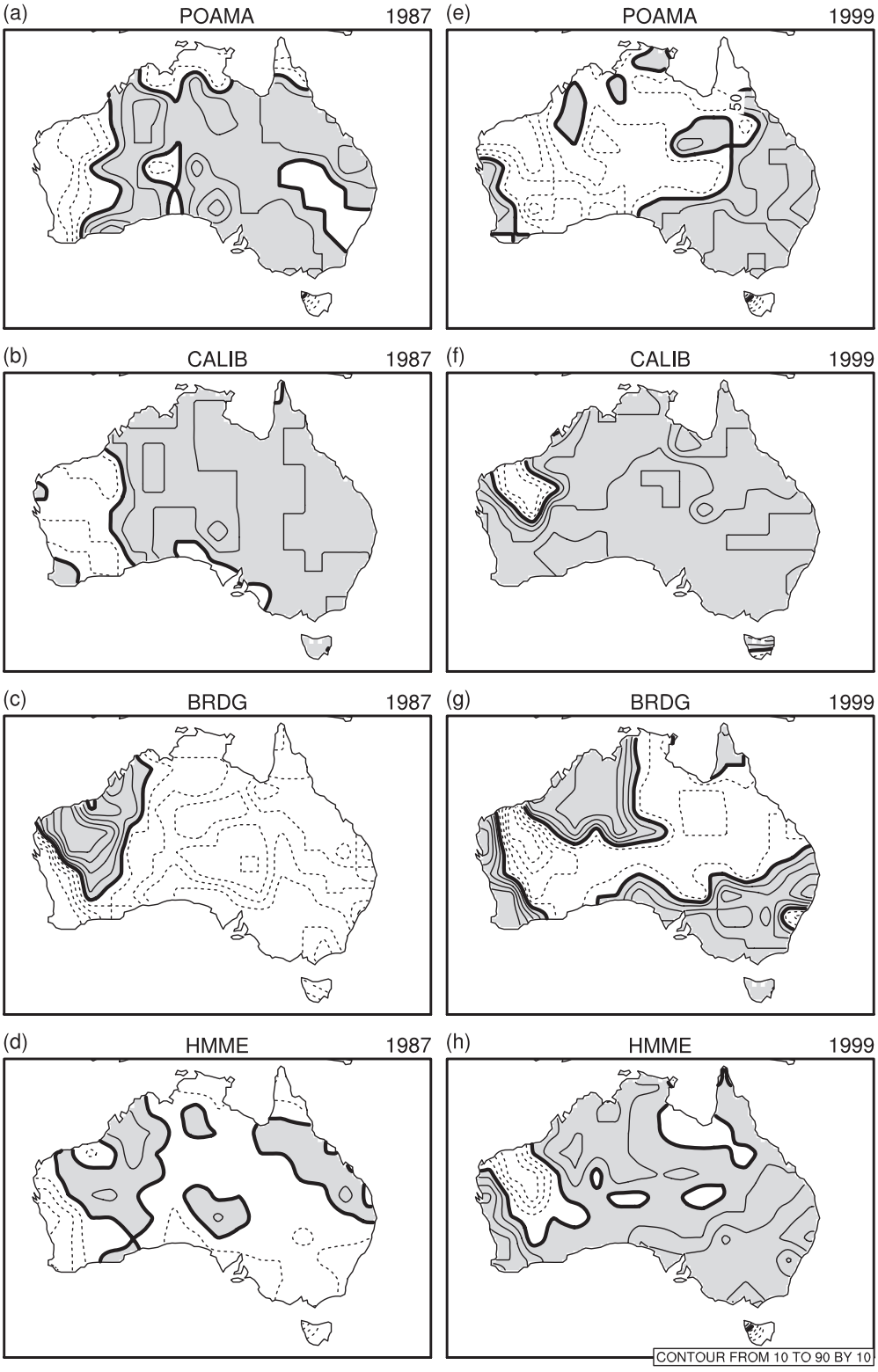

FIG. 11. Probabilistic forecasts for the chance of rainfall being above the median in 1987 and 1999 spring seasons from (a),(e) POAMA; (b),(f) statistical calibration; (c),(g) statistical bridging; and (d),(h) HMME. The contour interval is 10%. The solid (dashed) contours indicate higher (lower) than 50% chance for rainfall being above the median. Areas with more than 50% chance for rainfall being above the median are shaded with gray.

FIG. 12. BSS of HMME forecasts for above median rainfall, taking POAMA real-time forecasts as reference forecasts, in the 21-month period of September 2007–May 2009. The contour interval is 0.1. The solid (dashed) contours indicate positive (negative) changes made by HMME forecasts compared with POAMA forecasts. Gray shadings denote magnitudes of change. Hatching indicates the areas of statistically significant improvements by HMME in the BSS, tested by bootstrapping 2000 samples and rejecting a null hypothesis (H_0: BSS $\leq$ 0) at the 10% level.

increasing model horizontal and vertical resolution, improving model physics, and improved initialization so as to provide skillful prediction of regional climate variability. The potential for further skill improvement using statistical postprocessing seems to be somewhat limited because of a short length of hindcasts based on high-quality observed data (~30 yr) and the large level of climate variability and climate change in Australia (Timbal et al. 2007). However, our study suggests that even if statistical or statistical–dynamical forecasts are not very skillful, they can still contribute to improving seasonal forecast skill as components in a multimodel ensemble system. More sophisticated methods of combining single models (e.g., Bayesian approach) might be able to stretch the skill of the HMME further, which is worth exploring in the future.

Acknowledgments. This research was supported in part by the South Eastern Australian Climate Initiative (SEACI; http://www.mdbc.gov.au/subs/seaci/). The authors are grateful to Drs Andrew Marshall and Karl Braganza and two reviewers for their helpful comments on this manuscript.

REFERENCES

Ashok, K., S. K. Behera, S. A. Rao, and H. Weng, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.,* **112,** C1107, doi:10.1029/2006JC003798.

——, C.-Y. Tam, and W.-J. Lee, 2009: ENSO Modoki impact on the Southern Hemisphere storm track activity during extended austral winter. *Geophys. Res. Lett.,* **36,** L12705, doi:10.1029/2009GL038847.

Barnett, T. P., and R. W. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.,* **115,** 1825–1850.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate,* **5,** 541–560.

Coelho, C. A. S., S. Pezzulli, M. Balmaseda, F. J. Doblas-Reyes, and D. B. Stephenson, 2004: Forecast calibration and combination: A simple Bayesian approach for ENSO. *J. Climate,* **17,** 1504–1516.

Colman, R., and Coauthors, 2005: BMRC Atmospheric Model (BAM) version 3.0: Comparison with mean climatology. BMRC Res. Rep. 108, Bureau of Meteorology Research Centre, 32 pp.

Doblas-Reyes, F. J., M. Deque, and J.-P. Piedelievre, 2000: Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Quart. J. Roy. Meteor. Soc.,* **126,** 2069–2088.

——, R. Hagedorn, T. N. Palmer, and J.-J. Morcrette, 2006: Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. *Geophys. Res. Lett.,* **33,** L07708, doi:10.1029/2005GL025061.

Drosdowsky, W., and L. Chambers, 2001: Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *J. Climate,* **14,** 1677–1687.

Feddersen, H., A. Navarra, and M. N. Ward, 1999: Reduction of model systematic error by statistical correction for dynamical seasonal predictions. *J. Climate,* **12,** 1974–1989.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting–I. Basic concept. *Tellus,* **57A,** 219–233.

Hendon, H. H., D. W. J. Thompson, and M. C. Wheeler, 2007: Australian rainfall and surface temperature variations associated with the Southern Hemisphere annular mode. *J. Climate,* **20,** 2452–2467.

——, E. Lim, G. Wang, O. Alves, and D. Hudson, 2009: Prospects for predicting two flavors of El Nino. *Geophys. Res. Lett.,* **36,** L19713, doi:10.1029/2009GL040100.

Hudson, D., O. Alves, H. H. Hendon, and G. Wang, 2010: The impact of atmospheric initialisation on seasonal prediction of tropical Pacific SST. *Climate Dyn.,* doi:10.1007/s00382-010-0763-9, in press.

Jones, D. A., and G. Weymouth, 1997: An Australian monthly rainfall dataset. Tech. Rep. 70, Bureau of Meteorology, 19 pp.

Kanamitsu, M., R. E. Kistler, R. W. Reynolds, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP/DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1641–1643.

Kang, H., K.-H. An, C.-K. Park, A. L. S. Solis, and K. Stitthichivapak, 2007: Multimodel output statistical downscaling prediction of precipitation in the Philippines and Thailand. *Geophys. Res. Lett.,* **34,** L15710, doi:10.1029/2007GL030730.

Kang, I.-S., J.-Y. Lee, and C.-K. Park, 2004: Potential predictability of summer mean precipitation in a dynamical seasonal prediction system with systematic error correction. *J. Climate,* **17,** 834–844.

Karoly, D., 1989: Southern Hemisphere circulation features associated with El Niño–Southern Oscillation events. *J. Climate,* **2,** 1239–1251.

Kiladis, G. N., and K. C. Mo, 1998: Interannual and intraseasonal variability in the Southern Hemisphere. *Meteorology of the Southern Hemisphere, Meteor. Monogr.,* No. 49, Amer. Meteor. Soc., 307–336.

Lim, E.-P., H. H. Hendon, D. Hudson, G. Wang, and O. Alves, 2009: Dynamical forecast of inter–El Niño variations of tropical

SST and Australian spring rainfall. *Mon. Wea. Rev.,* **137,** 3796–3810.

Lin, H., and J. Derome, 2005: Correction of atmospheric dynamical seasonal forecasts using the leading ocean-forced spatial patterns. *Geophys. Res. Lett.,* **32,** L14804, doi:10.1029/2005GL023060.

Marshall, G. J., 2003: Trends in the Southern Annular Mode from observations and reanalyses. *J. Climate,* **16,** 4134–4143.

Mason, S. J., and D. B. Stephenson, 2008: How do we know whether seasonal climate forecasts are any good? *Seasonal Climate: Forecasting and Managing Risk,* A. Troccoli et al., Eds., Springer, 259–289.

McBride, J. L., and N. Nicholls, 1983: Seasonal relationships between Australian rainfall and the Southern Oscillation. *Mon. Wea. Rev.,* **111,** 1998–2004.

Meyers, G., P. MacIntosh, 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.

Mo, K. C., 2000: Relationships between low-frequency variability in the Southern Hemisphere and sea surface temperature anomalies. *J. Climate,* **13,** 3599–3610.

Mo, R., and D. M. Straus, 2002: Statistical–dynamical seasonal prediction based on principal component regression of GCM ensemble integrations. *Mon. Wea. Rev.,* **130,** 2167–2187.

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.

Nicholls, N., 1989: Sea surface temperatures and Australian winter rainfall. *J. Climate,* **2,** 965–973.

Oke, P. R., A. Schiller, D. A. Griffin, and G. B. Brassington, 2005: Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Quart. J. Roy. Meteor. Soc.,* **131,** 3301–3311.

Palmer, T. N., and Coauthors, 2004: Development of a European multimodel ensemble system for seasonal to interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872.

——, F. J. Doblas-Reyes, A. Weisheimer, and M. J. Rodwell, 2008: Toward seamless prediction: Calibration of climate change projections using seasonal forecasts. *Bull. Amer. Meteor. Soc.,* **89,** 459–470.

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.

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.

——, T. Ambrizzi, and S. E. T. Ferraz, 2005: Indian Ocean dipole mode events and austral surface air temperature anomalies. *Dyn. Atmos. Oceans,* **39,** 87–100.

Schiller, A., J. S. Godfrey, P. C. McIntosh, G. Meyers, N. R. Smith, O. Alves, G. Wang, and R. Fiedler, 2002: A new version of the Australian Community Ocean Model for Seasonal Climate Prediction. CSIRO Marine Res. Rep. 240, 82 pp.

Silvestri, G. E., and C. S. Vera, 2003: Antarctic Oscillation signal on precipitation anomalies over southeastern South America. *Geophys. Res. Lett.,* **30,** 2115, doi:10.1029/2003GL018277.

Smith, N. R., J. E. Blomley, and G. Meyers, 1991: A univariate statistical interpolation scheme for subsurface thermal analyses in the tropical oceans. *Prog. Oceanogr.,* **28,** 219–256.

Stephenson, D. B., C. A. D. S. Coelho, F. J. Doblas-Reyes, and M. Balmaseda, 2005: Forecast assimilation: A unified framework for the combination of multimodel weather and climate predictions. *Tellus,* **57A,** 253–264.

Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves, and M. A. Balmaseda, 1998: Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature,* **392,** 370–373.

——, D. Anderson, M. Balmaseda, F. Doblas-Reyes, L. Ferranti, K. Mogensen, F. Molteni, and F. Vitart, 2010: ECMWF Seasonal Forecast System 3 and its prediction of sea surface temperature. *Climate Dyn.,* in press.

Thompson, D. W. J., and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate,* **13,** 1000–1016.

Timbal, B., N. Nicholls, W. Cai, K. Hennessy, and P. Hope, 2007: Causes of past climate change, climate change in Australia. Tech. Rep. 2007, 29–35.

Tippett, M. K., L. Goddard, and A. G. Barnston, 2005: Statistical–dynamical seasonal forecasts of central-southwest Asian winter precipitation. *J. Climate,* **18,** 1831–1843.

Troccoli, A., M. Harrison, D. L. T. Anderson, and S. J. Mason, 2008: *Seasonal Climate: Forecasting and Managing Risk*. NATO Science Series, Vol. 82, Springer Academic Publishers, 499 pp.

Ummenhofer, C. C., A. Sen Gupta, A. S. Taschetto, and M. H. England, 2009: Modulation of Australian precipitation by meridional gradients in East Indian Ocean sea surface temperature. *J. Climate,* **22,** 5597–5610.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

Valke, S., L. Terray, and A. Piacentini, 2000: The OASIS coupled user guide version 2.4. Tech. Rep. TR/CMGC/00-10, CERFACS, 85 pp.

Voldoire, A., B. Timbal, and S. Power, 2002: Statistical–dynamical seasonal forecasting. BMRC Res. Rep. 87, Bureau of Meteorology Research Centre, 62 pp.

Wang, G., and H. H. Hendon, 2007: Sensitivity of Australian rainfall to inter–El Niño variations. *J. Climate,* **20,** 4211–4226.

——, O. Alves, D. Hudson, H. Hendon, G. Liu, and F. Tseitkin, 2008: SST skill assessment from the new POAMA-1.5 system. *BMRC Res. Lett.,* **8,** 1–6.

Ward, M. N., and A. Navarra, 1997: Pattern analysis of SST-forced variability in ensemble GCM simulations: Examples over Europe and the tropical Pacific. *J. Climate,* **10,** 2210–2220.

Weigel, A. P., M. A. Liniger, and C. Appenzeller, 2007: The discrete Brier and ranked probability skill scores. *Mon. Wea. Rev.,* **135,** 118–124.

——, ——, and ——, 2008: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quart. J. Roy. Meteor. Soc.,* **134,** 241–260.

Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 592 pp.

Zebiak, S. E., 2003: Research potential for improvements in climate prediction. *Bull. Amer. Meteor. Soc.,* **84,** 1692–1696.

Zhao, M., and H. H. Hendon, 2009: Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. *Quart. J. Roy. Meteor. Soc.,* **135,** 337–352.