Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

CSIRO

Dynamical seasonal prediction of tropical Indo-Pacific SST and Australian rainfall with improved ocean initial conditions

Eun-Pa Lim, Harry H. Hendon, Oscar Alves, Yonghong Yin, Guomin Wang, Debra Hudson, Mei Zhao and Li Shi

CAWCR Technical Report No. 032

October 2010

www.cawcr.gov.au

Dynamical seasonal prediction of tropical Indo-Pacific SST and Australian rainfall with improved ocean initial conditions

Eun-Pa Lim, Harry H. Hendon, Oscar Alves, Yonghong Yin,
Guomin Wang, Debra Hudson, Mei Zhao and Li Shi

The Centre for Australian Weather and Climate Research
- a partnership between CSIRO and the Bureau of Meteorology

CAWCR Technical Report No. 032

October 2010

ISSN: 1836-019X

National Library of Australia Cataloguing-in-Publication entry

Author: Eun-Pa Lim et al.

Title: Dynamical seasonal prediction of tropical Indo-Pacific SST and Australian rainfall with improved ocean initial conditions.

ISBN: 978-1-921826-09-2 (Electronic resource)

Series: CAWCR technical report; 32.

Notes: Included bibliography references and index.

Subjects: Ocean temperature.
Rain and rainfall--Australia.
Ocean-atmosphere interaction.

Other Authors / Contributors; Harry H. Hendon, Oscar Alves, Yonghong Yin, Guomin Wang, Debra Hudson, Mei Zhao, Li Shi and Keith Day (Ed)

Dewey Number: 551.5

Enquiries should be addressed to:
Eun-Pa Lim
Centre for Australian Weather & Climate Research
GPO Box 1289 Melbourne
Victoria 3001, Australia

E.Lim@bom.gov.au

Copyright and Disclaimer

All images reproduced in grayscale. A colour version of CAWCR Technical Report No.032
is available online: http://www.cawcr.gov.au

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The Australian Bureau of Meteorology (BoM) jointly with the Commonwealth Scientific and Industrial Research Organisation (CSIRO) has developed a dynamical intra-seasonal to seasonal climate forecast system, the Predictive Ocean Atmosphere Model for Australia (POAMA; Alves et al. 2003), in order to provide improved seasonal prediction of the Australian climate. The current operational version of the POAMA system (version 1.5b) generates a 9 month forecast each day. These forecasts are primarily used for tropical Pacific sea surface temperature (SST) variations associated with the El Niño-Southern Oscillation (ENSO), which is a main driver of the Australian seasonal climate.

The POAMA system is continuously being developed and improved. Because seasonal climate prediction is an initial value problem (whose solution depends on the conditions specified at the starting time (t_0)), the quality of the forecast depends directly on the quality of the depiction of the initial atmosphere-ocean state at the start of the forecast. A major upgrade for POAMA1.5b was the initialisation of the forecast system with more realistic conditions of the atmosphere and land surface. These improved atmosphere-land surface initial conditions were generated by an Atmosphere and Land Initialisation scheme (ALI; Hudson et al. 2010), which was developed during the South Eastern Australian Climate Initiative phase 1 (SEACI-1). The improved skill for predicting ENSO as a result of the improved atmosphere-land initial conditions provided by ALI is reported in detail in Hudson et al. (2010) and is briefly discussed in the next section.

The next major upgrade for POAMA is the improvement of the ocean initial conditions, which are at the heart of making seasonal climate predictions - it is the upper ocean that provides predictability to ENSO through equatorial wave propagations, and a model's ability to skilfully predict the evolution of the upper ocean significantly relies on the quality of the ocean initial conditions. In POAMA1.5b, the ocean initial conditions were produced from the POAMA Ocean Data Assimilation Scheme (PODAS). PODAS is based on a univariate optimum interpolation (OI) technique of Smith et al. (1991) that assimilates in situ temperature observations in the upper 500 m of the ocean (Wang et al. 2002).

The new state-of-the-art, ensemble-based ocean data assimilation scheme included in the newer version of POAMA (POAMA2) is called the POAMA Ensemble Ocean Data Assimilation System (PEODAS; Yin et al. 2010). PEODAS assimilates not only ocean temperature but also salinity. A unique feature of PEODAS is that for any point in time it produces an ensemble of equivalent ocean states that represent observational uncertainties. The ensemble of ocean states is then directly used as perturbed initial conditions for ensemble forecasts. Yin et al. (2010) show conclusively that the depiction of the upper ocean in PEODAS is significantly more realistic, accurate and dynamically/thermodynamically consistent, compared to that in PODAS.

The aim of this study is, therefore, to assess the impact of these improved ocean initial conditions on forecast skill of tropical Indo-Pacific SSTs and Australian rainfall. We will also explore the extent that multi-model ensembling, based on available versions of POAMA (currently three versions – POAMA1.5b and two versions of POAMA2), can improve regional rainfall forecasts.

We describe POAMA2 model features, the hindcast data sets and verification methods in section 2. In section 3, SST and Australian rainfall forecast skill using POAMA2 is evaluated in comparison to that of POAMA1.5b. Also, performance of a multi-model ensemble system consisting of POAMA1.5b and two different versions of POAMA2 is examined. Lastly, concluding remarks are provided in section 4.

2 POAMA MODEL FEATURES

The POAMA forecast system is based on coupled atmosphere and ocean models. The atmospheric model is the BoM Atmospheric Model version 3 (BAM3.0; Colman et al. 2005) and the ocean model is the Australian Community Ocean Model version 2 (ACOM2; Schiller et al. 2002, Oke et al. 2005). The horizontal structure of BAM3.0 is represented by spherical harmonics with a triangular truncation at wave number 47 (denoted T47, which has approximately 250 km resolution), and the vertical variation is represented by 17 sigma levels. BAM3.0 is used for both POAMA1.5b and POAMA2, but BAM3 in POAMA2 has improved atmospheric model physics associated with shallow convection.

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

2.1 Atmosphere and Land Initialisation

For both POAMA1.5b and POAMA2, atmosphere and land initial conditions are provided by ALI (Hudson et al. 2010). ALI nudges the BAM atmospheric model toward global analyses of zonal and meridional winds, temperature, and humidity. For the hindcast period, ALI nudges toward the ERA-40 reanalyses (Uppala et al. 2005) during 1980-2001 and to the global analyses from the BoM's numerical weather prediction system (GASP) during 2002-2006. For use in real time, ALI nudges to the analyses from GASP that are routinely produced at BoM in real time. Through this nudging toward a reanalysis, ALI produces atmosphere conditions that are very similar to the ERA-40/GASP analyses. However, less initial forecast shock is generated than if the ERA-40/GASP analyses were directly used as initial conditions in BAM3.0. Although the land surface fields are not directly nudged, land surface initial conditions are created indirectly in response to the nudged atmospheric conditions.

Hudson et al. (2010) document the increase in forecast skill resulting from the use of the realistic atmosphere/land initial conditions provided by ALI in POAMA1.5b as compared to POAMA1 (atmospheric initial conditions in POAMA1 contained no real information; they were generated as an atmospheric model response to imposed observed SST). In particular, Hudson et al. (2010) show that the use of ALI in POAMA1.5b results in improved skill for predicting the NINO3 and NINO3.4 SST indices with up to 6 month lead time[1].

[1] 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). For instance, a forecast for February at 1 month lead time is initialised on the 1st of January.

2.2 Ocean Initialisation

As mentioned earlier, in POAMA1.5b ocean initial conditions are provided by PODAS (Wang et al. 2002), which is based on univariate OI technique (Smith et al. 1991). The OI scheme is used to correct the ocean model background temperature field every three days using a 3-day observation window. Only subsurface temperature observations in the top 500 m are assimilated. Salinity is not updated, and ocean current velocity is updated by its geostrophic relation to the temperature increments computed in the assimilation scheme (Burgers et al. 2002). During the assimilation cycle, SST is strongly nudged to the observed SST analysis (Reynolds et al. 2002), which allows the model SST to be close to the observed.

POAMA2 ocean initial conditions are provided by PEODAS, which is a simplified form of an Ensemble Kalman Filter (Yin et al. 2010). In this system, in situ temperature and salinity observations are assimilated to a central run every three days. As it is expected that the important source of background error for the ocean model is observational errors in surface wind stress and surface fluxes on intra-seasonal time, an ensemble of ocean states is generated mainly by perturbing wind stress and surface fluxes. The ensemble of ocean states is then used to compute the background error covariances (i.e. covariance of the discrepancy between the true state and the model's background estimate of the true state) for temperature, salinity and currents. Using these background error covariances, observed data are assimilated to the central run. After each analysis, the ensemble members are updated by nudging toward the analysis from the central run. A resultant ensemble of ocean conditions is used to initialise the ensemble of seasonal forecasts.

According to Yin et al. (2010), PEODAS produces more realistic ocean initial conditions than does PODAS. As an example of an improvement in the ocean initial conditions, Fig. 1 displays correlation of monthly heat content anomaly in the upper 300m ocean (T300) from PEODAS and PODAS with a purely observational analysis, EN3 (based on the ENACT quality-controlled observation database; Ingleby and Huddleston 2007). The correlation of T300 from PEODAS is higher than that from PODAS over the tropical Pacific and western Indian Oceans, suggesting an improved depiction of the upper tropical oceans in PEODAS. Especially, the improvement in the initial conditions of T300 over the tropical Pacific implies the state of ENSO (which is as much a subsurface phenomenon as it is a surface phenomenon) to be better depicted by the POAMA2 system, which is confirmed in the later part of this report.

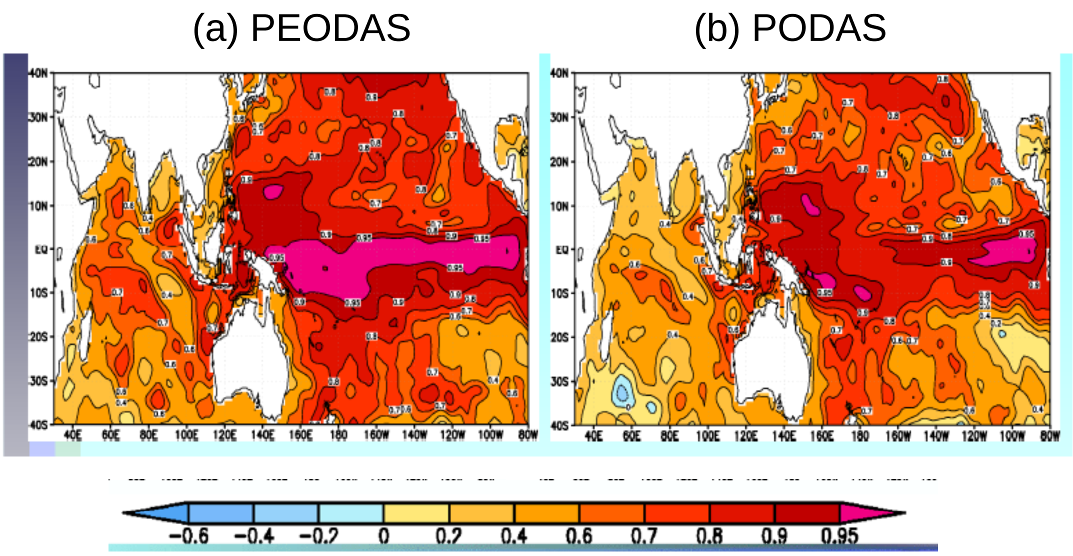


Fig. 1 Correlation of monthly T300 anomaly from (a) PEODAS and (b) PODAS with the observed analysis (EN3) for the period of 1982-2006.

2.3 Model bias correction

A common problem with coupled seasonal forecast models such as POAMA is that the simulated climate drifts as lead time increases. This is demonstrated in Figs 2a and 2b which show the bias of the climatology of SST from the POAMA1.5b and POAMA2 hindcasts (one forecast per month for the period 1980-2006) as a function of forecast lead time. Because the forecasts are initialized from the ocean conditions that are close to the observation, little bias is seen at 0 lead time. However, by three month lead time, a tropical-wide cold bias has developed, together with a warm bias off the coast of South America. By six month lead time the bias is nearly saturated with the bias ranging from -3°C to 6°C. Improved simulation of shallow convection in POAMA2 compared to POAM1.5b reduces the cold bias slightly (by ~ 0.5°C) over the Pacific (Fig. 2b). Nonetheless, the cold bias over most of the oceans and the warm bias over the eastern Pacific basin off South America are still pronounced in POAMA2. A direct result of the cold bias in the equatorial Pacific is that the maximum variability in tropical Pacific SST that is associated with ENSO is shifted westward from the South American coast at increased lead times. Such drift in the SST variability hinders the model's ability to discern differences in SST patterns between different types of ENSO events as lead time increases (Hendon et al. 2009).

In order to alleviate the mean SST bias, a version of POAMA2 was configured with an explicit flux correction scheme. This scheme corrects biases in shortwave radiation, total heat flux and wind stress by adjusting model generated fluxes to be close to the observed counterparts from the Coordinated Ocean Research Experiments (CORE) version 2 data sets (Large and Yeager 2009) for the hindcast period (see further details of the flux correction in Appendix A). As a result of this flux correction, the large cold and warm biases in the tropical Pacific are significantly reduced (Fig. 2c). Hereafter, we will refer to the flux corrected version as POAMA2.4b and the non-flux corrected version as POAMA2.4a.

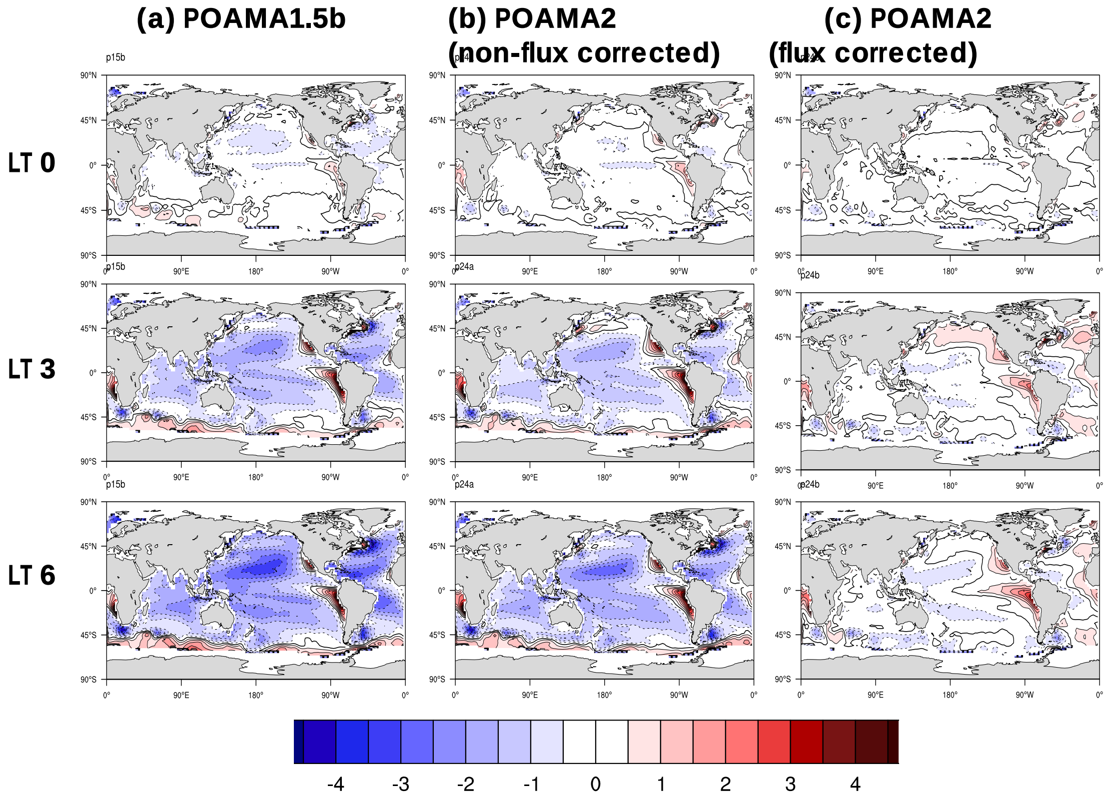


Fig. 2 Difference between the climatologies of predicted SST from (a) POAMA1.5b, (b) POAMA2.4a (non-flux corrected version) and (c) POAMA2.4b (flux corrected version) and observed mean SST (HadISST; Rayner et al. 2003) at 0, 3 and 6 months lead time (LT 0, LT 3 and LT 6, respectively) for the period 1980-2006. Positive (negative) values mean that POAMA predicts higher (lower) SST than observation on average. The contour interval is 0.5 °C.

2.4 Forecast verification

For this study, three different versions of POAMA are assessed – POAMA1.5b (P1.5b; non bias corrected) and two different versions of POAMA2 – non-flux corrected (P2.4a) and flux corrected (P2.4b) (Table 1). From each of these three versions, ten member ensemble retrospective forecasts were generated from the first day of each month for the period 1980-2006. From P1.5b, the ten ensemble members were generated by perturbing the atmospheric initial conditions by using six-hour consecutively earlier atmospheric analyses from the first day of each month. Identical ocean initial conditions from PODAS were used for all 10 members. In comparison, ten member ensemble hindcasts of each of the two versions of POAMA2 were perturbed by ten different sets of ocean initial conditions provided by PEODAS on the first day of each month.

Forecast skill for SST is evaluated using temporal correlation, normalized root-mean-square-error and normalized standard deviation. For rainfall over the Australian region we assess the skill of probabilistic forecasts (based on ten ensemble members) of above median rainfall, using proportion correct (also called hit rate or percent consistent score) and attributes diagrams. SST forecasts are verified against HadISST (Rayner et al. 2003) and Australian rainfall forecasts are verified against the National Climate Centre's gridded monthly analysis (Jones and Weymouth 1997).

Table 1 Overview of model configuration of POAMA1.5b and POAMA2.

	POAMA1.5b	POAMA 2	
Common features	• **Atmospheric model : BAM3.0 (T47 L17)** • **Ocean model : ACOM2 (zonal resolution 2°, meridional resolution 0.5° in the tropics gradually changing to 1.5° near the poles, 25 vertical levels)** • **Atmosphere and land initialised with ALI**		
Ocean Data assimilation System	**PODAS**	**PEODAS**	
Forecast ensemble generation	**Atmosphere perturbed**	**Ocean perturbed**	
Mean state bias correction	**No**	**P2.4a**	**P2.4b**
		No	**Yes**

In this study, we compare predicted anomalies against observed anomalies of SST and rainfall. For the predictions, the monthly climatology was computed as a function of forecast start month and lead time. By forming anomalies relative to the model's climatology, some aspects of the mean model bias are removed (Stockdale et al. 1998, Weigel et al. 2008). The observed anomalies were obtained relative to the observed climatology for the same 1980-2006 period.

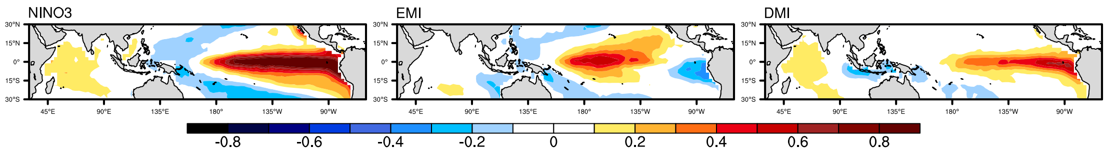
3 RESULTS

3.1 Indo-Pacific SST indices

Predictability of regional seasonal climate results mainly from ENSO via its teleconnections even though the impact of ENSO is different over different regions on the globe. Australia is one of the regions whose climate is highly sensitive to ENSO in its cool seasons (June to November) (McBride and Nicholls 1983).

Furthermore, Australian climate appears to respond differently to El Niño events whose maximum SST warming is found in the eastern Pacific (EP El Niño), to El Niño events whose maximum SST warming is found in the central Pacific (CP El Niño) and to Indian Ocean Dipole (IOD) events that captures the variation of the Indian Ocean east-west dipole mode, whether or not they occur in conjunction with ENSO in the Pacific (Lim et al. 2009, Risbey et al. 2009) (Fig. 3a). For instance, Australian area averaged rainfall is sensitive to CP El Niño, which is captured by El Niño- Modoki index[2] (EMI; Ashok et al. 2007), in autumn and spring, to EP El Niño depicted by NINO3 index[3] in spring, and to IOD depicted by Dipole Mode Index[4] (DMI; Saji et al. 1999) in late autumn to winter (Fig. 3b left panel). The influence of both types of El Niño is somewhat weaker over south eastern Australia but the influence of IOD in spring is stronger over south eastern Australia compared to its impact on all of Australia (Fig. 3b right panel).

(a)



(b)

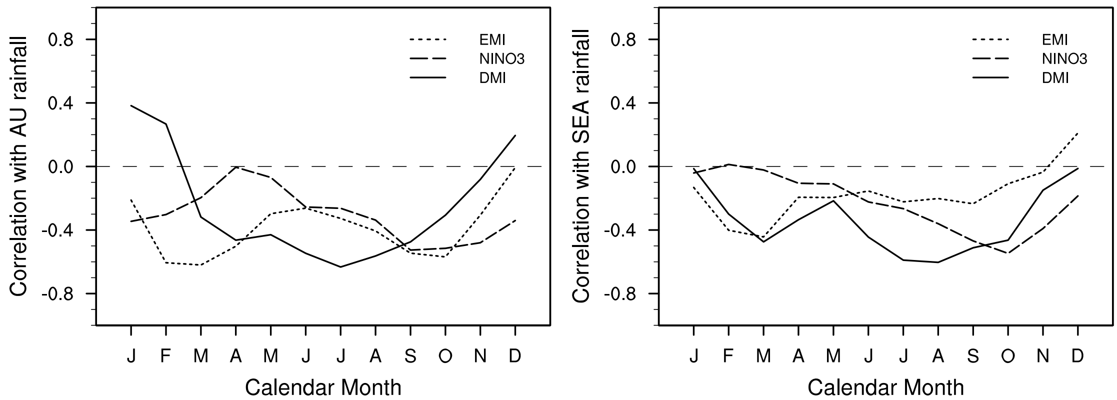


Fig. 3 (a) Regression of monthly SST anomaly on NINO3 index (left), El Niño Modoki Index (EMI; middle), and Indian Ocean Dipole Mode Index (DMI; right), (b) Correlation of seasonally averaged Australian area-averaged rainfall (left) and south eastern Australian area-averaged rainfall with NINO3, EMI and DMI.

[2] EMI= $\overline{SST}_{(165°E\text{-}140°W, 10°S\text{-}10°N)} - 0.5 * \overline{SST}_{(70°W\text{-}110°W, 15°S\text{-}5°N)} - 0.5 * \overline{SST}_{(125°E\text{-}145°E, 10°S\text{-}20°N)}$

[3] NINO3 index = $\overline{SST}_{(90°W\text{-}150°W,\ 5°S\text{-}5°N)}$

[4] DMI= $\overline{SST}_{(50°E\text{-}70°E,\ 10°S\text{-}10°N)} - \overline{SST}_{(90°E\text{-}110°E, 10°S\text{-}0°)}$

Having considered the important role of tropical Indo-Pacific SST in seasonal climate prediction and especially in Australian climate prediction, it is a necessary first step to evaluate a dynamical seasonal forecast system in terms of its ability to predict ENSO. Here, we examine skill of the three different versions of POAMA in predicting NINO3, EMI and DMI, taking persistence[5] as a reference forecast system.

Prediction skill is assessed by correlation, normalized root-mean-square error (NRMSE)[6] and normalized standard deviation (NSTDDEV)[7] from P1.5b, P2.4a and P2.4b (Figs 4-6). Predicted RMSE and STDDEV of an index are normalized by the STDDEV of the corresponding observed index. NRMSE <1 indicates that the forecast error is smaller than a climatological forecast. NSTDDEV shows if amplitude of an index is realistically predicted. NSTDDEV=1 indicates prediction of variability that has observed amplitude. A perfect forecast would have NRMSE= 0 and NSTDDEV=1, a climatological forecast would have NRMSE=1 and NSTDDEV=0, and a completely random forecast but with observed amplitude (e.g. a persistence forecast at long lead time) would have NRMSE=$\sqrt{2}$ and NSTDDEV=1.

In Figs 4 and 5, P2.4a shows improvements in predicting the NINO3 index at lead times longer than two-three months compared to P1.5b based on correlation and NRMSE. The improvement in skill is substantial. For instance, the same skill is achieved with two month longer lead times with P2.4a than with P1.5b. Interestingly, P2.4b shows only minor improvement over P1.5b, suggesting that correction of the mean state in P2.4b had little beneficial impact on the ability to predict EP El Niño (NINO3). P2.4b does show a modest improvement in predicting CP El Niño (EMI) at one-five month lead times compared to P1.5b and P2.4a, which perhaps reflects some beneficial impact of correcting the mean state on predicting this different flavoured El Nino. However, both P2.4a and P2.4b significantly underestimates the amplitudes of NINO3 and EMI as lead time increases (Figs 6a,b), especially compared to P1.5b. That is, the ENSO (both EP and CP ENSO) seems to be unrealistically damped in the POAMA2 model, regardless of whether the mean state bias was corrected. The reason for this is unknown, but the weaker amplitude of predicted ENSO variability in P2.4a and P2.4b means that teleconnections associated with ENSO will be weakened as well, especially at longer lead times. Hence, prediction of regional climate that is largely impacted by ENSO will be hampered at longer lead times.

In regard to predicting the IOD, skill drops rapidly as lead time increases in all three versions of POAMA. And, POAMA2 forecasts (both P2.4a and P2.4b) appear to be less skilful than POAMA1.5b forecasts (Figs 4c and 5c). The amplitude of predicted IOD is slightly more realistic in P2.4b, but all versions of POAMA overestimate the amplitude of the IOD (Fig. 6c). Here, it is worth noting that prediction skill of the IOD is highly sensitive to the observation data which the forecasts are verified against. For instance, Figs 4d, 5d and 6d display the correlation, NRMSE and NSTDDEV of predicted IOD verified against Reynolds SST

[5] Persistence forecast is to forecast the future climate condition to be the same as the present condition.

[6] $NRMSE = \sqrt{\frac{\sum_{i=1}^{I}(F_i - O_i)^2}{I}} \Big/ \sigma_O$, where F_i, O_i are the predicted and observed indices in the i-th year, respectively, I is the total number of years, and σ_O is the standard deviation of the observed index.

[7] $NSTDDEV = \sigma_F \big/ \sigma_O$, where σ_F is the standard deviation of the predicted index.

(Reynolds et al. 2002), respectively. Verifying against Reynolds SST generally results in higher skill and more realistic amplitude than when compared to HadISST. Interestingly, the correlation skill of persistence forecast of IOD is also very different between Reynolds SST and HadISST data sets (Figs 4c,d). This difference implies a significant amount of uncertainty in the observation over the Indian Ocean.

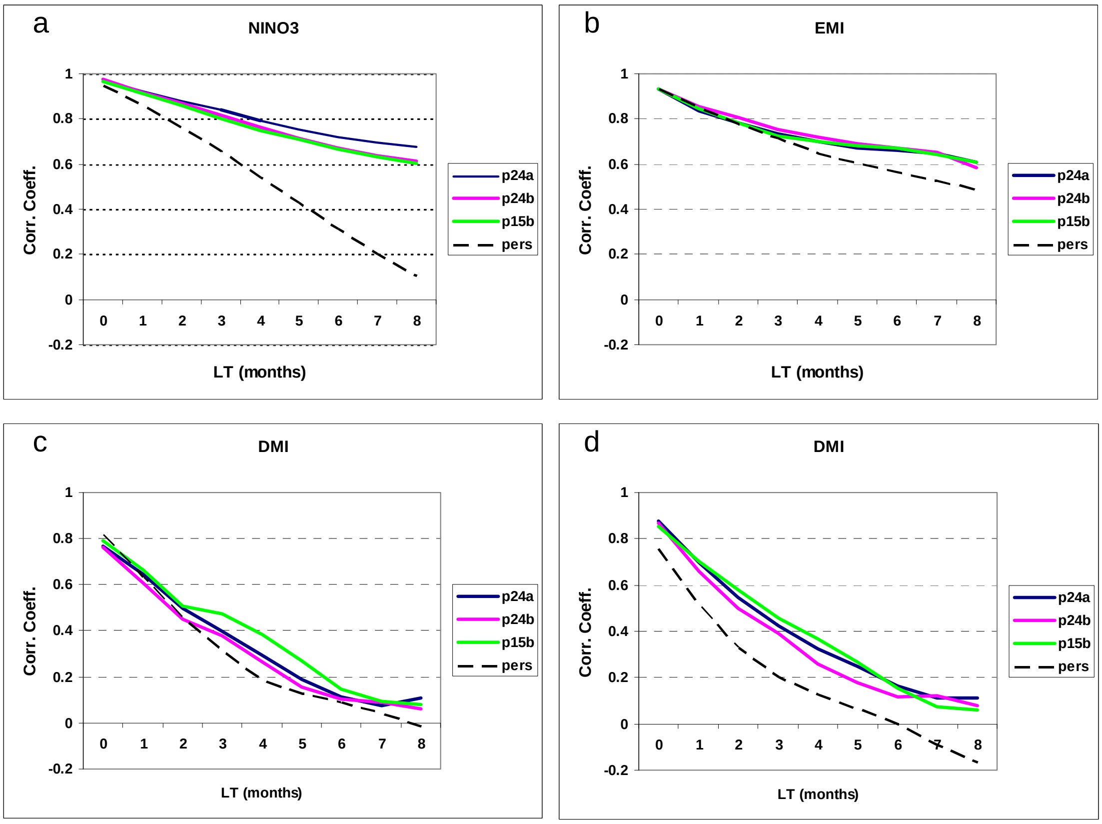


Fig. 4 Correlation of predicted (a) NINO3, (b) EMI and (c) DMI from POAMA1.5b (P1.5b), POAMA2 non-flux corrected version (P2.4a), POAMA2 flux corrected version (P2.4b) and persistence forecast verified against respective observed indices using the HadISST data set. Every month in 1980-2006 was used in forming the indices. (d) The same as (c) except verified against the observed DMI from the Reynolds SST data set.

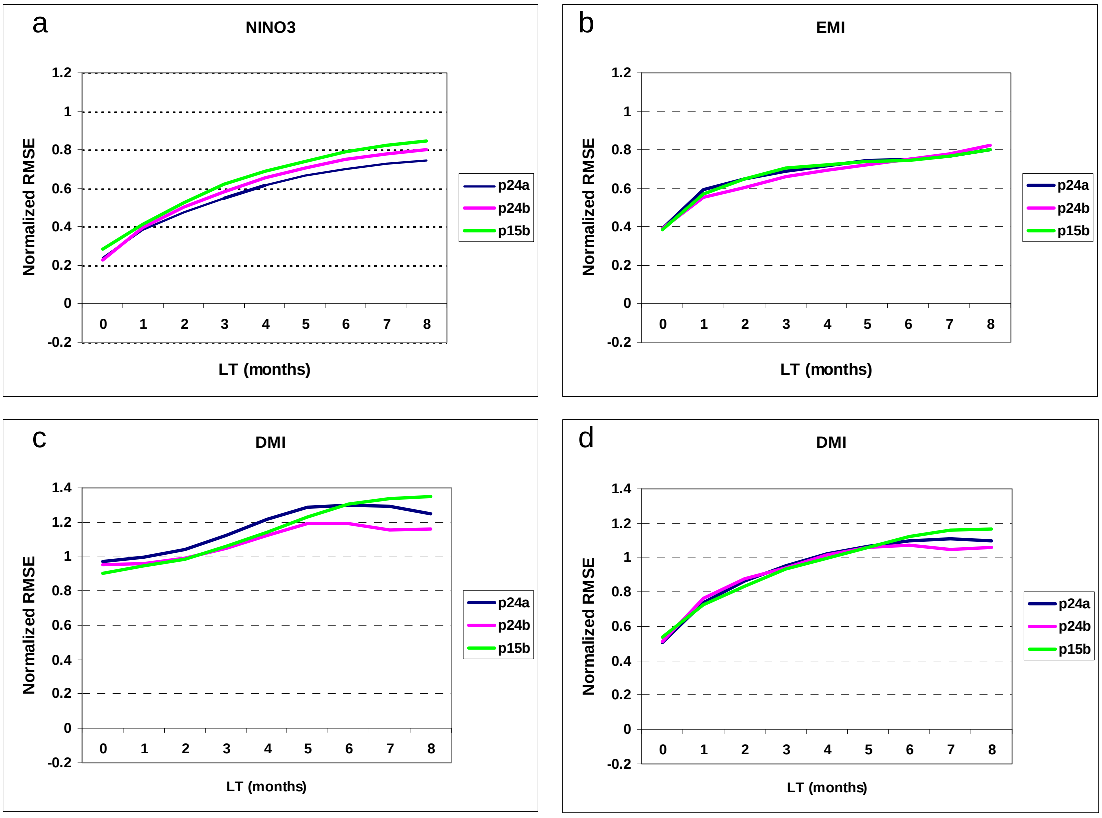


Fig. 5 Normalised root-mean-square-error (NRMSE) of predicted (a) NINO3, (b) EMI and (c) DMI from P1.5b, P2.4a and P2.4b. RMSEs of the predicted indices were normalised by the standard deviations of the respective observed indices from the HadISST data set. Every month in 1980-2006 was used in forming the indices. (d) The same as (c) except using the observed DMI from the Reynolds SST data set.

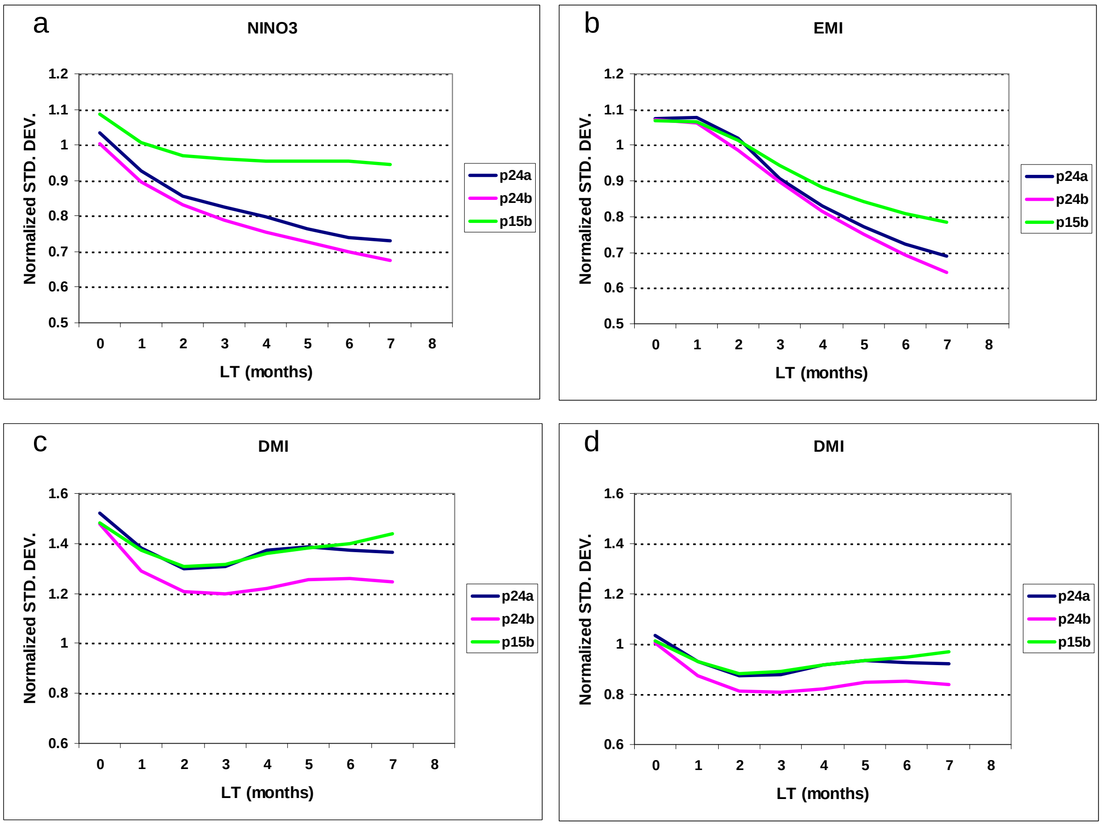


Fig. 6 Normalised standard deviation of predicted (a) NINO3, (b) EMI and (c) DMI from P1.5b, P2.4a and P2.4b (Standard deviations of the predicted indices were normalised by the standard deviations of the respective observed indices from the HadISST data set). Every month in 1980-2006 was used in forming the indices. (d) The same as (c) except using the observed DMI from the Reynolds SST data set.

3.2 Australian rainfall

Prediction skill for Australian rainfall is assessed by calculating the proportion of correct forecasts for exceeding median rainfall (Wilks 2006), based on the 2x2 contingency table shown in Fig. 7:

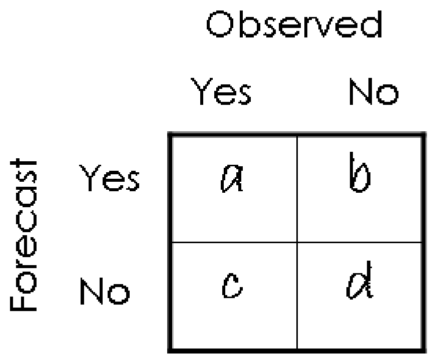


Fig. 7 The 2x2 contingency table of predicting a dichotomous event. The letters a, b, c, and d indicate the frequencies of four different types of forecast and observation pairs for a dichotomous event (taken from Lim et al. 2009, their figure 10).

$$proportion \quad correct \quad = \frac{a + d}{a + b + c + d}$$

Probabilistic forecasts for above median rainfall were computed using the 10 ensemble members for each version of POAMA. Medians of seasonal rainfall at each model grid point were obtained from the hindcasts in a cross-validated fashion at each lead time. The same procedure of finding the median and determining rainfall being above or below the median was applied to the observed gridded rainfall data set for the period 1980-2006.

Figure 8 displays proportion correct of predicting rainfall being above the median from P1.5b, P2.4a and P2.4b for lead time 0. In general, POAMA2 rainfall forecasts show higher proportion correct in late autumn (May-June-July; MJJ) to winter (June-July-August; JJA) and summer (December-January-February; DJF) compared to POAMA1.5b but lower proportion correct in late summer (February-March-April; FMA) to autumn (March-April-May; MAM). High proportion correct in south eastern Australia in winter-spring is consistent with this region being strongly influenced by ENSO/IOD, coupled with the good ability of POAMA to predict ENSO/IOD.

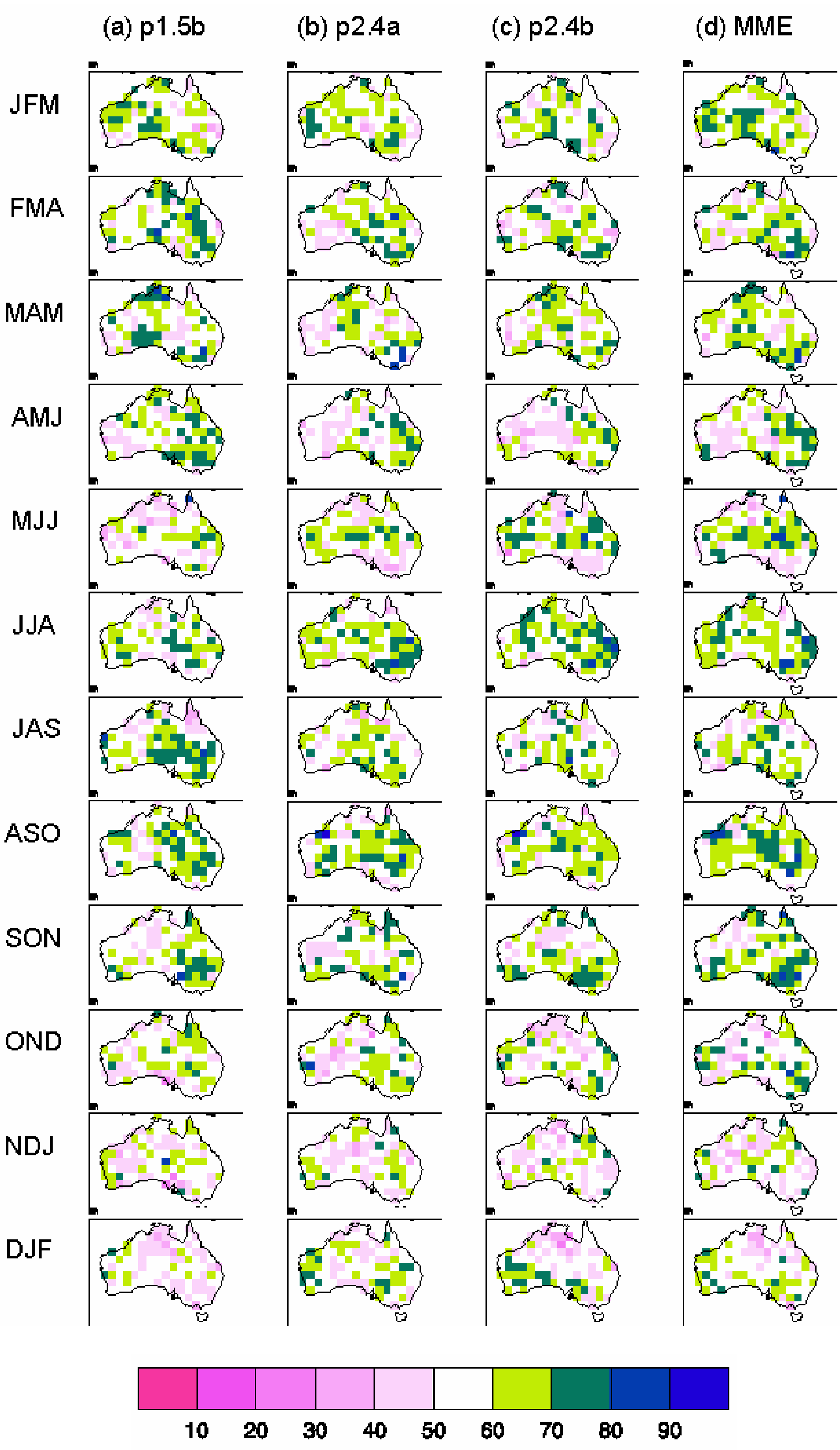


Fig. 8 Proportion correct of forecasts for rainfall being above the median at lead time 0 (LT 0) for the hindcast period (1980-2006). (a) P1.5b, (b) P2.4a, (c) P2.4b and (d) multi-model ensemble system consisting of the three versions of POAMA.

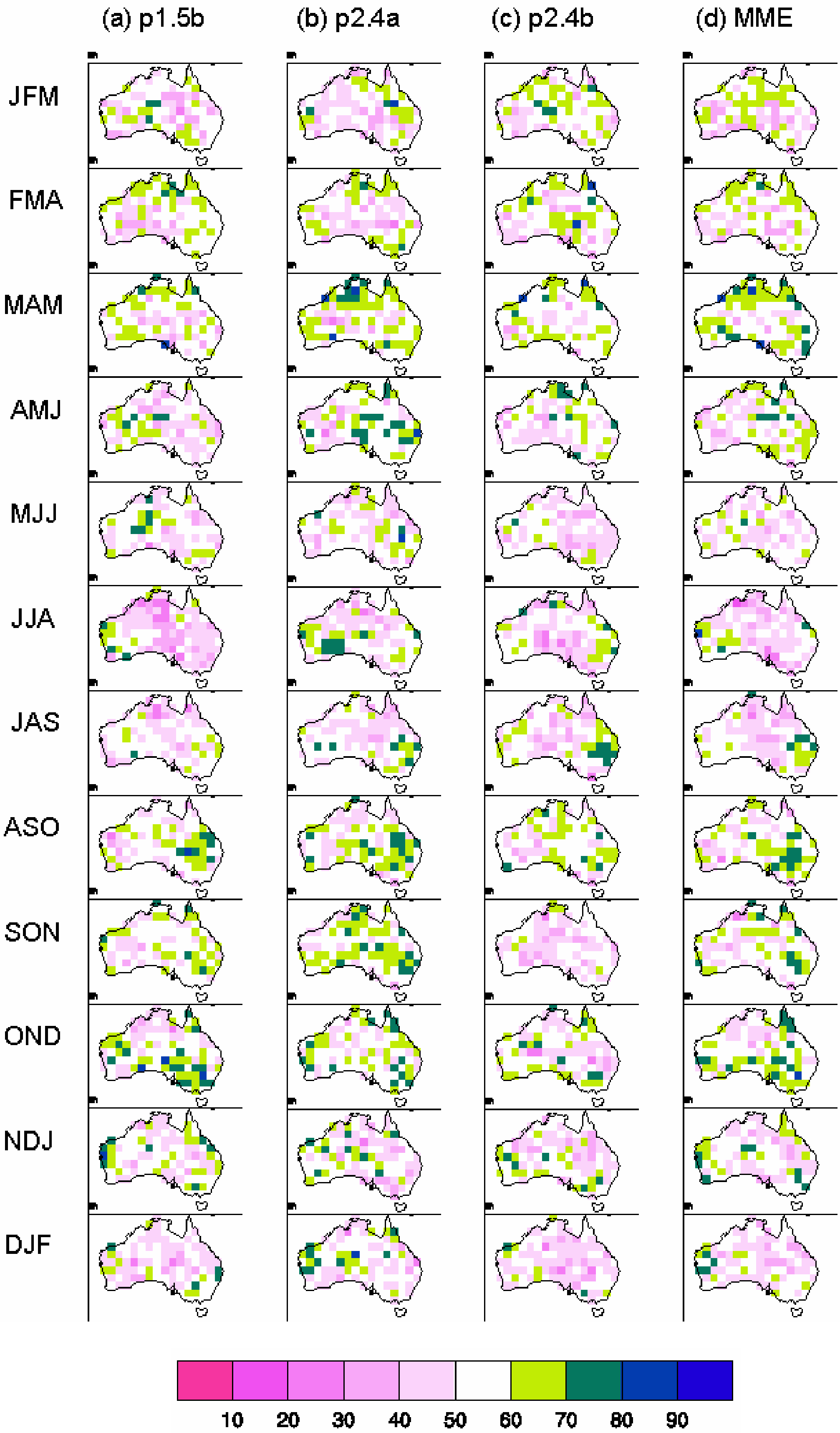


Fig. 9 The same as Fig. 8 except at lead time 3 (LT 3) months.

As it appears that all three versions demonstrate large areas of moderate to good proportion correct (>= 60%), and each version has its own strengths and weaknesses over different locations in different seasons, we combined all these three versions together to form a multi-model ensemble. The *overall* performance of a multi-model ensemble system is suggested to be better than the individual component models as a result of offsetting errors and increased ensemble spread (Doblas-Reyes et al. 2000, Palmer et al. 2004, Hagedorn et al. 2005, Weigel et al. 2008). Here, the POAMA multi-model ensemble system (POAMA MME) was formed simply by pooling all ensemble members of the three versions together (30 members in total). At lead time 0 POAMA MME forecasts appear to be more skilful than POAMA2 in all seasons except for DJF and more skilful than POAMA1.5b in all seasons except for FMA, April-May-June (AMJ) and July-August-September (JAS) (Fig. 8d).

By three month lead time (Fig. 9) all three versions of POAMA provide forecasts whose overall skill is no better than that of a climatological forecast. There is some indication that P2.4a is slightly more skilful than the other two versions (Figs 9a-c). Because the individual component versions do not have much skill in predicting rainfall at three month lead time, the effect of combining three component models is not obvious (Fig. 9d).

In order to make comparisons between the different forecast systems easier in all seasons at different lead times, we summarize the skill assessment based on proportion correct by computing the fraction of the grid points where proportion correct is equal to or greater than 60% (which is the threshold of a skilful forecast used in this study) multiplied by the mean proportion correct of those grid points. A perfect forecast at all grid points would yield a score of 1 by this measure (100% of the grid points multiplied by 100% proportion correct over those grid points). This weighted mean skill score is displayed in Fig. 10. Consistent with Fig. 8, at lead time 0 the weighted mean skill score for the two versions of POAMA2 is higher than POAMA1.5b in late autumn through to winter and in summer and is slightly higher in spring as well (Fig 10a). The weighted mean skill score of the POAMA MME is higher than that of any single POAMA version for most of the year at the shortest lead time. Improvement by this multi-model ensemble approach is achieved at the short lead times (0-1 month) for which the individual POAMA versions have reasonably good skill (Figs 10a,b). However, as lead time increases, the POAMA MME does not outperform the best single model, which appears to be P2.4a (Figs 10c-e).

In addition to proportion correct, reliability and resolution are important forecast qualities in probabilistic forecasts. The attributes diagram (Wilks 2006) is widely used as a compact way to display reliability and resolution as well as some other forecast features. Figure 11 displays the attributes diagrams of rainfall forecasts for above median using all grid points over Australia at lead time 0 based on the entire 27 year record of hindcasts. In Fig. 11, the diagonal line indicates perfect reliability (e.g. the accuracy of forecast of 70% chance of being above median rainfall should be 70%). Forecasts falling in the grey areas of Fig. 11 are considered to be reliable forecasts as they correctly indicate the occurrence/non-occurrence of the exceedance of median rainfall and have smaller magnitudes of error than a climatological forecast. The size of dots in Fig. 11 represents the frequency of forecasts in each probability category.

The overall distribution of the probabilistic forecasts from all three versions of POAMA suggests that the forecasts tend to be overconfident: The POAMA forecasts of above median rainfall occur more often than observed. Likewise, the POAMA forecasts of being below median occur more often then observed (i.e. POAMA is overly confident in its prediction that rainfall will not be above the median). In contrast to the poor forecast reliability, forecast resolution is good as indicated by different probability forecasts being followed by different observed outcomes (i.e. higher probability forecasts of an event (here, above median rainfall) are followed by higher frequency of the observed event).

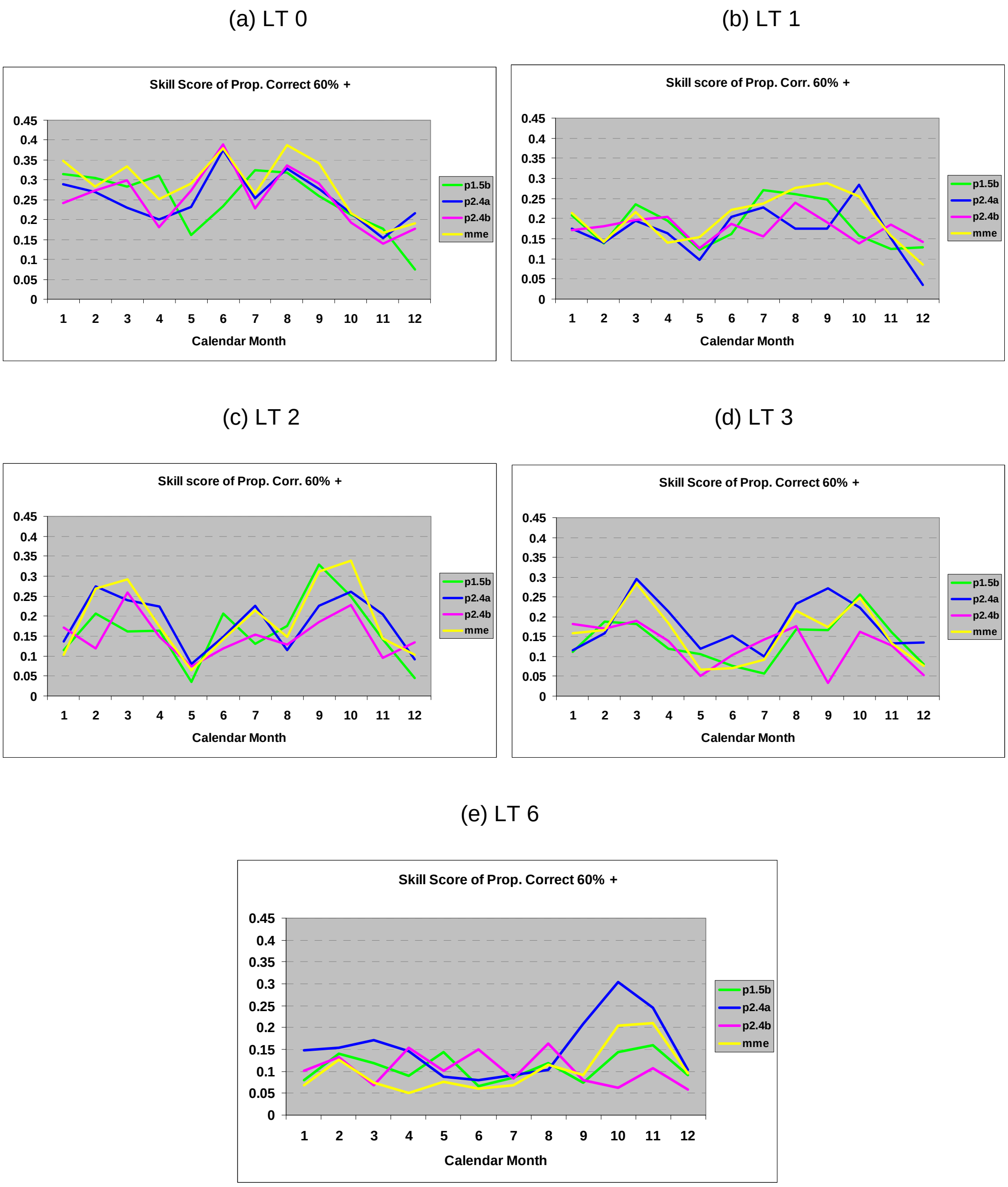


Fig. 10 Skill score using proportion correct that is computed by the multiplication of the fraction of the number of grid points over Australia whose proportion correct is equal to or greater than 60% by the average proportion correct over those grid points.

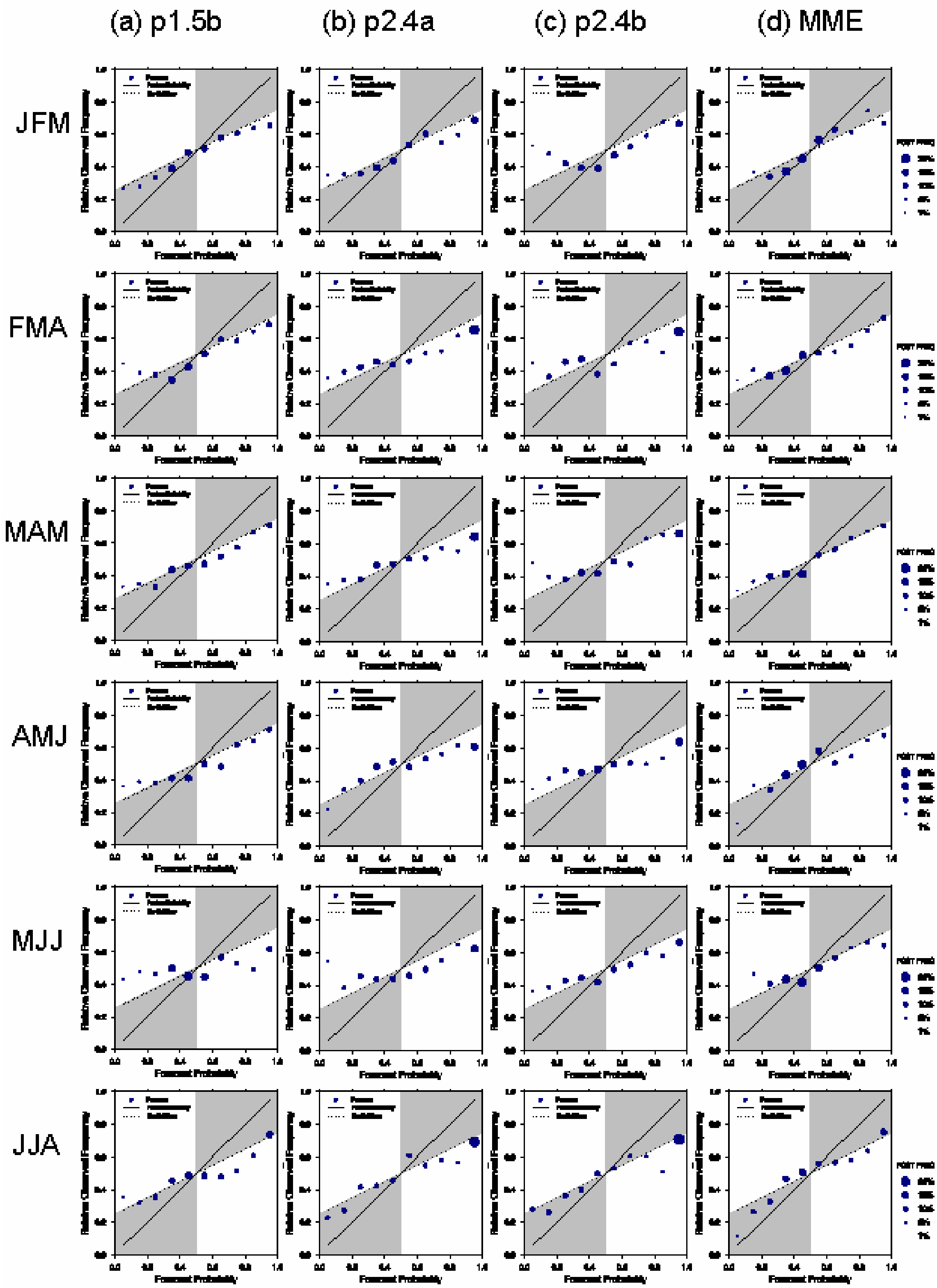


Fig. 11 Attributes diagram of POAMA prediction of above median rainfall at LT 0, using 27 years of hindcasts over all grid points of Australia. (a) P1.5b, (b) P2.4a, (c) P2.4b and (d) multi-model ensemble system consisting of the three versions of POAMA. The Y-Axis of each diagram indicates relative observed frequency, and the X-Axis indicates Forecast probability

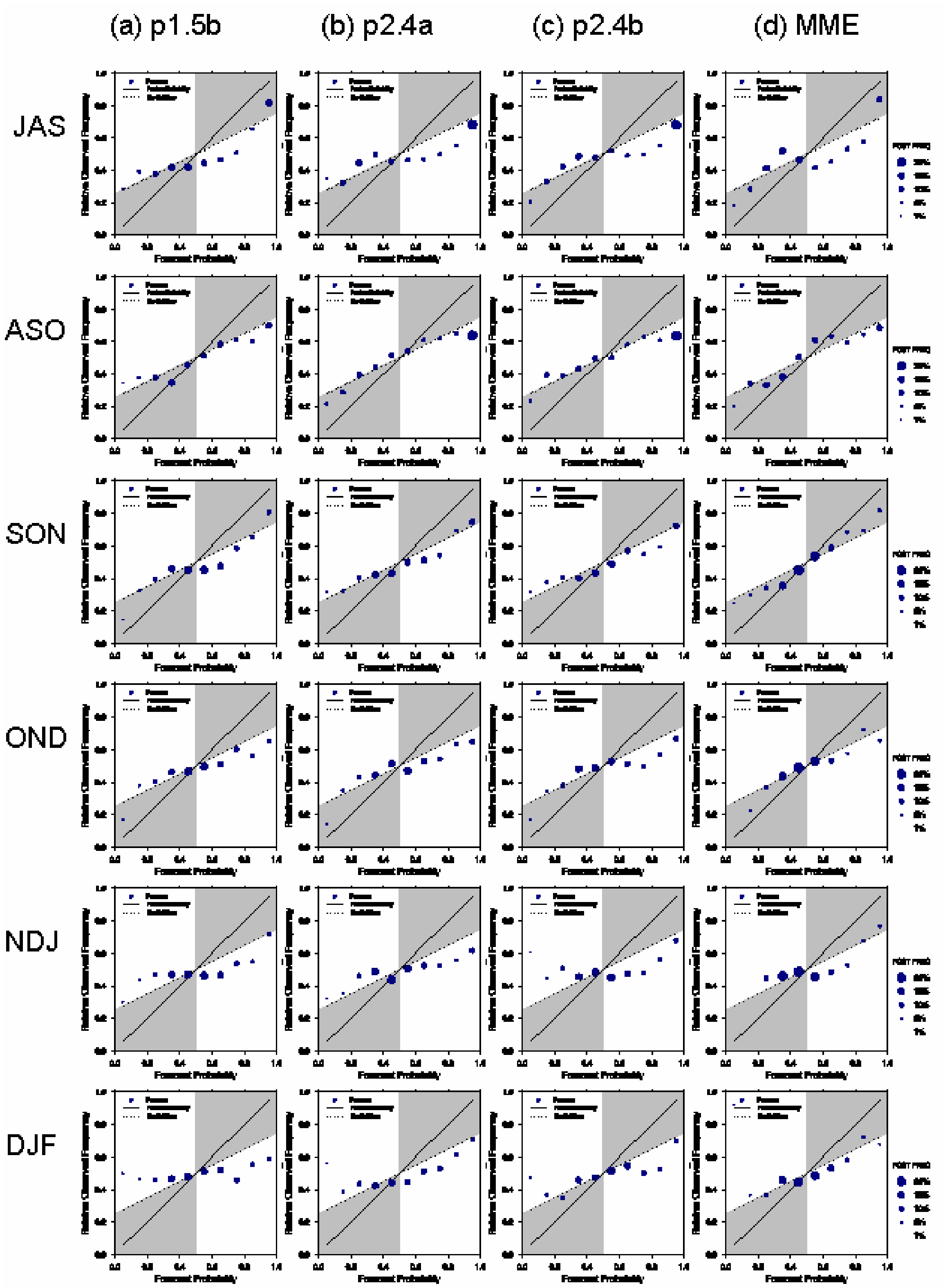


Fig. 11 *be continued.*

Figure 11d suggests that at the shortest lead time, the MME approach improves forecast reliability especially in late autumn to winter (MJJ, JJA) and spring (September-October-November, October-November-December) and improves forecast sharpness in summer (DJF, January-February-March).

In order to quantify reliability and resolution displayed in Fig.11, we computed mean reliability error (REL) and resolution (RES) following Wilks (2006):

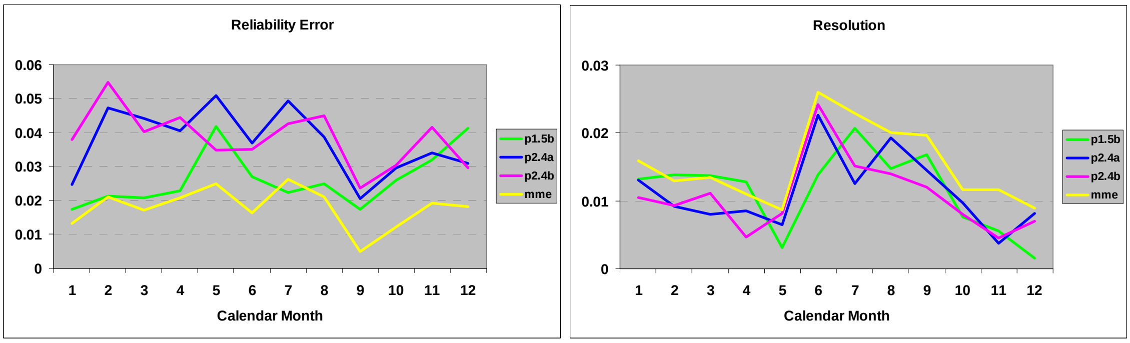
$$REL = \frac{1}{n}\sum_{i=1}^{I} N_i (F_i - \overline{O}_i)^2$$

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

where n is the total number of the forecasts, i indicates forecast categories (10% interval) and i runs from 1 to I (I=10 with 10% interval in this study), F_i is the forecast probability of rainfall being above the median that falls into the i[th] 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.

REL and RES of the three versions of POAMA and POAMA MME at lead time 0 are displayed in Figs 12a and 12b, respectively. Consistent with the above discussion, the POAMA MME significantly reduces reliability error while it improves forecast resolution. This is an encouraging result as it is difficult to increase forecast resolution without adding independent information to the prediction system (Stephenson et al. 2005, Doblas-Reyes et al. 2006). A more modest improvement in reliability made by POAMA MME over the individual versions of POAMA is found at lead time of three months (Fig 12c), but this improvement is achieved by having more forecasts between 40-60% forecast categories (not shown). Forecast resolution is significantly reduced by lead time 3 months in all four POAMA forecast systems (Fig 12d).

(a) LT 0



(b) LT 3

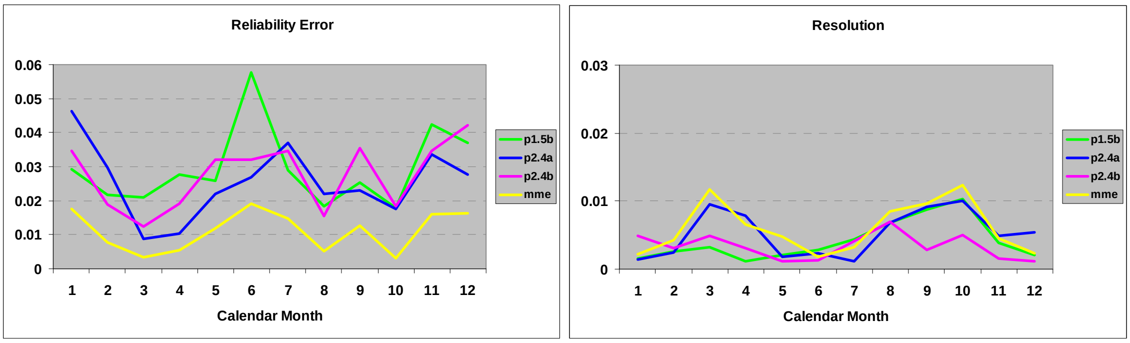


Fig. 12 Mean reliability error (left) and mean resolution (right) of the forecasts from the three versions of POAMA and the multi-model ensemble system in predicting above median rainfall at (a) LT 0 and (b) LT 3.

4 SUMMARY AND CONCLUDING REMARKS

We have examined the skill of hindcasts produced from the improved POAMA2 system. We have focused on predictions of tropical Indo-Pacific SST and Australian rainfall and made comparisons with the current operational version of POAMA (P1.5b). A major upgrade made for the POAMA2 system is to implement a new state-of-the art ensemble-based ocean data assimilation scheme, PEODAS (Yin et al. 2010). PEODAS assimilates in situ temperature and salinity, taking state-dependent background field errors into account. PEODAS provides more accurate and realistic ocean initial conditions than the previous univariate scheme employed in the POAMA1 system. The ensemble of these ocean states provided by PEODAS is used to initialize the atmosphere-ocean coupled model, and therefore, to generate forecast ensemble.

In addition to documenting the impact of improved ocean initial conditions provided by PEODAS for forecast skill, we also explored the impact on forecast skill of removing the mean state bias by applying an explicit flux correction scheme to POAMA2 (P2.4b). The scheme successively reduces the tropical-wide cold SST bias and more regional warm bias off the west coast of South America that develops with increasing lead time in P2.4a. Two sets of hindcasts from the different versions of POAMA2 – non-flux corrected P2.4a and flux corrected P2.4b – were analysed together with the hindcasts from P1.5b. Hindcasts for the period 1980-2006 were generated from each of these three versions of POAMA and were verified against HadISST and the National Climate Centre monthly rainfall analysis.

Our results suggest that there are strengths and weaknesses in POAMA2 in predicting tropical SST: P2.4a has improved skill in predicting the occurrence of El Niño in the eastern Pacific as captured by the NINO3 index. P2.4b (flux corrected) has improved skill in predicting El Niño Modoki (central Pacific El Niño) as depicted by EMI, compared to the non-flux corrected versions of POAMA - P1.5b and P2.4a. However, the amplitudes of NINO3 and EMI are significantly underestimated in the two versions of POAMA2. And, the overall skill in predicting the behavior of the IOD is less skilful with POAMA2 than with POAMA1.5b.

Australian rainfall is predicted with near equal skill with all three versions of the model at short lead times (0-1 month), but P2.4a slightly outperforms the other two at longer lead times. A multi-model ensemble system consisting of the three different versions of POAMA demonstrates higher proportion correct than any single version of POAMA, as well as improved reliability and resolution at short lead times of 0-1 month. However, at lead times longer than 3 months the POAMA multi-model ensemble is no better than the individual members. This is because at longer lead times each of the versions that goes into the multi model ensemble does not have skill and a multi-model ensemble cannot create skill although it can improve existing skill.

All versions of POAMA demonstrate high skill to predict tropical Pacific SST with up to 8 month lead times, and P2.4a shows a marked improvement over P1.5b. Therefore, there still remains a possibility to increase rainfall forecast skill at longer lead times by some statistical post-processing if hindcasts can be generated for a sufficiently long period. For instance, during the period 1980-2006, 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 for future work with POAMA.

ACKNOWLEDGMENTS

This research was supported in part by the South Eastern Australian Climate Initiative (SEACI; http://www.seaci.org Advancing Seasonal Predictions for SEA) and by the Managing Climate Variability project "Improving Seasonal Forecasts for SWWA" (Development of bias correction scheme and evaluation of impact on SWWA). The authors are grateful to Andrew Charles in the National Climate Centre and Claire Spillman in the CAWCR for their useful comments on the manuscript.

REFERENCES

Alves, O., et al., 2003. POAMA: Bureau of Meteorology operational coupled model seasonal forecast system. Science for drought - Proceedings of the National Drought Forum.,Brisbane, Australia, Queensland Government.

Ashok, K., Behera, S.K., Rao, S.A. and Weng, H. 2007. El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C1107, doi:10.1029/2006JC003 798.

Burgers, G., Balmaseda, M.A., Vossepoel, F.C., van Oldenborgh, G. and van Leeuwen, P. 2002. Balanced ocean-data assimilation near the equator. *J. Phys. Oceanogr.*, **32**, 2509–19.

Colman, R., Deschamps, L., Naughton, M., Rikus, L., Sulaiman, A., Puri, K., Roff, G., Sun, Z. and Embery, G. 2005. *BMRC Atmospheric Model (BAM) version 3.0: Comparison with mean climatology.* BMRC Research Report No. 108, Bureau of Meteorology Research Centre, 32 pp. (available from http://www.bom.gov.au/bmrc/pubs/researchreports/researchreports.htm).

Doblas-Reyes, F.J., Deque, M. and Piedelievre, J.-P. 2000. Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Quart. J. Roy. Meteor. Soc.*, **126**, 2069-88.

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

Hagedorn, R., Doblas-Reyes, F.J. and T.N. Palmer, 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept. *Tellus*, **57A**, 219-33.

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

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

Ingleby, B., and Huddleston, M. 2007. Quality control of ocean temperature and salinity profiles — Historical and real-time data. *J. Mar. Sys.,* **65**, 158-75.

Jones, D.A. and Weymouth, G. 1997. *An Australian monthly rainfall dataset.* Technical Report 70, Bureau of Meteorology.

Large, W.G. and Yeager, S.G. 2009. The global climatology of an interannually varying air-sea flux data set. *Climate Dyn.,* **33**, 341-64.

Lim, E.-P., Hendon, H.H., Hudson, D., Wang, G. and Alves, O. 2009. Dynamical forecast of inter-El Niño variations of tropical SST and Australian spring rainfall. *Mon. Wea. Rev.*, **137**, 3796-3810.

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

Oke, P.R., Schiller, A., Griffin, D.A. and Brassington, G.B. 2005. Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Q. J. Roy. Met. Soc.*, **131**, 3301-11.

Palmer, T.N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délécluse, P., Déqué, M., Díez, E., Doblas-Reyes, F.J., Feddersen, H., Graham, R., Gualdi, S., Guérémy, J.-F., Hagedorn, R., Hoshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A.P., Orfila, B., Rogel, P., Terres, J.-M. and Thomson, M.C. 2004. Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). *Bull. Amer. Met. Soc.*, **85**, 853-72.

Rayner, N.A., Parker, D.E., Horton, E., Folland, C.K., Alexander, L.V., Rowell, D.P., Kent, E.C. and Kaplan, A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108 (D14)**, 4407, doi:10.1029/2002JD002 670.

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C. and Wang, W. 2002. An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–25.

Risbey, J.S., Pook, M.J., McIntosh, P.C., Wheeler, M.C. and Hendon, H.H. 2009. On the remote drivers of rainfall variability in Australia. *Mon. Wea. Rev.*, **137**, 3233-53.

Saji, N., Goswami, B.N., Vinayachandran, P.N. and Yamagata, T. 1999. A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360–3.

Schiller, A., Godfrey, J.S., McIntosh, P.C., Meyers, G., Smith, N.R., Alves, O., Wang, G. and Fiedler, R. 2002. *A New Version of the Australian Community Ocean Model for Seasonal Climate Prediction*. CSIRO Marine Research Report No. 240.

Smith, N.R., Blomley, J.E. and Meyers, G. 1991. A univariate statistical inter- polation scheme for subsurface thermal analyses in the tropical oceans. *Prog. Oceanog.*, **28**, 219-56.

Stephenson, D.B., Coelho, C.A.D.S., Doblas-Reyes, F.J. and Balmaceda, M. 2005. Forecast assimilation: A unified framework for the combination of multimodel weather and climate predictions, *Tellus*, **57A**, 253-64.

Stockdale, T.N., Anderson, D.L.T., Alves, J.O.S. and Balmaseda, M.A. 1998. Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature,* **392**, 370-3.

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

Valke, S., Terray, L. and Piacentini, A. 2000. The OASIS coupled user guide version 2.4, *Technical Report* TR/ CMGC/00-10, CERFACS.

Wang G., Kleeman, R., Smith, N. and Tseitkin, F. 2002. The BMRC coupled general circulation model ENSO forecast system. *Mon. Wea. Rev*. **130,** 975-91.

Weigel, A.P., Liniger, M.A. and Appenzeller, C. 2008. Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Q. J. R. Meteorol. Soc.*, **134**, 241-60.

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

Yin, Yonghong., Alves, O. and Oke, P.R. 2010. An ensemble ocean data assimilation system for seasonal prediction. *Submitted to Mon. Wea. Rev.*

APPENDIX A – PROCEDURE OF FLUX CORRECTION

Step 1. The Bureau of Meteorology's Atmospheric Model v3.0 (BAM) is run with observed SST from Reynolds SST data set for the period 1984-2003, and the monthly climatologies of wind stress (τ_{BAM}) and shortwave radiation (Qs_{BAM}) from the model are computed ($\overline{\tau}_{BAM}$ and $\overline{Q}s_{BAM}$).

Step 2. The ocean and atmosphere coupled model (ACOM2 + BAM) is freely run for 20 years (being started with observed SST in January, 1984)

a. Wind stress and shortwave radiation in the atmosphere model in the coupled run, τ_C and $Qs_{C,}$ are corrected at each time step before being passed to the ocean model as below:

$$\tau_{mC} = \tau_C - (\overline{\tau}'_{BAM} - \overline{\tau}'_{OBS})$$

$$Qs_{mC} = Qs_C - (\overline{Q}s'_{BAM} - \overline{Q}s'_{OBS})$$

where $\overline{\tau}_{OBS}$ and $\overline{Q}s_{OBS}$ are observed monthly climatologies of wind stress and shortwave radiation that are obtained from the CORE data set (Large & Yeager 2009). $'$ indicates daily value interpolated from monthly climatology.

b. Total heat flux, $Q_{TotC,}$ is modified as

$$Q_{mTotC} = Q_{TotC} + \Delta Q = Q_{TotC} - \lambda(T_C - \overline{T}'_{OBS})$$

where λ= 40 W/m^2/°C, T_C is the coupled model SST at each time step and $\overline{T}'_{OBS}$ is observed SST interpolated daily from monthly climatology.

c. Compute the monthly climatology of ΔQ , $\Delta\overline{Q}$.

Step 3. Run the coupled model in forecast mode (i.e. initialize it on the 1st of each month and run for 9 months) for the hindcast period, having its wind stress, shortwave radiation and total heat flux, τ_F , Qs_F and Q_{TotF} corrected at each time step by daily values of monthly climatologies of $\overline{\tau}_{BAM}$, $\overline{\tau}_{OBS}$, $\overline{Q}s_{BAM}$, $\overline{Q}s_{obs}$, and $\Delta\overline{Q}$ as follows:

$$\tau_{mF} = \tau_F - (\overline{\tau}'_{BAM} - \overline{\tau}'_{OBS})$$

$$Qs_{mF} = Qs_F - (\overline{Q}s'_{BAM} - \overline{Q}s'_{OBS})$$

$$Q_{mTotF} = Q_{TotF} + \Delta\overline{Q}'$$

The Centre for Australian Weather and Climate Research is a partnership between CSIRO and the Bureau of Meteorology.