Reference No. 3.2.2

South Eastern Australian
Climate initiative

Review of techniques to bridge/calibrate dynamical seasonal predictions with focus on south eastern Australia

Harry Hendon, Eunpa Lim, Oscar Alves and Guomin Wang

h.hendon@bom.gov.au
BMRC, 700 Collins Street, Melbourne, Australia

15 January 2007

1. Introduction

The Bureau of Meteorology routinely makes dynamical seasonal predictions out to a 9 month lead time with the POAMA coupled ocean-atmosphere forecast system. POAMA (Predictive Ocean Atmosphere Model for Australia; Alves et al. 2003) is an intra-seasonal to inter-annual climate prediction system based on coupled ocean and atmosphere general circulation models. The first version (POAMA-1) was developed in a joint project involving the Bureau of Meteorology Research Centre (BMRC), CSIRO Marine Research (CMR) and Land and Water Australia. POAMA-1 became operational in October 2002. The main focus for POAMA is the prediction of sea surface temperature (SST) anomalies associated with El Niño/La Niña, for which POAMA-1's predictions are internationally competitive. El Niño/Southern Oscillation (ENSO) is the dominant driver of Australian climate variability, thus the POAMA forecasts have great value for anticipating the behavior of El Niño. However, low model resolution and model bias and drift hinder the direct utilization of regional climate prediction for Australia from POAMA-1.

The POAMA system is continually evolving and improving, and subsequent versions of POAMA will address the problematic bias and drift and will have improved horizontal resolution so as to provide skilful prediction of regional climate variability. Future development of the components of the POAMA system will be done as part of the ACCESS (Australian Community Climate and Earth System Simulator) project. ACCESS is a joint Bureau, CSIRO, and Australian Universities project that aims at coordinated development of core components of earth system models and data assimilation systems to support a range of applications, including POAMA's seasonal prediction.

In the meantime, bridging and calibrating techniques, which capitalize on those components of the climate system for which POAMA-1 provides skilful prediction and which have a tight connection to Australian climate (i.e., tropical sea surface temperature variations associated with El Niño), can be explored. This report focuses on statistical techniques to improve seasonal climate prediction in south eastern Australia as part of the South Eastern Australian Climate Initiative (SEACI). Practices at other national meteorological centres and research institutions are reviewed, and recommendations are made for some exploratory trials.

2. Current Status of POAMA-1

The POAMA system was designed to predict tropical coupled variability associated with ENSO. This focus on ENSO is motivated by the knowledge that ENSO is the dominant driver of the predictable interannual variations of Australian climate, especially in eastern Australia during winter-spring. For instance, Figure 1 shows composite winter-spring rainfall anomalies across Australia for the 12 strongest El Niño and La Niña events of the 20^{th} century. Much of eastern Australia experiences well below normal rainfall during El Niño and enhanced rainfall during La Niña. The basis for seasonal climate prediction in Australia is the strong persistence of El Niño or La Niña conditions from late winter into spring and summer. For example, Figure 2 shows the correlation of winter Niño3.4 SST index with the following spring rainfall across Australia. Over much of eastern Australia 10-20% of the springtime rainfall variance can be anticipated by the knowledge of the state of El Niño in winter. Hence, accurate prediction of the state of ENSO should result in skilful predictions of Australian rainfall one to two seasons in advance, especially in eastern Australia during winter and spring.

There is also growing evidence that rainfall variability in eastern Australia is sensitive not only to the occurrence of El Niño or La Niña but to the details of the SST variations during each warm or cold event. Figure 3, from a recent study of the sensitivity of Australian rainfall to inter-El Niño variations by Wang and Hendon (2006), shows that Australian-mean rainfall in spring is most sensitive to SST variations near the dateline, which is well west of where the largest SST anomalies during El Niño typically develop. Hence, stronger droughts are associated with El Niño events whose SST anomalies are displaced westward of their typical location. Australian rainfall is also sensitive to SST anomalies outside of the tropical eastern Pacific. For instance, Victorian-mean rainfall in winter-spring is as sensitive to SST anomalies in the eastern Indian Ocean/Coral Sea and the western Indian Ocean as it is to SST anomalies in the equatorial eastern Pacific (Figure 4). Thus, a successful seasonal prediction system must not only predict the occurrence of ENSO but also the details of the SST anomalies both locally in the equatorial eastern Pacific and remotely in the western Pacific and Indian Ocean.

In order to predict ENSO and its oceanic and atmospheric teleconnections, the forecast model should be able to capture the full range of physical processes relevant to atmosphere-ocean interactions associated with ENSO events. Furthermore, forecasts

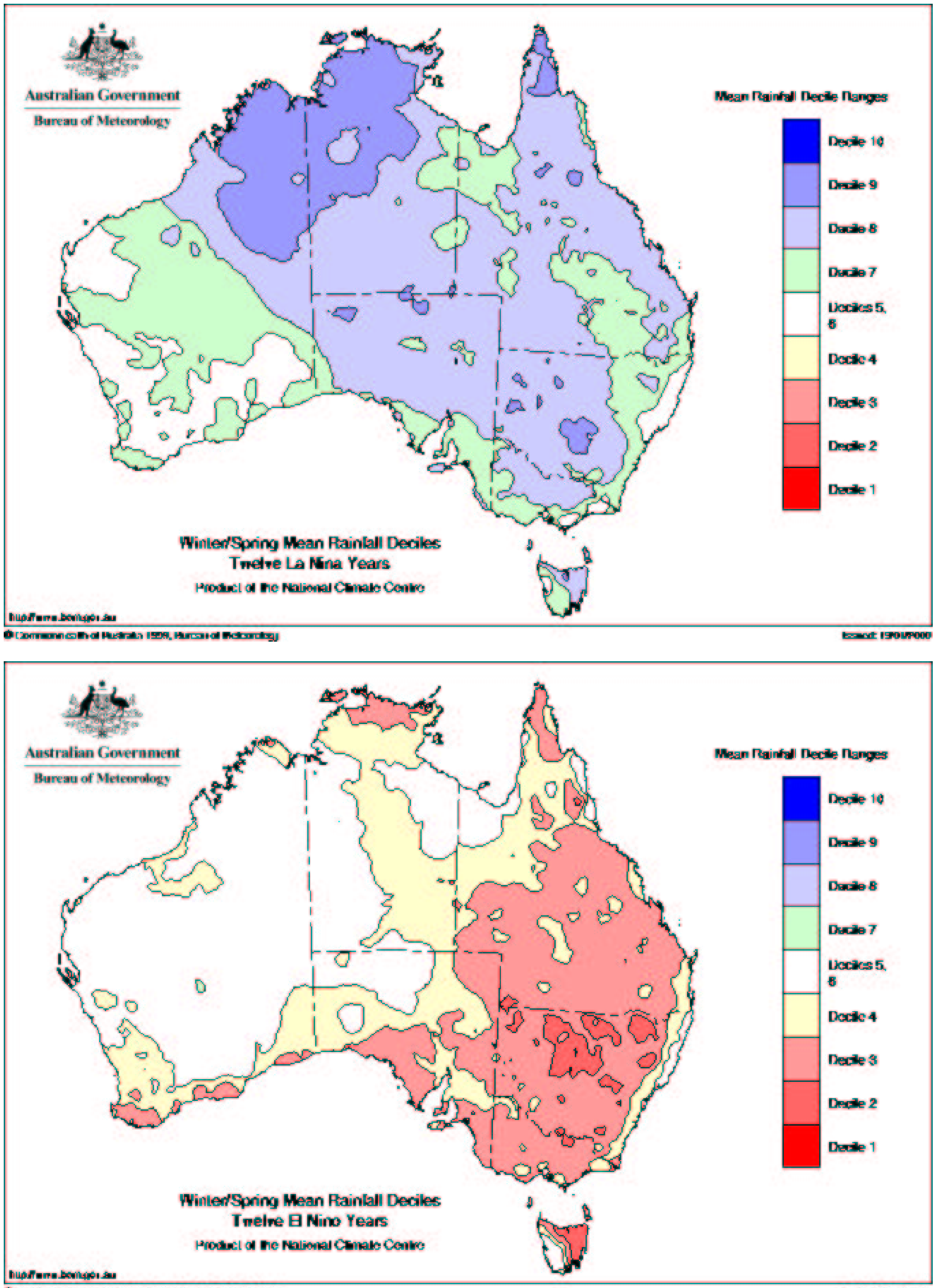

Figure 1: Winter-spring rainfall deciles for the twelve strongest La Niña (top) and El Niño (bottom) events in the 20^{th} century (La Niña years: 1910, 1916, 1917, 1938, 1950, 1955, 1956, 1971, 1973, 1975, 1988, 1998; El Niño years: 1905, 1914, 1940, 1941, 1946, 1965, 1972, 1977, 1982, 1991, 1994, 1997) (from http://www.bom.gov.au/climate/enso/ensorain.comp.shtml).

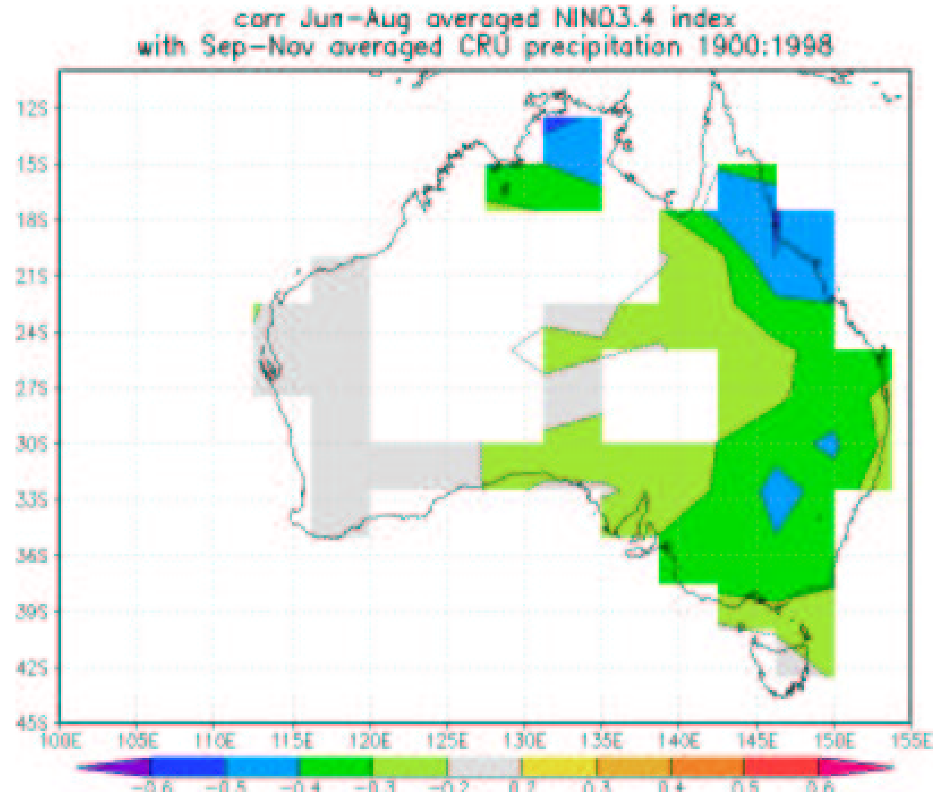

Figure 2: Correlation between June-July-August mean Niño3.4 SST index and the following September-October-November (SON) mean rainfall for the period 1900-1998. Calculation and data were provided by the KNMI Climate Explorer (from http://climexp.knmi.nl/start.cgi?someone@somewhere).

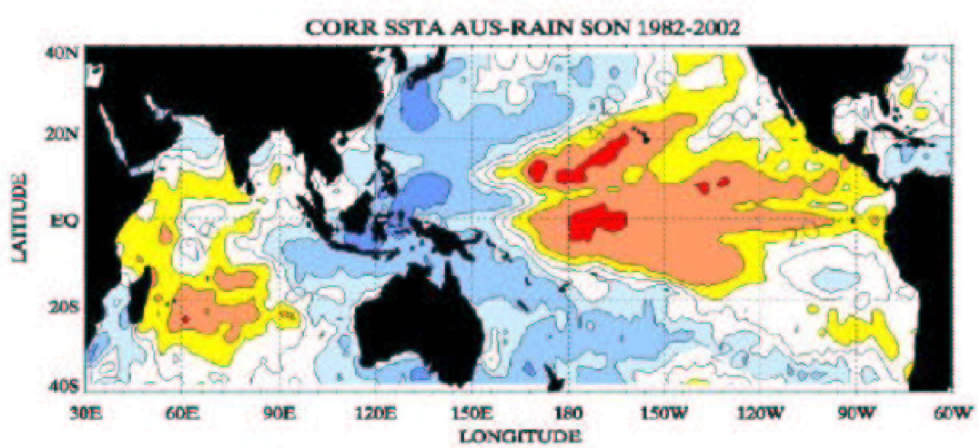

Figure 3: Correlation of SST with the negative of Australia-mean rainfall anomaly for the spring season (SON) 1982-2002 (from Wang and Hendon 2006). Blue shades indicate negative correlation and yellow/red shades indicate positive correlation. Shading level is 0.2.

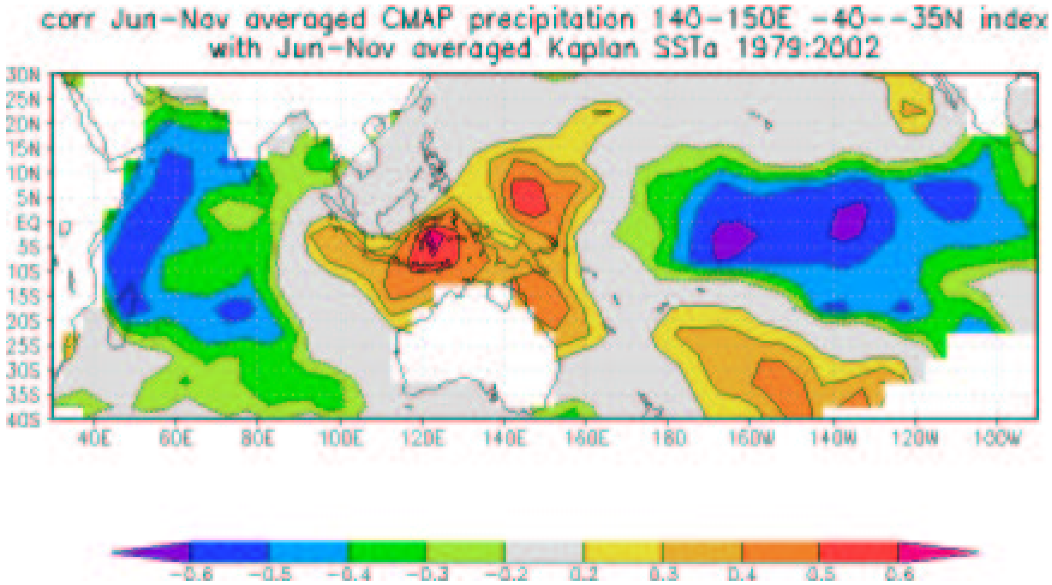

Figure 4: Correlation of SST with Victorian-mean rainfall in winter and spring for the period 1979-2002. Calculation and data were provided by the KNMI Climate Explorer (from http://climexp.knmi.nl/start.cgi?someone@somewhere).

need to be initialized with accurate atmospheric and upper ocean initial conditions in the Tropics where the predictability of El Niño stems. The POAMA system is based on comprehensive general circulation models of both the atmosphere and ocean, and incorporates a real time ocean data assimilation system. Atmospheric initial conditions are obtained in real time from the Bureau's global weather forecast model (GASP). The ocean data assimilation system provides the best estimate of the current state of the tropical upper ocean. It makes use of all available surface and subsurface temperature observations, including those from the TOGA-TAO moorings, drifters, and ARGO floats. An updated analysis is generated every 3 days.

POAMA-1 is based on version 3 of the Bureau's Atmospheric Model (BAM3; Colman et al. 2005) coupled to version 2 of the Australian Community Ocean Model (ACOM2; Schiller et al. 2002, Oke et al. 2005). The atmospheric model is run with modest horizontal resolution ($\sim$ 300 km resolution) and with 17 vertical levels. The ocean model is run with $\sim$ 200 km zonal resolution and telescoping meridional resolution to 0.5° latitude in the Tropics (i.e. the meridional resolution gradually decreases towards the Tropics). At the time of the development of POAMA-1 ($\sim$ 6 years ago) these resolutions were considered state of the art for a coupled seasonal forecast system based on general circulation models. At these resolutions, coupled behavior associated with ENSO, including atmospheric intra-seasonal variability that strongly interacts with ENSO, is well resolved, as are the associated global atmospheric and oceanic teleconnections. However, the primary deficiencies of this modest atmospheric resolution are the inability to resolve local climate variations associated with regional topography and orography (e.g. the Dividing Range is not well represented) and the fact that the extratropical storm tracks are too diffuse. The low zonal resolution of the ocean model results in an overly diffusive equatorial thermocline, which has a deleterious effect on the intensity of El Niño. Instability waves in the east Pacific and the narrow upwelling regions along some coastal boundaries are also not adequately resolved, but their effects in the climate system appear to be well represented by, for instance, parameterized diffusion (Wang et al. 2005).

A serious problem with coupled seasonal forecast models such as POAMA-1 is that the simulated climate drifts at long forecast lead times. This is demonstrated in Figure 5, which shows the bias of the climatology of SST from the POAMA-1 hindcasts (one forecast per month for the period 1987-2001) as a function of forecast lead times. Because the forecasts are initialized from observed ocean conditions, little bias is seen at lead times of 1-2 months . However, by the lead time of 3 months, a tropical-wide cold bias has developed, together with a warm bias off the coast of South America. By the lead time of 6 months the bias is nearly saturated. Of particular concern for ENSO prediction is the extension of the equatorial cold tongue into the western Pacific, and the warming of the SST in the stratus-upwelling region off the South American coast. A direct result of this bias is that the ENSO mode simulated by POAMA-1 tends to be shifted west relative to the observed mode and tends to exhibit an overly biennial tendency. Many of these issues are common to most climate models and resolving these issues is a major challenge to the international climate modeling community.

The atmospheric and oceanic teleconnections of ENSO into the Australian region are also negatively impacted by this climate drift. Subsequent versions of POAMA will strive to reduce the model systematic error, through improvements to the component models under the ACCESS project. Presently the model bias resulting from the mean SST is removed from the POAMA-1 forecasts by forming model anomalies without the lead-time dependent model climatology. However, more sophisticated methods are required to compensate for the deleterious impact of this bias on the teleconnection of ENSO into the Australian region, especially the rainfall teleconnection in south eastern Australia. We will review some techniques to compensate for systematic bias in Section 3.

Despite the significant bias of the mean state of the tropical climate in the POAMA-1 forecasts, especially at longer lead times, POAMA-1 exhibits useful skill for prediction of El Niño/La Niña with lead times of up to 9 months. This skill is indicated in Figure 6, which displays the anomaly correlation of the predicted Niño3 SST index as a function of forecast lead time in the period of 1987-2002. POAMA-1 readily beats persistence and demonstrates skill (correlation greater than 0.6) at lead times of 8-9 months. Such skill is typical of other dynamical and statistical forecast models run internationally. However, the direct prediction of rainfall from POAMA-1 appears to be of limited utility. The correlation of rainfall directly predicted from POAMA-1 with observed rainfall is shown in Figure 7. Only in western Australia is there any hint of skill. The lack of forecast skill for south eastern Australia is emphasized by

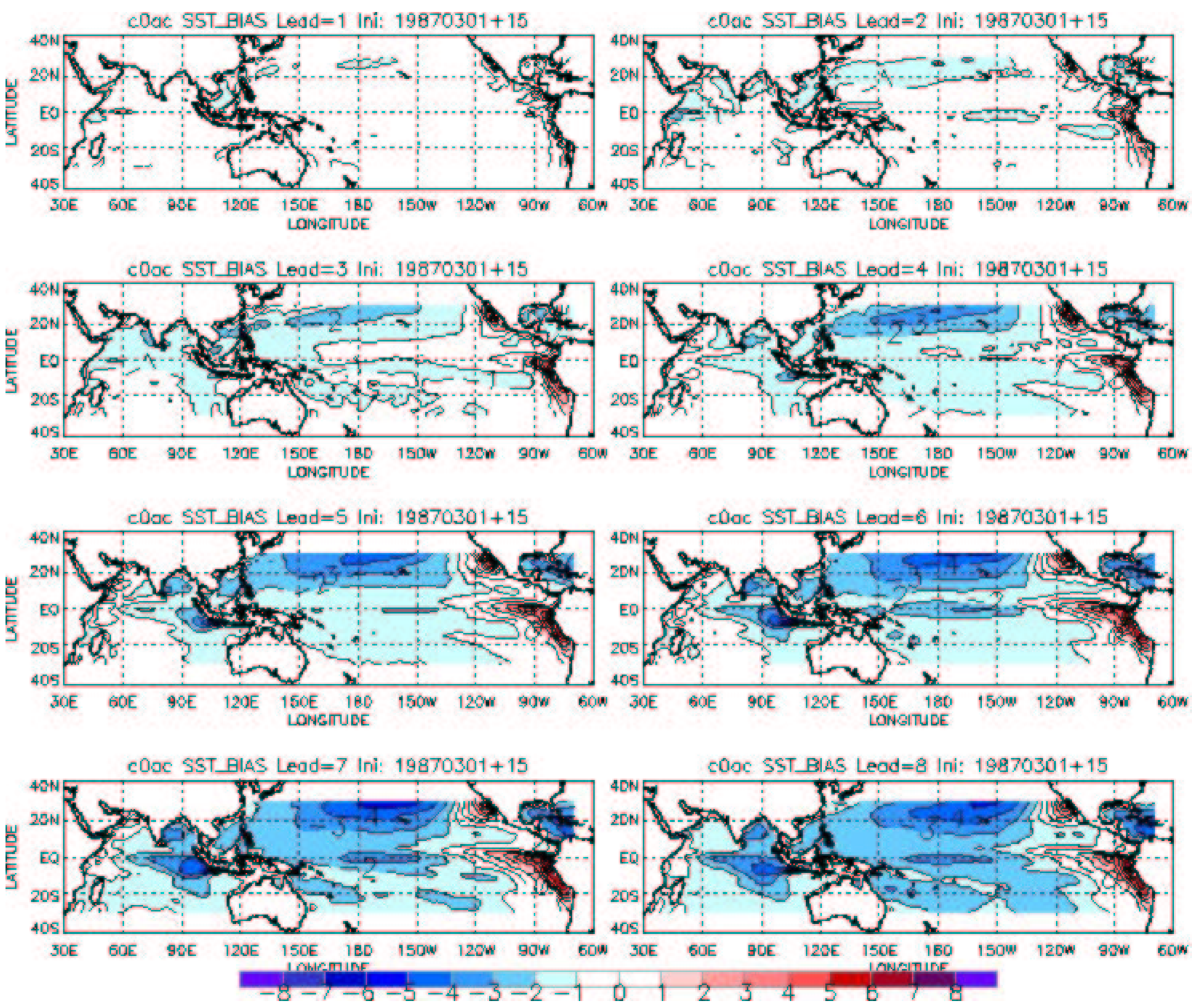

Figure 5: SST bias (POAMA SST - observed SST) from the POAMA-1 hindcasts as a function of forecast lead time (months) over the period of 1987-2001. Units are ° C.

comparing mean rainfall over south eastern Australia with the POAMA-1 predictions (Figure 8). At the lead time of 2 months south eastern Australian rainfall predicted from POAMA-1 has a weak but statistically significant correlation with the observed counterpart ($r \sim 0.16$), but POAMA-1's forecast with a 8 month lead time does not have any skill ($r \sim 0.04$). Nevertheless, skilful seasonal prediction of Australian rainfall (and regional climate) should be feasible with the current POAMA system because of its ability to predict El Niño/La Niña. Therefore, we will review commonly used statistical adjustment techniques to bridge/calibrate the predictable components of the climate system to regional climate, and we will recommend techniques that should be explored for application to the prediction of south eastern Australian rainfall.

3. Review of Bridging/Calibrating Techniques

Statistical post-processing of model forecasts can remove mean model bias, improve spatial patterns of predictable variability (calibrate/downscale), or exploit directly predictable components of the climate for prediction of associated variability (bridge). The basic approach of all statistical post-processing techniques is to develop relationships between forecasts and verification in a training period, and then to apply the statistical relationship to adjust/extend model forecasts for independent periods. Statistical post-processing suffers the same problems that pure statistical forecast-

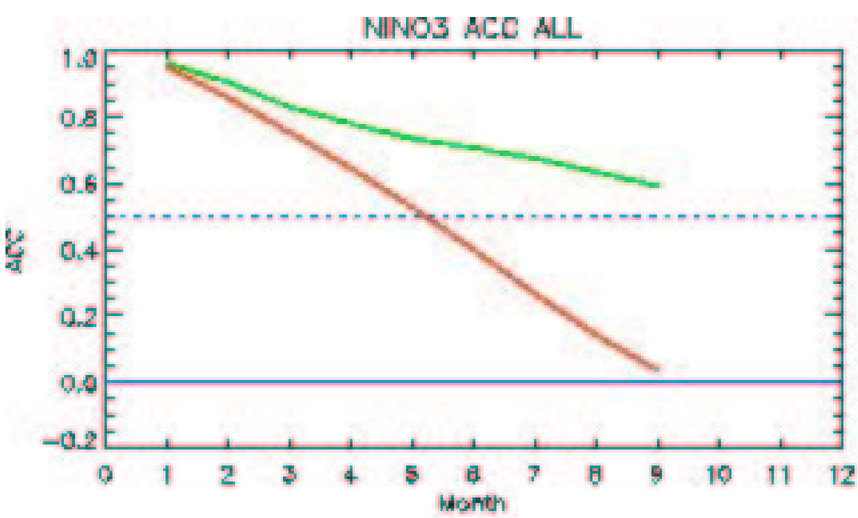

Figure 6: Anomaly correlation of observed Niño3 SST index with predicted index from the POAMA-1 hindcasts as a function of forecast lead time (green curve) in 1987-2002. For reference, correlation of persistence of observed index with itself (e.g. use this month's value of the index as a prediction for next month) is shown in red.

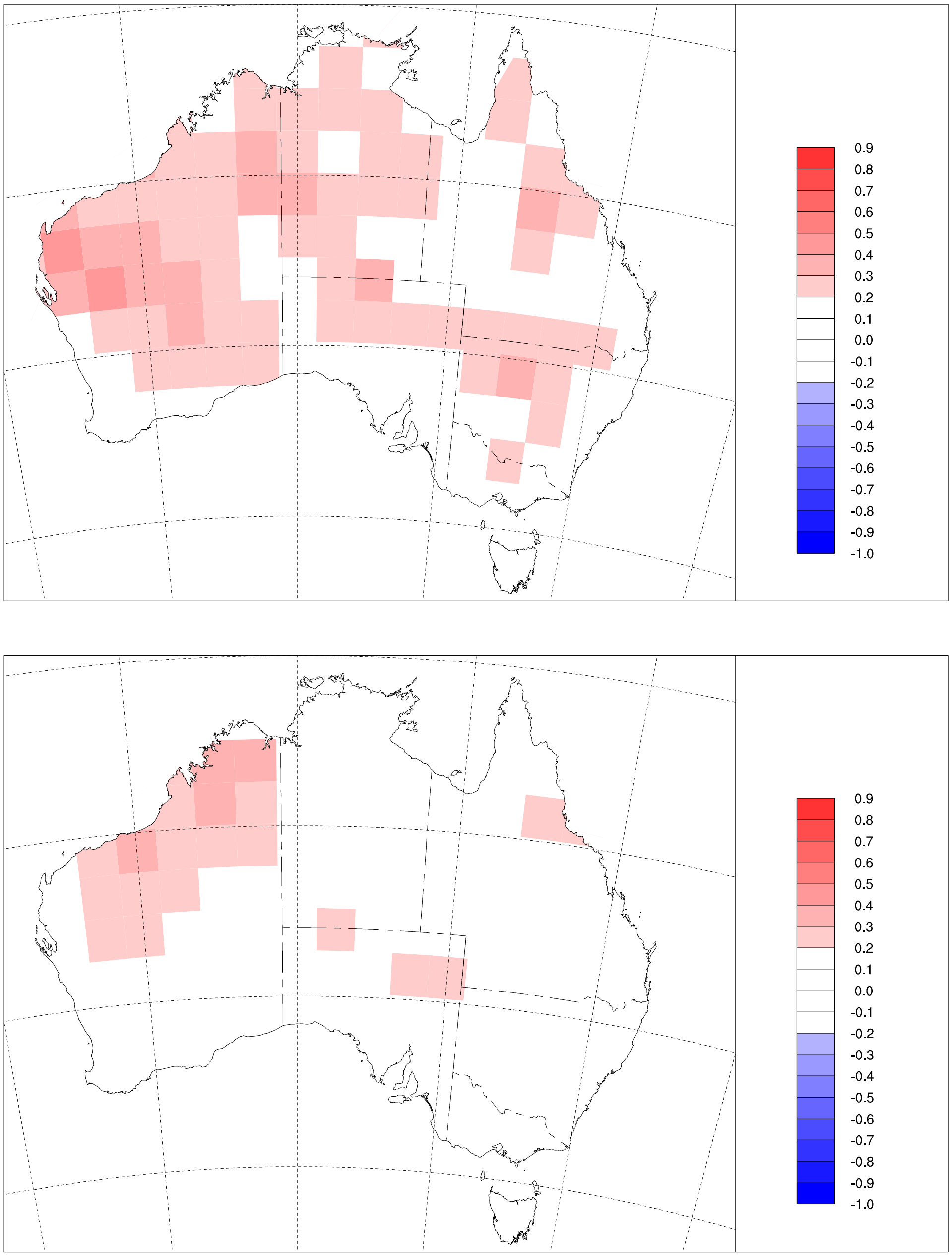

Figure 7: Correlation of observed monthly rainfall anomaly with monthly rainfall anomaly from the POAMA-1 hindcasts at the lead time of 2 months (top) and 8 months (bottom) for the period 1987-2002. The monthly rainfall anomaly is normalised by monthly mean standard deviation of rainfall.

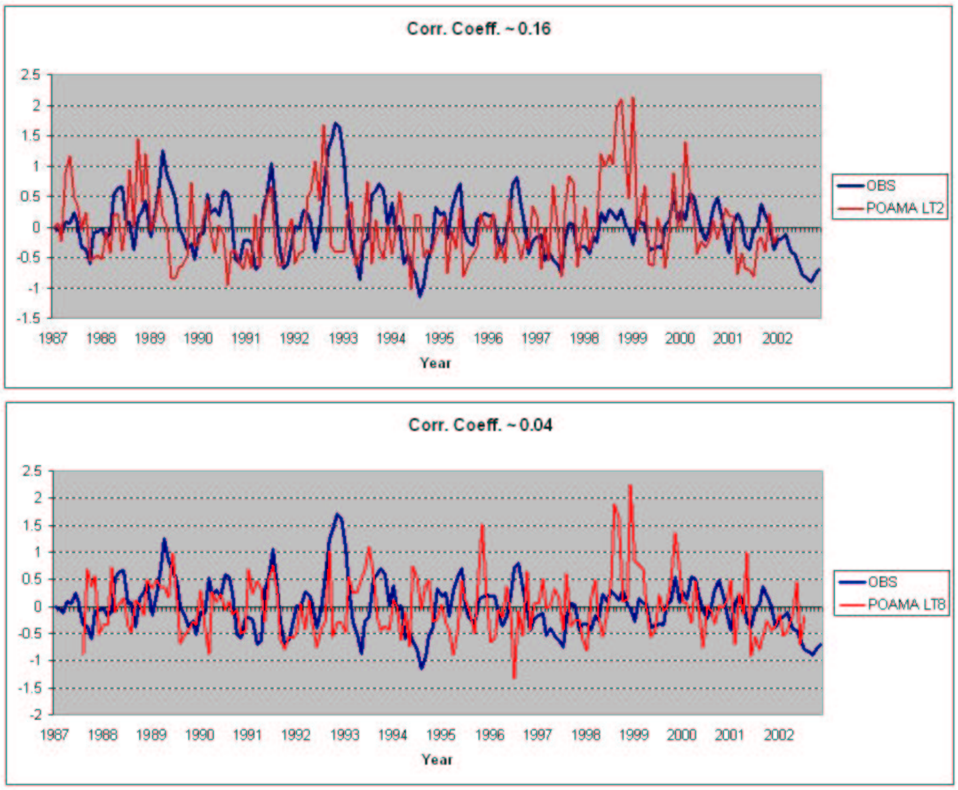

Figure 8: Time series of the normalised monthly rainfall anomalies over south eastern Australia from observation and the POAMA-1 hindcasts at the lead time of 2 months (top) and 8 months (bottom) over the period 1987-2002. Ordinate has no unit.

ing techniques do: relationships in the dependent period do not necessarily apply in independent periods, and there is the possibility of artificial skill. Also, statistical techniques are generally linear, and can not readily take into account nonlinear interactions from initial conditions and boundary forcing. Furthermore, statistical post-processing can not generate skill: the dynamical model must have skill in predicting some aspects of the climate. However, many of the observed drivers of climate variability are generally linear, and artificial skill can usually be assessed/avoided by judicious use of cross-validation. Nevertheless, the ultimate goal of dynamical seasonal prediction, which naturally accounts for nonlinearities and sensitivities to the full range of initial conditions, is to improve the dynamical forecasting system so that regional predictions can be directly used. Statistical post-processing should be viewed as an intermediate patch.

POAMA-1's ability to predict tropical Indo-Pacific SSTs with up to a 9 month lead time implies potential improvement for prediction of regional Australian climate variations associated with ENSO through statistical post-processing. We review here some possible methods to exploit this predictability of ENSO for Australian regional seasonal climate prediction.

a. Bridging

Bridging methods reported in the literature (e.g., Voldoire et al. 2002) are based on observed statistical relationships between a set of predictors (e.g., SST time series) and a set of target predictands (e.g., observed rainfall). The observed statistical relationship between predictor and predictand is generally a (multiple) linear regression relationship. There are a number of options for choosing the predictors/predictands and the associated spatial patterns. In general, the predictors must be fields for which the model has predictability. Predictands should be the fields of interest but only those that have a statistical relationship with the predictors (e.g., rainfall with SST as a predictor). The most common approaches to identifying a statistical relationship and the resultant spatial patterns of the predictors and predictands are to use singular value decomposition analysis (SVDA), canonical correlation analysis (CCA), or principal component analysis (PCA). These techniques expand predictors and predictands in terms of dominant patterns of variability and the time series of those patterns. For instance, the time series of a predictor field $\mathbf{x}_i$ (here the subscript i indicates time, and bold face indicates a spatial vector) from a training period are expanded in terms of spatial patterns ($\mathbf{g}_m$, where m is the number of patterns) and the time series (or principal components, $u_{m,i}$) of each pattern,

$$\mathbf{x}_i \approx \sum_{m=1}^{M} u_{m,i}\,\mathbf{g}_m \tag{1}$$

Likewise, the predictand time series $\mathbf{y}_i$ are expressed as

$$\mathbf{y}_i \approx \sum_{m=1}^{M} v_{m,i}\,\mathbf{h}_m \tag{2}$$

SVDA finds the spatial patterns ($\mathbf{g}_m$, $\mathbf{h}_m$) which maximize the temporal covariance between predictor and predictand (Bretherton et al. 1992, Ward and Navarra 1997). CCA looks for the spatial patterns of predictors and predictands that have maximum correlation (Barnett and Preisendorfer 1987, Bretherton et al. 1992). PCA is applied to the predictors (and sometimes also to the predictands) in order to identify the dominant patterns of variability in each field that account for the most variance (Bretherton et al. 1992, Barnett et al. 1993, Mo and Straus 2002). In this last case, the relationship between the dominant modes of predictors and predictands can be found by multiple linear regression.

According to Bretherton et al. (1992), SVDA and CCA have the advantage of providing a direct measure of association between a predictor and a predictand. However, a disadvantage of these two methods is the possibility that the covariances or correlations might artificially show a high fit to random noise (Ward and Navarra 1997). The leading modes of PCA are more reproducible in independent data (Feddersen et al. 1999). Also, as some model errors are due to noise, initial application of PCA can decrease such types of errors. However, Bretherton et al. (1992) reported that analysis based on principal components could display an undesirable mean bias towards the leading spatial patterns (which explain the most variance in one field but are not necessarily most related to another field). Despite the pros and cons of each method, Bretherton et al. (1992) and Feddersen et al. (1999) suggested that results from the three methods were similar when higher order modes of predictors and predictands were included.

Once predictors and predictands are determined and their association is found by any of these techniques, an estimate of the predictand at time i ($\hat{\mathbf{y}}_i$) is obtained by multiple linear regression of $\mathbf{y}_i$ on the time series $\mathrm{u}_{m,i}$ in (1), i.e.

$$\hat{\mathbf{y}}_i \approx \sum_{m=1}^{M_s} \mathbf{A}_m \, u_{m,i} \tag{3}$$

where M_s is the retained number of patterns, and $\mathbf{A}_m$ is a set of regression coefficients minimizing the expected root-mean-sqaure (rms) difference between $\hat{\mathbf{y}}_i$ and $\mathbf{y}_i$ from the training period. So, for instance, the $\mathrm{u}_{m,i}$ might be the principal component time series based on PCA of observed SST or of the predicted SST. The regression relation (3) then relates observed variations of the predictand (e.g., Australian rainfall at gridpoints) to these principal component time series. If (3) was developed using the principal components of observed SST and observed rainfall in a training period, then forecasts at time l, which is independent from the training period, are made by projecting predicted SST fields $\hat{\mathbf{x}}_l$ from POAMA-1 onto the respective spatial patterns from PCA of observed SST i.e.,

$$\hat{u}_{m,l} = \hat{\mathbf{x}}_l \, \mathbf{g}_m \tag{4}$$

The resulting principal components $\hat{u}_{m,l}$ are then substituted in the observed regression relation (3) to make the rainfall prediction ($\hat{\mathbf{y}}_l$). If (3) was developed using principal components of predicted SST and observed rainfall, then the predicted SST fields are projected onto the spatial patterns from PCA of predicted SST and the resultant principal components ($\hat{u}_{m,l}$) are plugged into (3) which has been developed between model SST and observed rainfall. Voldoire et al. (2002), using a similar bridging technique as described above (instead of principal component time series of the predictors, they used the Nino4 SST index), reported positive results for prediction of rainfall from an earlier version of the BMRC coupled forecast model. They showed that the model had no skill in directly predicting rainfall, but, via bridging, had comparable skill to the operational statistical scheme employed by the National Climate Center which uses the first two rotated principal components of observed SST anomaly as predictors (Drosdowsky and Chambers 2001).

b. Calibration

Calibration refers to the adjustment of spatial patterns of variability that are predicted from the forecast model against a reference data set (e.g. observational data). As in bridging, the relationship between the predicted patterns of variability and observed behavior is developed in a training period, and then this relationship is applied to forecasts from an independent period. There are various options for calibration techniques, and SVDA, CCA and PCA are also useful tools for calibration as well as for bridging. In particular, because PCA concentrates data containing a large number of variables into a small number of new variables, PCA is widely used for filtering of noise in a data set (Barnett and Preisendorfer 1987, Tippett et al. 2003). PCA also produces time series and spatial patterns that are uncorrelated from one another (Wilks 1995). Hence PCA lends itself nicely to the development of calibration based on correlation between predictors and predictands. Therefore, we assume here that a predictor (e.g. POAMA rainfall) and a predictand (e.g. observed rainfall) are expanded with PCA. Each time series of a predictor field $\mathbf{x}_i$ and a predictand field $\mathbf{y}_i$ from the training period is expanded in terms of spatial patterns ($\mathbf{g}_m$ and $\mathbf{h}_m$) and their time series ($\mathrm{u}_{m,i}$ and $\mathrm{v}_{m,i}$)

$$\mathbf{x}_i \approx \sum_{m=1}^{M} u_{m,i}\,\mathbf{g}_m \tag{5}$$

$$\mathbf{y}_i \approx \sum_{m=1}^{M} v_{m,i}\,\mathbf{h}_m \tag{6}$$

Then the predictand principal components are estimated by multiple regression onto the $\mathrm{u}_{m,i}$ time series, i.e.

$$\hat{v}_{m,i} \approx \sum_{m=1}^{M_s} C_m\,u_{m,i} \tag{7}$$

where C_m are the multiple linear regression coefficients. If the predictand time series are not expanded with spatial and temporal coefficients, then an estimate of the predictand is directly obtained by multiple linear regression on the $\mathrm{u}_{m,i}$ time series, i.e.

$$\hat{\mathbf{y}}_i \approx \sum_{m=1}^{M_s} \mathbf{B}_m\,u_{m,i} \tag{8}$$

Forecasts at time l are then either made (i) by using the predictor principal components $\mathrm{u}_{m,l}$ in (7) to obtain the calibrated principal components of the predictand ($\hat{\mathrm{v}}_{m,l}$), with which the predictand field ($\hat{\mathbf{y}}_l$) is reconstructed in (6), or (ii) by using $\mathrm{u}_{m,l}$ in (8) to directly obtain the best estimate of the predictand field $\hat{\mathbf{y}}_l$.

A common problem associated at longer lead times with statistical calibration using a linear regression is the tendency to lose variance. This can be easily fixed by multiplying the calibrated value by the ratio between the standard deviations of the observed and the adjusted values with some weighting factors - i.e. "inflation" (Kang et al. 2004, Doblas-Reyes et al. 2005).

With both bridging and calibration, if the data period is relatively short, cross-validation is required to estimate true skill. For example, to produce a forecast for time t, the data at t are removed from a complete data set, and the remaining data are used to construct a statistical model. Then, a forecast is made with this model for the data at t and verified against observation (von Storch and Navarra 1999).

c. Applications of bridging/calibrating techniques

Using statistical bridging and calibrating methods, a number of studies have reported significant forecast skill improvements. Feddersen et al. (1999) performed a single model experiment to test the prediction skill of winter precipitation in northeast Brazil and North America. Model forecasts were made with three different initial atmospheric conditions, but all were forced by identical observed SST for 34 years (1961-1994). Feddersen et al. found that, compared to the direct model output, the SVDA based adjustment improved prediction of the locations of the dominant rainfall anomalies over Brazil and reduced the amplitude error over northeast Brazil and North America.

Kang et al. (2004) investigated potential predictability of summer mean precipitation, using the KMA-SNU seasonal prediction system comprised of 10 ensemble integrations for 21 years (1979-1999). After correction with SVDA, the summer rainfall predictability was significantly enhanced over most of the domain (20°-50°N).

The above two studies utilized ensembles initialized with different initial conditions from a single model. Feddersen and Andersen (2005) and Lin and Derome (2005) used multi-model ensembles which can smooth the errors from individual model deficiencies. A good example of multi-model ensemble forecast is the one from DEMETER project (Development of a European Multi-model Ensemble system for seasonal to inTERannual prediction, Hagedorn et al. (2005)).

Feddersen and Andersen (2005) used 40 year hindcast data (1961-2000) from the DEMETER multi-model ensemble to improve forecasts of precipitation and 2-m temperature over Scandinavia, Europe, Northwest America, the U.S. and Australia. They downscaled the rainfall and temperature predicted by the DEMETER system from the model grids to regional station observations and from seasonal to daily resolution in order to make the information from the seasonal climate prediction available for crop yield models. They showed the downscaling based on SVDA improved the October-November-December rainfall prediction over Australia in terms of the mean anomaly correlation with observations. Also, the 40 year mean probability distribution of temporally downscaled rainfall exhibited a more consistent pattern with the observed counterpart than the probability distribution of the rainfall directly forecasted from the model ensemble.

Lin and Derome (2005) formed a multi-model ensemble to examine the forecast skill of the Pacific North America pattern (PNA) and the North Atlantic Oscillation (NAO) in 1969-1999. SVDA was conducted to bridge the ensemble mean Northern Hemisphere (NH) 500 geopotential height (Z500) in January and February and the observed tropical Pacific SST in the previous November. The resultant time series of the first 3 leading modes of the model Z500 were then regressed onto the time series of observed Z500, resulting in a calibrated prediction of Z500. The correlation scores between the adjusted and the observed PNA and NAO increased from 0.51 and 0.26 to 0.59 and 0.57, respectively.

In comparison, Tippett et al. (2003), Tippett et al. (2005) and Doblas-Reyes et al. (2005) attempted to bridge and calibrate their dynamical model output based on CCA. Tippett et al. (2003) sought to improve the rainfall prediction skill over the central south western Asia (CSW Asia), comparing three different predictors - CSW Asian precipitation, 200 hPa level wind over eastern Asia, and the tropical west Pacific precipitation - as simulated in a 24 member ensemble using the ECHAM 4.5 atmospheric GCM. With the tropical western Pacific rainfall as a predictor variable, the bridged and calibrated CSW Asian rainfall with CCA had a significantly improved mean anomaly correlation of 0.54 with the observed rainfall, compared to the direct rainfall prediction from the model having a correlation of 0.1.

Tippett et al. (2005) extended the work of Tippett et al. (2003) by using a multi-model ensemble with four additional models. They demonstrated that each of the five atmospheric GCMs obtained better scores in its rainfall prediction after the statistical correction using CCA based on the tropical western Pacific rainfall and the CSW Asian rainfall. The statistical adjustment enhanced the probability for below-normal rainfall for the four drought years of 1999-2002 and for above-normal rainfall over the northern part of the region in 2002. Another interesting result worth noting is that the differences in predictions from the single model ensemble and the multi-model ensemble predictions significantly decreased after each model went through the statistical calibration.

Doblas-Reyes et al. (2005) used the DEMETER multi-model ensemble with the application of CCA for the prediction of mean sea level pressure, precipitation and 2-m temperature over the tropical band (20°S- 20°N), Pacific/North America and North Atlantic-Europe in December-January-February. Their results showed that the statistical calibration could improve the accuracy of probabilistic prediction from each single model ensemble and the multi-model ensemble. Geographically, the skill enhancement was more obvious over the Tropics than for two other extratropical regions. According to their comparison of the prediction skill from the single model and the multi-model ensembles, the statistical adjustment reduced the differences in prediction skills between the single model and the multi-model ensembles, which is consistent with the finding of Tippett et al. (2005).

Barnett et al. (1993) and Chen et al. (2000) used a similar approach with PCA to correct their model systematic bias (rather than forecast error) by identifying the dominant patterns of systemtic error associated with observed variations of SST. Then this error was removed from the model based on the temporal behavior of the SST. In the study of Barnett et al. (1993) who used a dynamical ocean model coupled with a statistical atmospheric model, the error in their model's SST was predicted by the principal components of the model SST and sea level, and then the predicted error was subtracted from the model SST. The error-corrected model SST showed a higher correlation with the observed SST over the tropical Pacific Ocean.

Chen et al. (2000), using similar techniques, also obtained improved prediction skill for model SST and wind stress over the tropical Pacific Ocean. Their corrected model simulated more consistent wind and SST patterns, for instance, as occurred in the 1982-83 El Niño. It also produced a better forecast for the 1997-98 El Niño.

Finally, Mo and Straus (2002) presented a comprehensive study with respect to the application of PCA with multiple linear regression. They calibrated prediction of NH winter Z500 by regressing the leading principal components of the observed Z500 onto those of the model Z500 simulated in a GCM with 9 different initial conditions. The results indicated that the statistical correction could enhance prediction skill for Z500 in the NH winter. Also, mean square error of the forecast was significantly reduced through the adjustment when skilful predictors were chosen out of the principal components of the model Z500. In addition, they applied the same calibration technique to five other GCMs and confirmed that the statistical correction could decrease the forecasting error for the NH Z500.

In summary, the literature suggests that statistical bridging and calibration are promising ways to improve skills of seasonal prediction of regional climate, given the present limitations of a dynamical model. However, it should be borne in mind that the improvement over one region can be made at the expense of the skill over other regions (Mo and Straus 2002), and the degree of improvement can be different according to the region, season, and variable (Doblas-Reyes et al. 2005). Importantly, skill improvement in prediction of regional climate is only possible if the model can predict some relevant components of climate variability. For example, if the model has no skill in predicting SST variations associated with El Niño, no amount of statistical correction will be able to recover regional rainfall variability associated with El Niño.

4. Recommendations for application to seasonal prediction in south eastern Australia

In conclusion, previous work has reported significant improvement in regional climate prediction with bridging and calibration techniques. In regard to south eastern Australian climate prediction, explorative work to bridge and calibrate predictions from the POAMA-1 hindcasts appears to be worthwhile because there is a significant statistical relationship between south eastern Australian rainfall variability and the tropical Indo-Pacific SST variability for which POAMA-1 provides skilful prediction (Nicholls 1989, Drosdowsky and Chambers 2001, Wang and Hendon 2006). Therefore, recommended techniques to be investigated are: (i) bridging of POAMA-1's predictions of the leading modes of tropical SST variability to regional rainfall with CCA, SVDA or PCA and (ii) calibration of POAMA's predictions of rainfall using CCA, SVDA, or PCA together with linear regression onto the observed rainfall.

A major limitation of the POAMA-1 hindcasts is the short record and lack of a multi-member ensemble, which makes detection/calibration/bridging of the predictable signals difficult. Currently there is only one forecast made per month for the period 1987-2001. We strongly recommend that an extended hindcasts set be generated by a multi-member ensemble system. Work is now underway to produce an ensemble of hindcasts ($\sim$ 8-12 forecasts per month) and to extend the hindcast period 1980-2005. The generation of the ensemble hindcasts should be given the highest priority because the true benefit of calibration and bridging can not be realized with a single member ensemble.

References

Alves, O., G. Wang, A. Zhong, M. Smith, F. Tzeitkin, G. Warren, A. Schiller, S. Godfrey, and G. Meyers, 2003: POAMA: Bureau of Meteorology operational coupled model seasonal forecast system. *Science for drought - Proceedings of the National Drought Forum.* Queensland Government, Brisbane.

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.

Barnett, T. P., M. Latif, N. Graham, M. Flügel, S. Pazan, and W. White, 1993: ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean-atmosphere model. *J. Climate*, **6**, 1545–1566.

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.

Chen, D., M. A. Cane, S. E. Zebiak, R. C. nizares, and A. Kaplan, 2000: Bias correction of an ocean-atmosphere coupled model. *Geophys. Res. Lett.*, **27**, 2585–2588.

Colman, R., L. Deschamps, M. Naughton, L. Rikus, A. Sulaiman, K. Puri, G. Roff, Z. Sun, and G. Embery, 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).

Doblas-Reyes, F. J., R. Hagedorn, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting - II. Calibration and combination. *Tellus*, **57A**, 234–252.

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.

Feddersen, H. and U. Andersen, 2005: A method for statistical downscaling of seasonal ensemble predictions. *Tellus*, **57A**, 398–408.

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.

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.

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

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.

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. *Q. J. Roy. Met. Soc.*, **131**, 3301–3311.

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 Research Report No. 240.

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.

Tippett, M. K., M. Barlow, and B. Lyon, 2003: Statistical correction of central southwest Asia winter precipitation simulations. *Int. J. Climatol.*, **23**, 1421–1433.

Voldoire, A., B. Timbal, and S. Power, 2002: *Statistical-dynamical seasonal forecasting.* BMRC Research Report No. 87, Bureau of Meteorology Research Centre, 62 pp.

von Storch, H. and A. Navarra, 1999: *Analysis of climate variability.* 2nd, Updated and Extended ed., Springer, 342 pp.

Wang, G. and H. H. Hendon, 2006: Sensitivity of Australian rainfall to inter-El Niño variations. *Submitted to J. Climate*.

Wang, G., O. Alves, and N. Smith, 2005: *BAM3.0 tropical surface flux simulation and its impact on SST drift in a coupled model.* BMRC Research Report No. 107, Bureau of Meteorology Research Centre, 39 pp.

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

Wilks, D. S., 1995: *Statistical methods in the atmospheric sciences.* Academic Press, 467 pp.