

# The skill of empirical and combined/calibrated coupled multi-model South American seasonal predictions during ENSO

C. A. S. Coelho<sup>1</sup>, D. B. Stephenson<sup>1</sup>, F. J. Doblas-Reyes<sup>2</sup>, and M. Balmaseda<sup>2</sup>

<sup>1</sup>Department of Meteorology, University of Reading, Reading, UK

<sup>2</sup>European Centre for Medium-range Weather Forecasts, Reading, UK

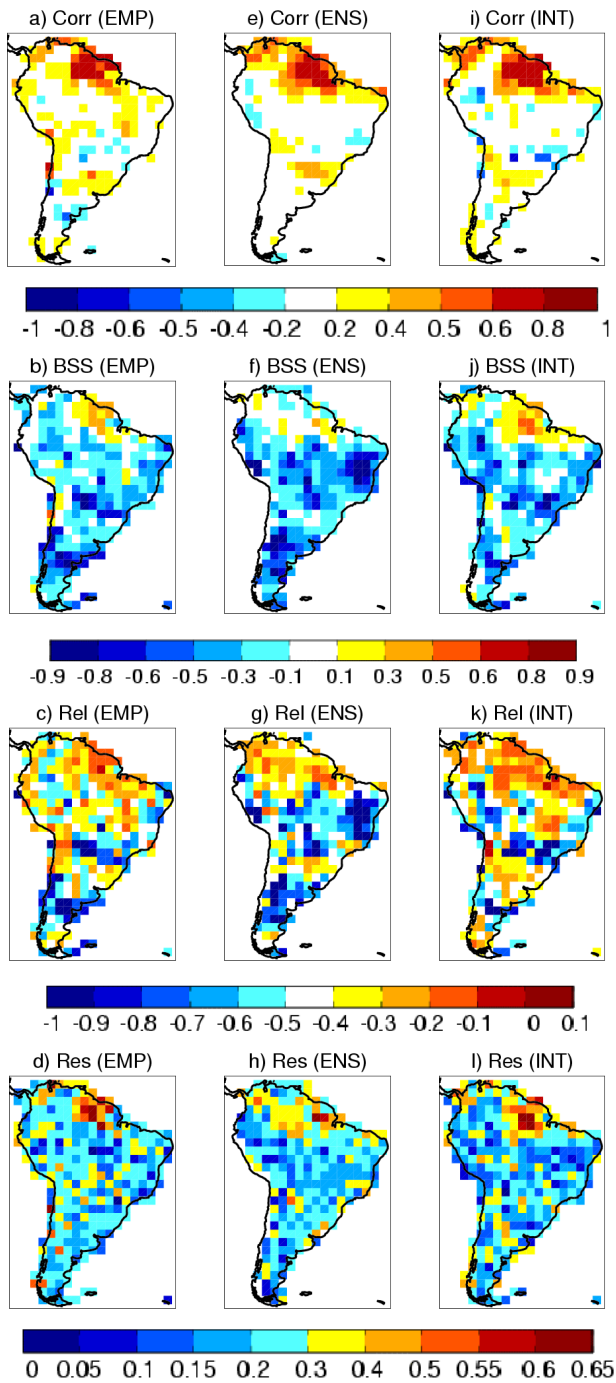
Received: 24 May 2005 – Revised: 12 December 2005 – Accepted: 15 December 2005 – Published: 9 January 2006

**Abstract.** This study addresses seasonal predictability of South American rainfall during ENSO. The skill of empirical and coupled multi-model predictions is assessed and compared. The empirical model uses the previous season August–September–October Pacific and Atlantic sea surface temperatures as predictors for December–January–February rainfall. Coupled multi-model 1-month lead December–January–February rainfall predictions were obtained from the Development of a European Multi-model Ensemble system for seasonal to inTERannual prediction (DEMETER) project. Integrated (i.e. combined and calibrated) forecasts that incorporate information provided by both the empirical and the coupled multi-model are produced using a Bayesian procedure. This procedure is referred to as forecast assimilation. The skill of the integrated forecasts is compared to the skill of empirical and coupled multi-model predictions. This comparison reveals that when seasonally forecasting December–January–February South American rainfall at 1-month lead-time the current generation of coupled models have a level of deterministic skill comparable to those obtained using simplified empirical approaches. However, Bayesian combined/calibrated forecasts provide better estimates of forecast uncertainty than the coupled multi-model. This indicates that forecast assimilation improves the quality of probabilistic predictions. The tropics and the area of South Brazil, Paraguay, Uruguay and Northern Argentina are found to be the two most predictable regions of South America. ENSO years are more predictable than neutral years, the latter having nearly null skill.

## 1 Introduction

This work has been developed as part of the Ph. D. of the first author of this paper (Coelho, 2005), under supervision of the second and third authors, in collaboration with the fourth author. The ultimate aim of this study was to produce improved integrated (i.e. combined and calibrated) probability forecasts of seasonal rainfall for South America that incorporate all available forecast information (i.e. provided by empirical and physically-derived dynamical models). Such forecasts allow local governments to plan their actions prior to the occurrence of climate anomalies such as those observed during El Niño–Southern Oscillation (ENSO) events. ENSO is one of the most important modes of climate variability affecting rainfall over South America. Therefore, the improvement of ENSO seasonal forecasts can indirectly help to improve the quality of rainfall forecasts for South America.

We compare the skill of South American austral summer (December–January–February, DJF) rainfall predictions produced by an empirical model with the skill of predictions produced by a coupled multi-model ensemble system and forecasts of an integrated system that incorporate information provided by both the empirical model and the coupled multi-model ensemble. The austral summer is when most of South America receives most of its annual rainfall. Therefore, good quality predictions for DJF are crucial for those sectors that depend on seasonal rainfall for future planning (e.g. agriculture, electricity generation).



**Fig. 1.** Maps of correlation (first row), BSS (second row) and its reliability (third row) and resolution (fourth row) components for empirical (EMP), coupled multi-model ensemble (ENS), and integrated (INT) forecasts. The BSS is for the event “rainfall anomaly less than or equal to zero”. The reliability component of the BSS is given by  $-Rel/BSc$  and the resolution component of the BSS is given by  $Res/BSc$ , where  $BSc$  is the Brier Score (Brier, 1950) of the climatological forecast, and  $Rel$  and  $Res$  are the reliability and resolution components of the Brier Score shown in panels c, g, and k and d, h and l, respectively.

## 2 Methodology

The Bayesian methodology used here, which allows the calibration and combination of coupled model with empirical predictions to produce a single integrated forecast, was developed and tested in three stages. The Bayesian method (Bayes, 1763) is a procedure for updating prior information when new information becomes available. See Coelho et al. (2003, 2004, 2005) and Stephenson et al. (2005) for further information. First, the Bayesian procedure was developed for the calibration of predictions of an ENSO index (Niño-3.4) obtained from an individual coupled model (Coelho et al., 2003, 2004). Next, the method was extended for the calibration and combination of equatorial Pacific sea surface temperature anomaly predictions from seven DEMETER coupled models (Stephenson et al., 2005). Hence, in the second stage the method dealt with multi-model predictions and acquired the first spatial (longitudinal) component. Finally, in the third stage the Bayesian multi-model method was applied to the calibration and combination of spatial field forecasts of rainfall over South America (Coelho et al., 2005). The Bayesian procedure is referred to as forecast assimilation (Coelho, 2005; Stephenson et al., 2005). Results show that integrated (INT) forecasts (i.e. combined and calibrated) obtained with forecast assimilation are better calibrated and more reliable than both empirical (EMP) and coupled multi-model ensemble (ENS) forecasts alone.

The multi-model ensemble investigated here is composed of three coupled models (ECMWF, Météo-France and Met Office), and was produced as part of the DEMETER project (<http://www.ecmwf.int/research/demeter>). Each model provides 9 members to compose the 27 multi-model ensemble. One-month lead predictions for DJF are investigated. The empirical model uses the previous season August–September–October (ASO) ERA-40 (<http://www.ecmwf.int/research/era>) sea surface temperature anomalies of the Pacific and Atlantic as predictors for DJF rainfall anomalies for the entire South American continent. Multi-model ensemble predictions are obtained by pooling all 27 ensemble members. All results shown here were obtained using the cross-validation method (Wilks, 1995). Skill assessment is performed for the period 1959–2001. Additional information about South American DEMETER hindcasts can be found at <http://www.met.reading.ac.uk/~swr01cac>.

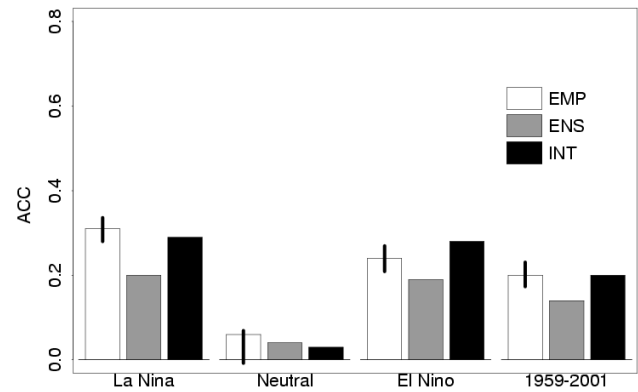
## 3 Results and discussion

Figure 1 shows correlation maps (Figs. 1a, e and i), Brier Skill Score (BSS) (Wilks, 1995) maps (Figs. 1b, f and j), as well as the reliability (Figs. 1c, 1g and 1k) and the resolution (Figs. 1d, h and l) components of the Brier Skill Score for rainfall anomaly predictions of empirical, coupled multi-model ensemble and integrated forecasts obtained with forecast assimilation for the period 1959–2001. Correlation maps show the correlation between observed and predicted anomalies at each grid point. The BSS is for the event “rainfall

anomaly less than or equal to zero". The BSS represents the level of improvement of the Brier score (Brier, 1950) compared to the Brier score of the reference/climatological forecast (BSc). The climatological probability of the event "rainfall anomaly less than or equal to zero" is about 50%, and the BSc is about 0.25. The BSS is designed to range from one for perfect predictions, through zero for predictions that provide no improvement over the reference forecast, to negative values for predictions that are worse than the reference forecast. The Guianas region in northern South America, is the most skilful with correlations between 0.6 and 0.8 and BSS between 0.1 and 0.6. The subtropics (southern Brazil, Paraguay, Uruguay and northern Argentina) also show some skill. Correlations between 0.2 and 0.5 are found in this region. These two regions are influenced by ENSO (e.g. Diaz et al., 1998 and Mestas-Nuñez and Enfield, 2001). This suggests that most of the skill of South American rainfall predictions is ENSO derived.

It is noteworthy, however, that the Guianas is not the traditionally well-known northern South American region associated with ENSO. In fact, the traditional well-known northern South American region associated with ENSO and subject to many studies is the northeast region of Brazil (e.g. Uvo et al., 1998 and Coelho et al., 2002). The effect of ENSO in the atmospheric circulation of the Guianas in DJF is to reduce convection over the region during El Niño years (Mestas-Nuñez and Enfield, 2001), whereas the reduction of convection which is observed over the northeast regions of Brazil in March–April–May is association with the warming of the tropical North Atlantic that moves the Inter Tropical Convergence Zone away from the region, favouring subsidence over the northeast of Brazil.

Empirical, coupled multi-model ensemble and integrated predictions have similar correlation maps (Figs. 1a, e and i), indicating that these three approaches have comparable level of deterministic skill. The probabilistic measure of skill (Figs. 1b and f) shows that empirical predictions are more skilful than coupled multi-model ensemble predictions, particularly in the tropical region where empirical predictions have higher BSS. Integrated forecasts obtained with forecast assimilation (Fig. 1j) have higher BSS than uncalibrated coupled multi-model ensemble predictions (Fig. 1f). This indicates that the calibration provided by forecast assimilation improves the skill of the multi-model predictions. This increase in BSS is mainly due to improvements in the reliability of the predictions (Figs. 1c, g and k), with the tropical regions also showing improvements in resolution (Figs. 1d, h and l). Integrated predictions obtained with forecast assimilation (Fig. 1j) now have a level of probabilistic skill comparable to empirical predictions (Fig. 1b). The predominance of negative BSS in Figs. 1b, f and j is due to certain properties of this score. Mason (2004) has shown that the expected value of the BSS is less than zero if nonclimatological forecast probabilities are issued. As a result, negative skill scores can often hide useful information content in the forecasts. Therefore, negative skill scores need to be interpreted with caution.



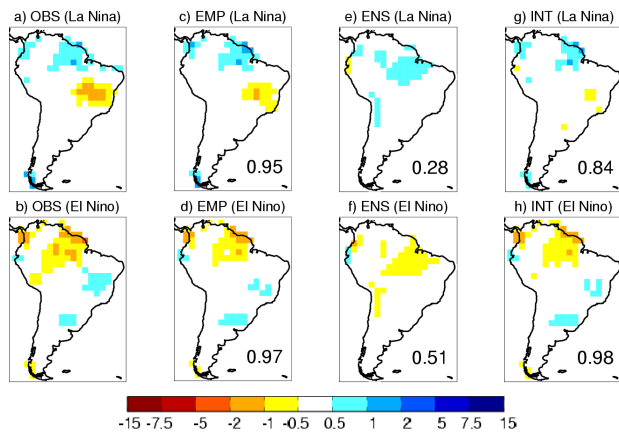
**Fig. 2.** Mean DJF anomaly correlation coefficient (ACC) for empirical (EMP), coupled multi-model ensemble (ENS) and integrated (INT) forecasts of La Niña, neutral, El Niño years (listed in Table 1) and all 1959–2001 years. The vertical solid line on the top of the white bars indicate the 95% confidence interval for the mean ACC of empirical predictions which were obtained using a bootstrap resampling procedure (Wilks, 1995, Sect. 5.3.2).

Figure 2 shows the mean anomaly correlation coefficient (ACC) (Jolliffe and Stephenson, 2003), sometimes referred to as pattern correlation, for La Niña, neutral and El Niño years occurred during 1959–2001 (Table 1) and all years. The ACC of each year is given by the correlation between the observed and predicted spatial anomaly pattern over the South American domain. The geographic conformity is generally low (c.f. ACC less than 0.31). La Niña and El Niño years have higher mean ACCs than neutral years, indicating that predictions for ENSO years are more geographically consistent than predictions for neutral years. El Niño and La Niña integrated forecasts obtained with forecast assimilation show an increase in the mean ACC compared to the uncalibrated coupled multi-model ensemble predictions. The mean ACC of El Niño and La Niña integrated forecasts are now comparable to the mean ACC of empirical predictions. This comparison is better quantified by examining the vertical solid lines on the top of the white bars of Fig. 2. These lines indicate the 95% confidence interval for the mean ACC of empirical predictions. These intervals were obtained using a bootstrap resampling procedure as described in section 5.3.2 of Wilks (1995). Mean ACCs that are within the range of values of the 95% confidence interval of the mean ACC of empirical predictions cannot be considered different from the mean ACC of empirical predictions from the statistical point of view at the 5% significance level. Figure 2 shows that empirical and integrated forecasts have comparable ACCs. Neutral years have nearly null mean ACC, indicating that rainfall anomalies of these years are virtually unpredictable.

Figure 3 shows observed and forecast austral summer rainfall anomaly composites for those La Niña and El Niño years listed in Table 1. Empirical prediction composites (Fig. 3c–d) are in good agreement with the observed composites (Fig. 3a–b). Coupled multi-model ensemble composites (Fig. 3e–f) fail to reproduce the observed anomalies in

**Table 1.** La Niña, neutral and El Niño years occurred during 1959–2001 as defined by the Climate Prediction Center (<http://www.cpc.noaa.gov>).

	Years
La Niña	1964/65, 1970/71, 1971/72, 1973/74, 1974/75, 1975/76, 1983/84, 1984/85, 1988/89, 1995/96, 1998/99, 1999/00, 2000/01
Neutral	1959/60, 1960/61, 1961/62, 1962/63, 1966/67, 1967/68, 1978/79, 1980/81, 1981/82, 1985/86, 1989/90, 1993/94, 1996/97, 2001/02
El Niño	1963/64, 1965/66, 1968/69, 1969/70, 1972/73, 1976/77, 1977/78, 1979/80, 1982/83, 1986/87, 1987/88, 1990/91, 1991/92, 1992/93, 1994/95, 1997/98

**Fig. 3.** DJF rainfall anomaly composites (mm/day) for those La Niña and El Niño years listed in Table 1. Observed (first column), empirical (second column), coupled multi-model ensemble (third column) and integrated (fourth column) forecast composites. Correlation between observed and forecast composites appears on the right hand corner of the forecast composites panels.

eastern Brazil, Patagonia (in southern South America) and the Plate river region (in southeastern South America). Besides the Guianas anomalies in northern South America are shifted to the southeast and unobserved Andean anomalies are introduced. Integrated forecast composites (Fig. 3g–h), obtained when multi-model ensemble predictions were calibrated and combined with empirical predictions, are now comparable to empirical composites (Fig. 3c–d). This is because, in addition to incorporating well-calibrated empirical predictions, the forecast assimilation procedure used to generate integrated forecasts also corrects biases in the coupled models predictions and shifts multi-model ensemble prediction patterns towards the observed patterns.

#### 4 Conclusions

This study addresses seasonal predictability of South American rainfall. The skill of empirical, a coupled multi-model ensemble and integrated forecasts obtained with forecast assimilation has been assessed and compared. This comparison revealed that when seasonally forecasting DJF South American rainfall at 1-month lead-time the current generation of coupled models have comparable level of deterministic skill

to those obtained using a simplified empirical approach, as illustrated by the correlation maps of Fig. 1. This result is in agreement with findings of previous comparison studies.

Bayesian forecast assimilation has been shown to be a powerful tool for the calibration and combination of multi-model ensemble and empirical predictions. Integrated predictions obtained with forecast assimilation have been shown to have improved BSS compared to the simple multi-model prediction. This is because forecast assimilation provides better estimates of forecast uncertainty than coupled multi-model. Additionally, forecast assimilation ENSO composites have been shown to be in much better agreement with observed composites than multi-model ensemble composites.

The tropics and the area of south Brazil, Paraguay, Uruguay and northern Argentina have been found to be the two most predictable regions of South America. South American rainfall is generally only predictable in ENSO years rather than in neutral years, which exhibit very little skill.

*Acknowledgements.* CASC was sponsored by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) process 200826/00-0. FJDR was supported by DEMETER (contract EVK2-1999-00024).

Edited by: P. Fabian and J. L. Santos

Reviewed by: D. Enfield and J. L. Santos

#### References

- Bayes, T.: An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society of London*, 53, 370–418, 1763.
- Brier, G. W.: Verification of forecasts expressed in terms of probability, *Mon. Wea. Rev.*, 78, 1–3, 1950.
- Coelho, C. A. S., Uvo, C. B., and Ambrizzi, T.: Exploring the impacts of the tropical Pacific SST on the precipitation patterns over South America during ENSO periods, *Theoretical and Applied Climatology*, 71, 185–197, 2002.
- Coelho, C. A. S.: Forecast calibration and combination: Bayesian assimilation of seasonal climate predictions, PhD thesis, University of Reading, 178pp, 2005.
- Coelho, C. A. S., Stephenson, D. B., Balmaseda, M., Doblas-Reyes, F. J., and von Oldenborgh, G. J.: Towards an integrated seasonal forecasting system for South America, *J. Clim.*, accepted, also published as a ECMWF Technical Memorandum, No. 461, 26pp., 2005.

- Coelho, C. A. S., Pezzulli, S., Balmaseda, M., Doblas-Reyes, F. J., and Stephenson, D. B.: Skill of coupled model seasonal forecasts: A Bayesian assessment of ECMWF ENSO forecasts, ECMWF Tech. Memo., 426, 16pp., 2003.
- Coelho, C. A. S., Pezzulli, S., Balmaseda, M., Doblas-Reyes, F. J., and Stephenson, D. B.: Forecast calibration and combination: A simple Bayesian approach for ENSO, *J. Clim.*, 17, 1504–1516, 2004.
- Diaz, A. F., Studzinski, C. D., and Mechoso, C. R.: Relationships between precipitation anomalies in Uruguay and southern Brazil and sea surface temperature in the Pacific and Atlantic oceans, *J. Clim.*, 11, 251–271, 1998.
- Jolliffe, I. N. and Stephenson, D. B.: Forecast verification: A practitioner's guide in atmospheric science, Wiley and Sons, First edition, 240pp., 2003.
- Masson, S. J.: On using “Climatology” as a reference strategy in the Brier and ranked probability skill scores, *Mon. Wea. Rev.*, 132, 1891–1895, 2004.
- Mestas-Núñez, A. M. and Enfield, D. B.: Eastern equatorial Pacific SST variability: ENSO and non-ENSO components and their climatic associations, *J. Clim.*, 14, 391–402, 2001.
- Stephenson, D. B., Coelho, C. A. S., Doblas-Reyes, F. J., and Balmaseda, M.: Forecast Assimilation: A unified framework for the combination of multi-model weather and climate predictions, *Tellus*, 57, 253–264, 2005.
- Uvo, C. R. B., Repelli, C. A., Zebiak, S. E., and Kushnir, Y.: On the relationship between tropical Pacific and Atlantic SST in northeast Brazil monthly precipitation, *J. Clim.*, 13, 287–293, 1998.
- Wilks, D. S.: Statistical methods in atmospheric sciences: An introduction, 1st Edition, Academic Press, 467pp., 1995.