From multi-model ensemble predictions to well-calibrated probability forecasts: Seasonal rainfall forecasts over South America 1959-2001

Caio A. S. Coelho*, David B. Stephenson*, Francisco J. Doblas-Reyes**, and Magdalena Balmaseda**.
*Department of Meteorology, University of Reading – Reading, U.K.
**ECMWF – Reading, U.K.
Corresponding author: c.a.d.s.coelho@reading.ac.uk

INTRODUCTION

South American rainfall seasonal forecasts are currently produced using either physically derived numerical climate models or using empirical (statistical) relationships based on historical data. Only a few studies have compared the skill of these two approaches for some regions of South America, indicating that more comparison studies are required. This paper aims to compare the skill of an empirical model with coupled multi-model December-January-February (DJF) South American rainfall anomaly predictions. The austral summer season is when most 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).

METHODOLOGY

The multi-model ensemble investigated here is composed by three coupled models (ECMWF, Meteo-France and UKMO), and was produced as part of the DEMETER[1] project. 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[2] sea surface temperature anomalies of the Pacific and Atlantic as predictors for DJF rainfall anomalies (Chen *et al.*, 2002) for the entire South American continent. The skill of the empirical model is compared to the skill of both the simple multi-model ensemble, obtained by pooling all 27 ensemble members, and the skill of combined/calibrated forecasts obtained using the Bayesian forecast assimilation (FA) procedure as described in Stephenson *et al.* (2004). Coelho *et al.* (2004) and Coelho *et al.* (2003) provide additional information about the Bayesian method of calibration/combination of forecast. All results shown here were obtained using the cross-validation method (Wilks 1995). Skill assessment is performed for the period 1959-2001, which is the common period of hindcasts produced by the three DEMETER coupled models. Additional information about South American DEMETER hindcasts can be found at http://www.met.reading.ac.uk/~swr01cac

RESULTS AND DISCUSSION

Figure 1 shows correlation maps (Figs. 1a-c) and Brier Skill Score (Wilks, 1995) maps (Figs. 1d-f) of rainfall anomaly predictions for empirical, multi-model and forecast assimilation (FA) for the period 1959-2001. These 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 that of a reference forecast (in this case climatological probability of the event). 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 tropical region, in Northern South America, is the most skilful region with correlations between 0.6 and 0.8 and BSS between 0.1 and 0.6. The subtropics (south 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 well known to be influenced by the El Niño Southern Oscillation (ENSO). This suggests that most of the skill of South American rainfall predictions is ENSO derived. Empirical, multi-model and forecast assimilation predictions have similar correlation maps (Figs. 1a-c), indicating that these three approaches have comparable level of deterministic skill. The probabilistic measure of skill (Figs. 1d and 1e) shows that empirical predictions are more skilful than multi-model predictions, particularly in the tropical region where empirical predictions have higher BSS. Bayesian combined/calibrated predictions obtained with forecast assimilation (Fig. 1f) have higher BSS than uncalibrated multi-model predictions (Fig. 1e). 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, with the tropical regions also showing improvements in resolution (not shown). Combined/calibrated predictions obtained with forecast assimilation (Fig. 1f) have now comparable level of probabilistic skill as

[1] http://www.ecmwf.int/research/demeter
[2] http://www.ecmwf.int/research/era

empirical predictions (Fig. 1d). The predominance of negative BSS in Figs. 1d-f is due to some 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.

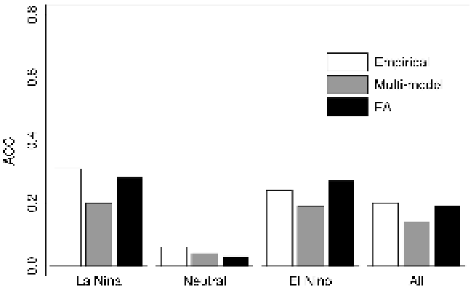

Figure2: Mean ACC for empirical, multi-model and forecast assimilation (FA) predictions of La Niña, neutral and El Niño years of Table 1 and all (1959-2001) years.

Figure 2 shows the mean anomaly correlation coefficient (ACC) 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. La Niña and El Niño years have higher mean ACC than neutral years, indicating that predictions for ENSO years are more skilful than predictions for neutral years. El Niño and La Niña predictions obtained with forecast assimilation show an increase in the mean ACC compared to the uncalibrated multi-model. The mean ACC of El Niño and La Niña forecast assimilation predictions are now comparable to the mean ACC of empirical predictions. Neutral years have nearly null mean ACC, indicating that rainfall anomalies of these years are hardly predicted.

	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

Table 1: La Niña, neutral and El Niño years occurred during 1959-2001.

Figure 3 shows observed and predicted composites of DJF rainfall anomalies for those La Niña and El Niño years listed in Table 1. Figures 3b and 3f show that the empirical model reproduces remarkably well both La Niña and El Niño observed composites patterns (Figs. 3a and 3e). The correlation between the predicted and observed patterns is 0.95 for the La Niña composite and 0.97 for the El Niño composite. The multi-model composites (Figs. 3c and 3g) partially reproduce the observed pattern in equatorial South America and fail to reproduce the observed pattern in the other regions of the continent. The correlation between the multi-model composite and the observed composite is 0.28 for La Niña and 0.51 for El Niño. Bayesian combined/calibrated composites produced with forecast assimilation (Figs. 3d and 3h) are in much better agreement with the observations. The correlation between FA composites and observed composites is 0.82 for La Niña and 0.97 for El Niño, being now comparable to the correlation values of empirical composites. These results suggest that additional skill can be gained by calibrating multi-model predictions with forecast assimilation.

CONCLUSIONS

This study addressed seasonal predictability of South American rainfall. The skill of empirical, DEMETER coupled multi-model and combined/calibrated predictions obtained with forecast assimilation has been assessed and compared. This comparison revealed that when seasonally forecasting Dec-Jan-Feb South American rainfall at 1-month lead-time the current generation of coupled models have comparable level of skill to those obtained using a simplified empirical approach. The same conclusion still holds for longer (e.g. 3-month) lead times. 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 of multi-model predictions. Forecast assimilation predictions 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. Forecast assimilation predictions (and FA alone) produce probability forecasts with skill in Southeastern South America – an important region for hydroelectricity production. Additionally, forecast assimilation ENSO composites have been shown to be in much better agreement with observed composites than multi-model 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.

ACKNOWLEDGMENTS

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).

REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1-3.

Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin, 2002: Global land precipiation: A 50-yr monthly analysis based on gauge observations, 2002, *J. of Hydrometeorology*, **3**, 249-266.

Coelho, C. A. S., S. Pezzulli, M. Balmaseda, F. J. Doblas-Reyes and D. B. Stephenson, 2003: Skill of coupled model seasonal forecasts: A Bayesian assessment of ECMWF ENSO forecasts. *ECMWF Tech. Memo. 426*.

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

Stephenson D. B., C.A.S. Coelho, F. J. Doblas-Reyes and M. Balmaseda, 2004: Forecast Assimilation: A unified framework for the combination of multi-model weather and climate predictions. *Tellus A – DEMETER special issue (in press).*

Wilks, D. S., 1995: Statistical methods in atmospheric sciences: An introduction. 1st Edition. Academic Press. 467pp.

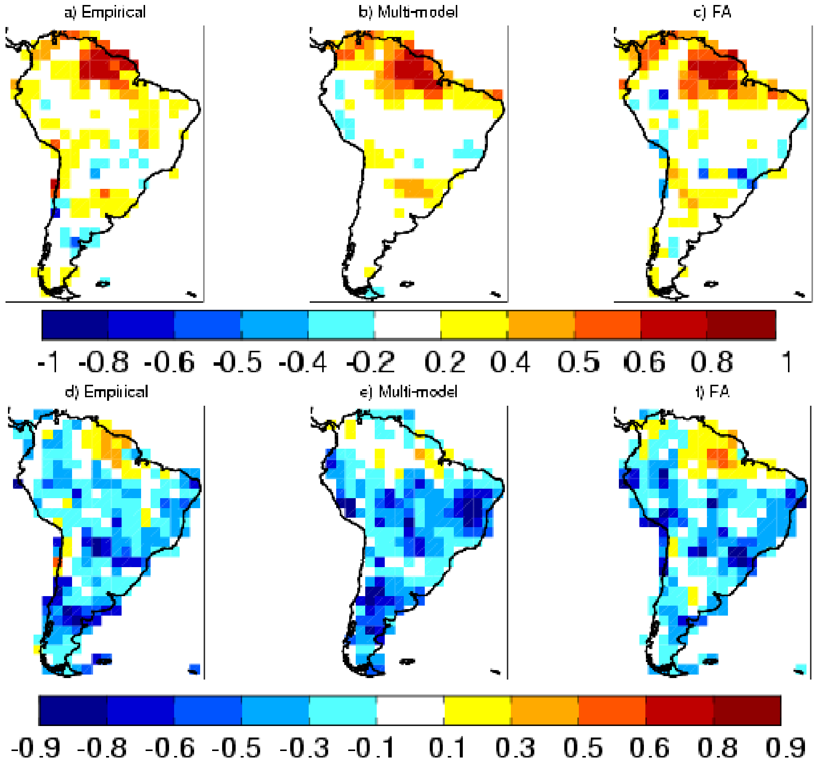

Figure 1: Correlation maps (panels a-c) and Brier Skill Score maps (panels d-f) of empirical, multi-model and forecast assimilation (FA) DJF predictions for the period 1959-2001.

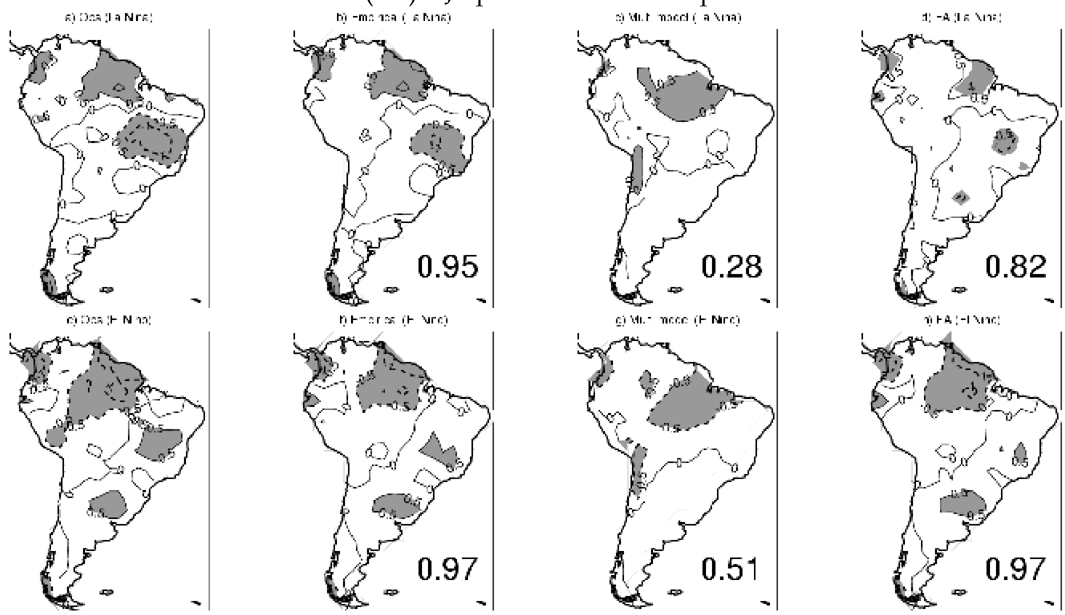

Figure 3: Observed (panels a and e) and predicted (panels b-d and f-h) DJF rainfall anomaly composites of La Niña and El Niño years (in mm.day^{-1}) by empirical, multi-model and forecast assimilation (FA). The number in the bottom right hand corner of panels b-d) and f-h) is the correlation between the observed composite and the predicted composite. Contour interval is 0.5 mm.day^{-1}. Rainfall anomalies below -0.5 mm.day^{-1} and above 0.5 mm.day^{-1} are grey shaded.