

Towards an integrated seasonal forecasting system for South America

C. A. S. Coelho¹ and D. B. Stephenson

Department of Meteorology, University of Reading, Reading, United Kingdom

M. Balmaseda and F. J. Doblas-Reyes

*European Centre for Medium-Range Weather Forecasts, Reading,
United Kingdom*

G. J. van Oldenborgh

Royal Dutch Meteorological Institute, De Bilt, Netherlands

Submitted to the Journal of Climate

May 2, 2005

¹C. A. S. Coelho, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, U. K., E-mail: c.a.d.s.coelho@reading.ac.uk

Abstract

This study proposes an integrated seasonal forecasting system for producing improved and well-calibrated probabilistic rainfall forecasts for South America. The proposed system has two components: a) an empirical model that uses Pacific and Atlantic sea surface temperature anomalies as predictor for South American rainfall; and b) a multi-model system composed of three European coupled ocean-atmosphere models. Three-month lead austral summer rainfall predictions produced by the components of the system are integrated (i.e. combined and calibrated) using a Bayesian forecast assimilation procedure. The objective calibration and combination of empirical and multi-model coupled predictions makes this a first step towards an integrated forecasting system for issuing South American seasonal forecasts. The skill of empirical, coupled multi-model and integrated forecasts obtained with forecast assimilation is assessed and compared. This comparison reveals that the simple multi-model ensemble of the current generation of coupled models has comparable level of skill to that obtained using a simplified empirical approach. As for most regions of the globe, seasonal forecast skill for South America is low. However, when empirical and coupled multi-model predictions are combined and calibrated using forecast assimilation, more skillful integrated forecasts are

obtained than with either empirical or coupled multi-model predictions alone. Both the reliability and resolution of the probability forecasts have been improved by forecast assimilation in several regions of South America. The tropics and the area of southern Brazil, Uruguay, Paraguay and northern Argentina have been found to be the two most predictable regions of South America during the austral summer. Skillful South American rainfall forecasts are generally only possible during El Niño or La Niña years rather than in neutral years.

1. Introduction

South American seasonal forecasts are currently produced with either empirical (statistical) or physically-derived dynamical models (see brief literature review in section 2). It is noteworthy, however, that seasonal forecasts provided by several different sources may be of no practical use for end-users. This is because end-users are usually interested in a single well-calibrated forecast of the event of their interest. Providing a finite set of forecasts from different models will not always help the decision making process of the end-user. Conversely, a single well-calibrated probability forecast that incorporates all available model prediction information may be more beneficial for end-users. The need of an objective method for combining different pieces of available forecast information in South America has recently been recognised by Berri and Antico (2005). This paper addresses and proposes a solution for this problem by objectively producing a single integrated (i.e. combined and calibrated) forecast of seasonal rainfall for South America that gathers prediction information from four different sources (three coupled ocean-atmosphere models and an empirical model).

Good quality seasonal forecasts are fundamental for local governments to plan their actions in order to minimize human and economical losses that may be caused by anomalous climate events such as those observed during El Niño-Southern Oscillation (ENSO) episodes. In South America these forecasts are useful for civil defence, agricultural, fishery and water resources (reservoir man-

agement) planning. Brazil, the largest and most populated country of South America, produces more than 90% of its electricity from hydropower stations², emphasising the need for good quality seasonal rainfall forecasts. The provision of improved seasonal rainfall forecasts could help the Brazilian government to better plan its management actions in order to have a more efficient control of its national electricity production program.

Figure 1a shows the total annual mean rainfall for South America. Figure 1b shows the mean rainfall in austral summer defined here as November-December-January (NDJ). These figures show that a large amount of the total annual rainfall falls during the austral summer, which defines the wet season for most South America. Tropical and southern South America have the largest annual and austral summer rainfall variability (Fig. 1c-d). Although good quality dry season forecasts may be as important as wet season forecasts for some sectors, wet season forecasts are of most relevance for electricity generation. This study focusses on predictions of November-December-January total rainfall for South America produced with initial conditions of the first day of the preceding August (i.e. 3-month lead). These are the longest lead austral summer predictions that were available for investigation from the European Union funded project entitled Development of a European Multi-model Ensemble system for seasonal to inTERannual prediction (DEMETER³) (Palmer *et al.* 2004). Therefore, all results shown in this paper are 3-month lead forecasts

²For more information see <http://www.ons.org.br>

³For more information about this project refer to <http://www.ecmwf.int/research/demeter>

for November-December-January. Three-month lead traditional austral summer December-January-February forecasts could not be examined because DEMETER simulations have only been produced four times each year, starting on the first day of February, May, August and November.

[Figure 1 about here.]

South America is a region with strong atmospheric teleconnections linked to ENSO (Wallace and Gutzler 1981; Trenberth *et al.* 1998). Because of these teleconnections, there is promising skill in seasonal forecasts for some regions of South America. Figure 2 shows La Niña and El Niño composites of austral summer South American rainfall. These figures illustrate regions strongly affected during ENSO events (shaded areas). During La Niña years, positive anomalies are observed in northern South America and negative anomalies are observed in southern/southeastern South America. This pattern is reversed during El Niño years. Seasonal forecasts in these regions have some predictive skill, as will be discussed in more detail later.

[Figure 2 about here.]

[Table 1 about here.]

Climate model predictions usually produced at coarse $2.5^{\circ} \times 2.5^{\circ}$ resolution are not able to (and are not expected to) simulate the observed climate perfectly.

This problem is further aggravated by the lack of comprehensive observational datasets to initialise the models appropriately and the lack of a complete physical understanding of the climate system. All these limitations contribute to uncertainties in climate predictions and therefore calibration against past observations is required. To our knowledge, no studies have been published with the aim of improving the quality of South America physically-derived climate model seasonal predictions by statistical calibration based on past observations. Most previous studies (e.g. Cavalcanti *et al.* 2002; Marengo *et al.* 2003; Moura and Hastenrath 2004) investigated the ability of atmospheric general circulation models forced with observed sea surface temperature in simulating climatological features such as the annual and seasonal cycles of rainfall for some regions of South America. These studies have identified systematic forecast errors, yet have not suggested approaches for correcting these errors so as to improve the forecasts. The errors arise from a combination of factors such as the chaotic evolution of the atmosphere, errors in the initial conditions of the model, and errors in model formulation/parameterisation. This study uses a Bayesian approach for statistical calibration of coupled model South American rainfall seasonal predictions.

We present a new integrated rainfall seasonal forecasting system for South America that consists of an empirical model and a coupled multi-model ensemble prediction system. Integrated forecasts are produced using a probabilistic Bayesian forecast assimilation procedure (Coelho 2005; Stephenson *et*

al. 2005). Appendix A provides more information about Bayesian forecast assimilation. This procedure allows the production of well-calibrated reliable probability estimates of rainfall by statistically correcting model errors based on past history of predictions and observations. It also allows the combination of forecasts produced by different sources, and is used here for combining both empirical and coupled multi-model predictions. The resulting combined and calibrated forecasts are summarized by the mean and the variance of a normal (Gaussian) distribution at each grid point. Seasonal averages of rainfall over several regions of South America are found to be close to following a normal distribution (see Fig. 3).

A particular aim of this study is to answer the question of whether or not forecasts produced by the proposed integrated system are better than those produced by a simple empirical model or by the simple multi-model ensemble alone. In order to address these questions, the skill of predictions produced by each individual component of the proposed integrated system is compared to the skill of the integrated forecasts.

Section 2 reviews the current state of South American rainfall seasonal forecasting and introduces the idea of combining all available prediction information in order to produce the best possible estimate of the future likely climate conditions. Section 3 briefly introduces and describes the two components of the proposed integrated system. Section 4 describes how predictions are objectively combined and calibrated in the proposed integrated system. The skill of

combined and calibrated forecasts of the integrated system as well as empirical and multi-model ensemble predictions alone are assessed and compared in section 5. Finally, section 6 concludes the paper with a summary of the major findings and a discussion of possible future areas of research.

2. Review of South American rainfall seasonal forecasting

Several studies have used atmospheric general circulation models forced with observed sea surface temperatures to simulate seasonal rainfall over South America (e.g. Folland *et al.* 2001; Cavalcanti *et al.* 2002; Marengo *et al.* 2003; Moura and Hastenrath 2004). These studies have demonstrated that atmospheric models have some predictive skill when forecasting rainfall in the tropical region of South America and over the southern part of Brazil, Uruguay, Paraguay and northeastern Argentina. All other areas of South America showed poor predictive skill. They all found that forecast skill is highly conditioned on the presence of ENSO events, with neutral years having less predictive skill. Both tropical South America and the southern region of Brazil, Uruguay, Paraguay and northeastern Argentina have strong ENSO signals (Fig. 2).

Studies by Pezzi *et al.* (2000), Folland *et al.* (2001), Greischar and Hastenrath (2000) and Martis *et al.* (2002) have developed empirical models that

relate observed rainfall to sea surface temperature over the Atlantic and Pacific oceans as well as the meridional surface wind component over the tropical Atlantic. These models have been used to predict seasonal rainfall over the south and northeast regions of Brazil and the Netherlands Antilles. Empirical models have been primarily developed for these regions because of the higher predictability of these regions compared to the other areas of South America (Cavalcanti *et al.* 2002; Marengo *et al.* 2003). Empirically based rainfall predictions for the northeast region of Brazil are skillful during the period March-April-May, which is the rainy season for most parts of this region (Greischar and Hastenrath 2000; Folland *et al.* 2001; Moura and Hastenrath 2004). The empirical predictions of Pezzi *et al.* (2000) for the south of Brazil are generally less skillful than the predictions for the northeast region of Brazil, and El Niño years were found to be more predictable than neutral and La Niña years. It is worth noticing, however, that the majority of these studies produced deterministic forecasts. Very little effort has been put into producing probabilistic rainfall seasonal forecasts for South America, emphasizing that this is an area of research that deserves further attention. This paper produces probabilistic rainfall seasonal forecasts for South America.

The comparative skill of physically-derived dynamical and empirically based seasonal predictions of South American rainfall has not been fully explored, and further systematic comparisons are desirable (Moura and Hastenrath 2004). Only a few comparison studies (Folland *et al.* 2001; van Oldenborgh *et al.*

2005; and Moura and Hastenrath 2004), focussing on rainfall forecasts for South America have been carried out. Van Oldenborgh *et al.* (2005) concluded that physically-derived dynamical predictions slightly outperform empirical predictions over tropical South America, northeast Brazil and Uruguay in December-January-February. Folland *et al.* (2001) and Moura and Hastenrath (2004) focussed on rainfall forecasts for the northeast of Brazil and concluded that physically-derived dynamical predictions do not outperform empirically based predictions. Their conclusions are in accordance with other comparative skill assessment studies for other target regions outside South America (e.g. Barnston *et al.* 1999; Anderson *et al.* 1999). Section 5 of this paper will contribute to this skill comparison exercise. We compare the skill of an empirical model that uses observed Pacific and Atlantic sea surface temperature anomalies to predict South American austral summer rainfall anomalies, with the skill of 3-month lead austral summer rainfall anomalies predictions produced by a coupled multi-model ensemble for the period 1959-2001. The empirical model developed in this study differs from those of previous studies (Greischar and Hastenrath 2000; Pezzi *et al.* 2000; Folland *et al.* 2001; Moura and Hastenrath 2004) in that it predicts rainfall anomalies for the entire South American continent, while the previous studies focussed only on specific sub-regions.

Combining the predictions from these two approaches can help yield better forecasts of future climate. The International Research Institute for Climate Prediction (IRI) subjectively combines physically-derived dynamical and em-

pirically based seasonal predictions for several continental regions (Barnston *et al.* 2003). Objective combination has still not been implemented operationally by any seasonal climate prediction centre. This study demonstrates an objective forecast assimilation scheme for combining physically-derived coupled model and empirically based predictions of South American rainfall anomalies. The skill of integrated forecasts obtained with forecast assimilation is shown to exceed the skill of each individual prediction approach.

3. Empirical and coupled model predictions

a. Empirical prediction model

South America is bordered by the Pacific and Atlantic oceans. Surface conditions of these two oceans are potential sources of predictability for South American climate (Moura and Shukla 1981; Mechoso *et al.* 1990; Marengo 1992; Nobre and Shukla 1996; Diaz *et al.* 1998; Uvo *et al.* 1998; Barros and Silvestri 2002; Coelho *et al.* 2002; Peagle and Mo 2002 among others). These studies have identified regions of South America where rainfall is sensitive to sea surface temperature anomalies in the Pacific and Atlantic oceans. This section exploits these relationships by using a lagged regression empirical model based on maximum covariance analysis (MCA) (von Storch and Zwiers 1999) – sometimes referred to as singular value decomposition (SVD). This empirical

model uses May-June-July (MJJ) Pacific and Atlantic sea surface temperature anomalies (140°E-10°E; 15°N-60°S) as predictors for South American rainfall anomalies of the following November-December-January during 1959-2001. The previous season May-June-July is used for consistency with the initial conditions of the first day of August that are used by the three DEMETER coupled models here investigated to predict November-December-January rainfall (see section 3b). In this way both empirical and coupled models use initial conditions observed up until the last day of July. Sea surface temperature anomalies were obtained from ERA-40 reanalysis (ECMWF 40 years reanalysis project, Uppala *et al.* 2005). ERA-40 provides global analysis of variables for the atmosphere and land for the period 1958-2001. More information is available at <http://www.ecmwf.int/research/era/>. Precipitation anomalies were obtained from the 50-year 1950-2001 global monthly precipitation reconstruction over land (PREC/L) version 1.0 dataset⁴ (Chen *et al* 2002), which is based on gauge observations. These datasets have been chosen because they are among the most complete with the longest records available for climate research.

A simple way to predict a South American rainfall anomaly y uses multivariate linear regression on the preceding sea surface temperature anomaly z :

$$y = M(z - z_0) + \epsilon_T \quad (1)$$

where y is a q -dimensional vector, z is a p -dimensional vector, z_0 is a p -dimensional

⁴Available at <ftp://ftp.ncep.noaa.gov/pub/precip/50yr/gauge/2.5deg>

bias vector, M is a $p \times p$ matrix of parameters and ϵ_T is a (multivariate) normally distributed error with zero mean and $q \times q$ error covariance matrix T . The equation can be written as the following probability model:

$$y|z \sim N(M(z - z_0), T), \quad (2)$$

where the standard statistical symbol ' $|$ ' denotes “given” (conditional upon) and $\sim N(\cdot)$ means (multi-variate) normally distributed.

The normality assumption is generally valid for seasonal rainfall anomalies. Figure 3a shows the Yule-Kendall skewness statistics for seasonal mean rainfall anomalies. This statistics provides a resistant measure of asymmetry of the distribution and is defined as

$$\gamma = \frac{y_{0.25} - 2y_{0.5} + y_{0.75}}{y_{0.75} - y_{0.25}}, \quad (3)$$

where $y_{0.25}$, $y_{0.5}$ and $y_{0.75}$ are the lower quartile, the median and the upper quartile of y , respectively. Several regions of South America have γ close to zero, indicating that seasonal mean rainfall anomalies for these regions are closely approximated by a normal distribution. The normality assumption substantially simplifies both modelling and parameter estimation.

[Figure 3 about here.]

The matrices M and T and the bias vector z_0 can be obtained using ordinary

least squares estimation:

$$\hat{M} = S_{yz}S_{zz}^{-1} \quad (4)$$

$$\hat{T} = S_{yy} - S_{yz}S_{zz}^{-1}S_{yz}^T \quad (5)$$

$$\hat{z}_0 = -(\bar{y} - \bar{z}M^T)M(M^TM)^{-1} \quad (6)$$

where S_{zz} is the $(p \times p)$ covariance matrix of sea surface temperature anomalies, S_{yy} is the $(q \times q)$ covariance matrix of rainfall anomalies, and S_{yz} is the $(q \times p)$ cross-covariance matrix. Overbars denote time means of y and z , X^T denotes the transpose of matrix X and X^{-1} the inverse of matrix X . The common period of rainfall and sea surface temperature anomalies used for estimation in this study is 1959-2001 ($n = 43$ summers).

Reliable parameter estimation is difficult because of the large dimensionality of gridded data sets (e.g. $p=2761$ grid points of sea surface temperature anomalies over the Pacific and Atlantic and $q=312$ grid points of rainfall anomalies over South America) and the strong dependency between values at neighbouring grid points. Poor conditioning of matrices such as S_{zz} makes parameter estimation unreliable. This problem can be circumvented using multivariate dimension reduction techniques to reduce the dimensionality of the data sets. Instead of considering grid point variables, one can project the data onto a small set of leading spatial patterns to obtain a small number of indices. In the example presented here, MCA has been used to extract leading co-varying

modes of sea surface temperature and rainfall anomalies. See Stephenson *et al.* (2005) for an explanation of other dimensional reduction techniques that can be used to improve parameter estimation. A large number of MCA modes have been retained and tested. It was found that MCA with 6 modes gave the best cross-validated forecast results, which are hereafter shown in this paper. The first 6 modes account for 84.8% of the squared covariance between sea surface temperature and rainfall anomalies. Figure 4 shows the squared covariance fraction (SCF) as a function of the number of modes for the MCA between sea surface temperature and rainfall anomalies. This figure reveals that the SCF drops monotonically until 6 modes. After 6 modes a very small amount of the squared covariance is accounted for by each additional MCA mode.

[Figure 4 about here.]

Figure 5 shows correlation maps (spatial patterns) and the expansion coefficients (time series) of the first mode of the MCA analysis between the sea surface temperature anomalies over the Pacific and Atlantic oceans and the South American rainfall anomalies over the period 1959-2001. Correlation maps are obtained by correlating the expansion coefficient time series of one field (e.g. rainfall) with the observed grid point values of the other field (e.g. sea surface temperature). Correlations with magnitude greater than 0.3 are statistically significant at the 5% level using a two-sided Student's t-test. This first mode accounts for a large amount (51.5%) of the squared covariance between sea sur-

face temperature and rainfall. The sea surface temperature pattern (Fig. 3a) shows basin-wide correlations in the equatorial Pacific related to ENSO. Warm (El Niño) years are evident as positive peaks in the time series of Fig. 5c and cold (La Niña) years are evident by minima in these time series. The rainfall pattern (Fig. 5b) has negative correlations over northern South America and positive correlations over southern Brazil, Uruguay, Paraguay, northern Argentina and Ecuador. This figure reveals a dipole pattern that during El Niño years is marked by deficit of rainfall in northern South America and excess of rainfall in southern Brazil, Uruguay, Paraguay and northern Argentina. During La Niña years this pattern is reversed. A similar ENSO pattern to Fig. 5 has been identified by Ropelewski and Halpert (1987; 1989), Kiladis and Diaz (1989) and Peagle and Mo (2002) and is in accordance with the ENSO composites shown in Fig. 2. The second and the third MCA modes (not shown) account for 14.7% and 7.2% of the squared covariance, respectively, and relate sea surface temperature variability in the Atlantic ocean and rainfall over South America. The second mode has positive correlation between sea surface temperatures in the equatorial Atlantic and rainfall over the northeast region of Brazil. The third mode positively relates sea surface temperatures over a large area of the Atlantic with rainfall over central and northwestern South America.

[Figure 5 about here.]

The empirical predictions (EMP) are performed as follows:

1. To avoid artificial skill, the model parameters are estimated for each summer using only data for all the other summers (cross-validation).
2. Time means are subtracted from the $n - 1$ sea surface temperature and rainfall observations to make anomalies stored in a $(n - 1 \times q)$ data matrix Y and a $(n - 1 \times p)$ data matrix Z , respectively.
3. An SVD analysis is performed of the matrix $Y^T Z$ to determine the leading MCA modes U and V in $Y^T Z = U \Sigma V^T$.
4. A multivariate regression of the k -leading MCA rainfall and sea surface temperature modes is performed in order to estimate M , z_0 , and T .
5. The estimated quantities M , z_0 , and T are then used to predict the rainfall anomalies for the removed year using the sea surface temperature anomalies available for that year.

b. Coupled multi-model ensemble predictions

An ensemble forecast of an individual coupled model samples uncertainties in the initial conditions used to produce the forecast. Uncertainties in the model formulation are not sampled by this single-model ensemble approach. However, different models use different numerical and parameterisation schemes to represent mathematically the same physical processes. The multi-model ensemble, consisting of ensemble predictions produced by different climate research

institutions, helps sample uncertainties due to model formulation.

The multi-model forecasting system used here consists of the three state-of-the-art European coupled ocean-atmosphere models listed in Table 2. These models were run as part of the DEMETER project at ECMWF (Palmer *et al.* 2004; Hagedorn *et al.* 2005) to produce multi-model ensemble hindcasts (i.e. retrospective forecasts produced after the events are observed) for the period 1959 to 2001 (43 years). In fact a total of seven coupled models were run in DEMETER. The reason for using only predictions of the three models of Table 2 is because they produced the longest time series of hindcasts. The coupled models were run four times per year, starting on the 1st day of February, May, August and November at 00:00 GMT. A nine member ensemble forecast was made for each coupled model for the following six months. Wind stress and sea surface temperature perturbations were used to generate the ensemble members for each model. Atmospheric and land-surface initial conditions were taken from the ERA-40 reanalysis. All model predictions were bi-linearly interpolated onto the same $2.5^{\circ} \times 2.5^{\circ}$ grid. More details of the experiments are described in Palmer *et al.* (2004) and Hagedorn *et al.* (2005).

[Table 2 about here.]

The 27 member ensemble of the three coupled models of Table 2 is used here to produce predictions of austral summer rainfall for South America. These predictions are obtained by computing the ensemble mean and variance of all

members of the ensemble and are referred to as coupled multi-model ensemble predictions (ENS) or simply multi-model predictions.

4. Calibration and combination of forecasts

The objective forecast assimilation procedure of Coelho (2005) and Stephenson *et al.* (2005) is used here for the calibration and combination of South American rainfall anomaly predictions produced by:

- the three DEMETER coupled models listed in Table 2; and
- the empirical model of section 3a.

A useful feature of forecast assimilation is that it allows predicted patterns to be shifted spatially in order to correct for model biases in coupled model predictions. In other words, the procedure accounts for inter-grid point dependencies, whereas the combination methods used by Rajagopalan *et al.* (2002), Mason and Mimmack (2002) and Robertson *et al.* (2005) are performed at grid points individually. Appendix A summarises the forecast assimilation procedure.

Three different forecast experiments have been produced with forecasts assimilation:

- *Integrated forecasts* produced with prior distribution estimated using rainfall observations over the calibration period 1959-2001; and using empirical

predictions in addition to coupled model predictions in the forecast assimilation procedure. In other words, the matrix of forecasts used in forecast assimilation consisted of predictions of the three DEMETER coupled models in addition to empirical model predictions (see Appendix A for further details)

- *Integrated forecasts with empirical prior* produced with empirical predictions as estimates of the prior distribution; and only the three DEMETER coupled model predictions are used in the forecast assimilation procedure
- *Coupled model integrated forecasts* produced with prior distribution estimated using rainfall observations over the calibration period 1959-2001; and only the three DEMETER coupled model predictions are used in the forecast assimilation procedure. Note that these forecasts do not incorporate empirical predictions.

The second and third experiments above have been performed as sensitive tests for the first experiment. The second experiment was designed to check whether or not the use of empirical predictions as estimates for the prior distribution could provide better quality forecasts than integrated forecasts of the first experiment. The third experiment aimed to check if the exclusion of empirical predictions of the first experiment would impact in gain or loss of forecast skill. Results indicate that integrated forecasts obtained in the first experiment have slightly better skill than the forecasts of the other two experiments. For this reason only integrated forecasts of the first experiment will be shown and discussed

hereafter. These forecasts are referred to as integrated forecasts (INT).

5. Forecast skill

This section assesses the skill of austral summer rainfall forecasts. Both deterministic and probabilistic skill measures are presented. The skill of integrated forecasts obtained with forecast assimilation is compared to the skill of both coupled multi-model ensemble predictions and empirical predictions previously defined in section 3.

Figure 6 shows observed austral summer rainfall anomalies during 1982/83, 1988/89 and 1998/99 (first row), and the forecast anomalies for these three years produced by the empirical model (second row), the coupled multi-model ensemble (third row) and the integrated system (fourth row). The spatial correlation between the observed and the forecast anomalies is shown in the bottom right hand corner of Figs. 6d-l. Empirical predictions are able to reproduce the observed pattern of negative anomalies in northern South America and positive anomalies in southern Brazil during 1982/83 (Fig. 6d) and a reverse pattern during 1988/89 (Fig. 6e). In 1998/99 the empirical model failed to successfully reproduce the observed pattern (Fig. 6f). The coupled multi-model prediction was able to reproduce the observed pattern of negative anomalies in northern South America and positive anomalies in coastal Peru and Ecuador during 1982/83, but failed to reproduce the pattern of positive anomalies observed in southern

Brazil (Fig. 6g). In 1988/89 the coupled multi-model partially reproduced the observed pattern of positive anomalies in northern South America and negative anomalies in southern Brazil (Fig. 6h). In 1998/99, as for the empirical prediction (Fig. 6f), the coupled multi-model failed to reproduce the observed pattern (Fig. 6i). When empirical and coupled multi-model predictions were combined and calibrated with forecast assimilation, much better integrated forecasts were obtained. Integrated forecasts (Figs. 6j-l) are in better agreement with the observations (Figs. 6a-c). Integrated forecast for 1982/83 have a spatial correlation of 0.66 (Fig. 6j), which is larger than the values of 0.37 and 0.58 obtained for empirical and coupled multi-model predictions, respectively (Figs. 6d and 6g). The integrated forecast for 1988/89 have a spatial correlation of 0.42 (Fig. 6k), which is larger than the value of 0.33 of the empirical prediction (Fig. 6e) and slightly smaller than the value of 0.44 of the coupled multi-model prediction (Fig. 6h). Note, however, that the integrated forecast reproduces much better the observed negative anomalies in southern Brazil than the coupled multi-model prediction. Integrated forecast for 1998/99 have a spatial correlation of 0.24 (Fig. 6l), which is much larger than the values of 0.05 obtained for empirical and coupled multi-model predictions (Figs. 6f and 6i).

[Figure 6 about here.]

Figure 7 shows Brier score (Brier 1950) maps and its reliability and resolution components for empirical (first row), multi-model (second row) and inte-

grated forecasts (third row) for the period 1959-2001. See Appendix B for definition and further information about the Brier score and its decomposition. The Brier score is for the binary event defined by negative seasonal mean rainfall anomalies. The tropical region, in northern South America, and the subtropics (southern Brazil, Uruguay, Paraguay and northern Argentina) are the most skillful regions with Brier scores less than 0.25 (Figs. 7a, 7d and 7g). Interestingly, these two regions have the Yule-Kendall skewness statistics of rainfall anomalies closer to zero than the other regions of South America (Fig. 3). This suggests that forecast skill for the other regions of South America might be improved if another distribution, different from the multi-variate normal distribution, is used in the forecast assimilation procedure. As previously discussed, the tropical and southeastern South America are influenced by ENSO. Figs. 7a, 7d and 7g suggest that most of the skill of South American rainfall predictions is ENSO derived in accordance with Figs. 5 and 10, which show that most of the predictability of South American austral summer rainfall is ENSO-related.

The comparison of Brier score maps of Figs. 7a and 7d reveals that empirical and multi-model predictions have similar level of skill. Integrated forecasts obtained with forecast assimilation have improved skill over both empirical and multi-model predictions alone (Fig. 7g). Integrated forecasts (Fig. 7g) have larger areas with Brier scores below 0.25 than empirical (Fig. 7a) and coupled-model (Fig. 7d) predictions. Additionally, integrated forecasts also show predictive skill over the south of the northeast region of Brazil, whereas empirical

and coupled-model forecasts have poor predictive skill over this region. The improved predictive skill obtained with the integrated system is mainly due to improvements in the reliability of the forecasts (Figs. 7b, 7e and 7h), although improvements in the resolution of the forecasts are also noticed in tropical South America and southern Brazil (Figs. 7c, 7f and 7i).

[Figure 7 about here.]

Figure 8 shows the mean anomaly correlation coefficient (ACC) for La Niña, neutral and El Niño years occurred during 1959-2001 (listed in Table 1) and all (1959-2001) years. The ACC of each year is given by the correlation between the observed and the predicted spatial anomaly pattern (Jolliffe and Stephenson 2003, their section 6.3.1). The most striking feature of Fig. 8 is the little magnitude of the mean ACC (inferior to 0.3). This illustrates that austral summer South American rainfall forecast skill is low. La Niña and El Niño years have higher mean ACC than neutral years, indicating that predictions for ENSO years are more skillful than predictions for neutral years. Note, however, that predictions for El Niño years are more skillful than predictions for La Niña years. El Niño and La Niña integrated forecasts obtained with forecast assimilation show an increase in the mean ACC compared to empirical and multi-model predictions. This is because forecast assimilation has shifted model predicted patterns towards observed patterns. Neutral years have very little mean ACC, indicating that rainfall anomalies of these years are hardly

predicted. The higher predictability of ENSO years compared to neutral years supports the argument that most of the skill of austral summer South American rainfall forecasts is ENSO derived. The vertical solid lines on the top of the white bars of Fig. 8 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. This means that empirical and multi-model predictions have similar level of skill when forecasting rainfall of La Niña years and the all (1959-2001) years.

[Figure 8 about here.]

6. Conclusions

This study has addressed predictability of austral summer mean South American rainfall by proposing an integrated seasonal forecasting system for South America. The proposed integrated system has two components: a) an empirical model that uses Pacific and Atlantic sea surface temperature anomalies as predictor for South American rainfall, and b) a multi-model coupled system composed by ECMWF, CNRM and UKMO models. These models constitute the operational European multi-model seasonal forecasting system, which is

hosted at ECMWF. Empirical and coupled model predictions were combined and calibrated using a Bayesian forecast assimilation procedure (Coelho 2005; Stephenson *et al.* 2005) in order to produce a single integrated probabilistic forecast. The objective calibration and combination of empirical and multi-model coupled predictions proposed here makes this a first step towards an integrated forecasting system for issuing South American seasonal forecasts.

The proposed integrated system could feasibly be implemented at any operational weather service in South America (e.g. the Centre for Weather Prediction and Climate Studies (CPTEC), in Brazil). The proposed system can be expanded by the inclusion of CPTEC coupled model predictions in addition to ECMWF, CNRM and UKMO coupled model predictions in the multi-model system. However, the feasibility of the implementation of such a new system depends on the establishment of international cooperation between CPTEC, ECMWF, CNRM and UKMO as well as other South American weather services. Such a system could be used as an additional tool for producing objective probabilistic climate forecasts during regional climate outlook forums, which are regularly sponsored by the World Meteorological Organization. A similar approach could be of use in other regions of the world (e.g. North America).

In order to answer the question of whether or not forecasts produced by the proposed integrated system are better than those produced by a simple empirical model or by the simple multi-model ensemble alone, the skill of empirical, coupled multi-model and integrated forecasts obtained with forecast assimila-

tion has been assessed and compared. This comparison revealed that when seasonally forecasting South American austral summer rainfall at 3-month lead-time the current generation of coupled models has comparable level of skill to those obtained using a simplified empirical approach. The same conclusion holds for shorter (e.g. 1-month) lead times (Coelho *et al.* 2005). This result is in agreement with findings of previous comparison studies (e.g. Folland *et al.* 2001; Moura and Hastenrath 2004). This implies that both empirical and coupled model predictions are comparable to each other. However, when empirical and coupled multi-model forecasts have been combined and calibrated with forecast assimilation more skillful integrated forecasts than either empirical or coupled multi-models predictions alone have been obtained. This result demonstrates that forecast assimilation can be used for improving the quality of South American seasonal predictions. The resulting integrated forecasts have been shown to have improved Brier scores compared to both empirical and the simple multi-model prediction over some regions of South America. Forecast assimilation improved both the reliability and resolution of the predictions in tropical South America. Southeastern South America – an important region for South American hydroelectricity and crop yield production – and the northeast region of Brazil also had the reliability of the predictions improved. Recent results demonstrate that forecast assimilation is also useful for local downscaling of rainfall and river flow anomalies for southeastern South America (work in progress).

The tropics and southern Brazil, Uruguay, Paraguay 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. It is worth stressing that the skill of austral summer South American rainfall predictions produced with the current generation of coupled and empirical models is still low (c.f. values of mean ACC less than 0.3). This suggests that a large amount of research is still required in order to improve the quality of these predictions.

It would be interesting in the future to extend the method used here for combining and calibrating predictions to non-normally distributed data and to deal with weather and climate extremes. The use of non-normal distributions might improve forecast skill over those regions of South America where the normality assumption is not strictly valid. Non-stationarity of climate can also affect forecast calibration. The development of more generalised methods capable of dealing with non-stationary time series might also help to improve forecast skill. Model selection prior to the combination of predictions might also help to improve forecast skill.

Acknowledgements

We wish to thank Prof. D. L. T. Anderson and Dr. T. N. Palmer, who kindly provided the ECMWF coupled model hindcasts used in this research. CASC was sponsored by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) process 200826/00-0. FJDR was supported by DEMETER (EVK2-1999-00024).

Appendix

A. Bayesian forecast assimilation

Forecast assimilation (Coelho 2005 and Stephenson *et al.* 2005) is based upon Bayesian updating of prior information when new information becomes available (Bayes 1763). If one has a first guess of the prior distribution $p(y)$ of a particular variable of forecast interest y (e.g. rainfall) and additional (new) prediction information x then becomes available (e.g. an ensemble of predictions), then it is possible to update $p(y)$ to obtain the posterior conditioned probability density function $p(y|x)$ by making use of Bayes' theorem

$$p(y|x) = \frac{p(y)p(x|y)}{p(x)}. \quad (7)$$

The distribution of climate model ensemble predictions gives an estimate of $p(x)$. However, one is really interested in the posterior distribution $p(y|x)$ not $p(x)$. Because of uncertainties in model formulation and in initial conditions, climate predictions in model space deviate away from the true evolution in observation space. Model predictions x should then be considered as proxy information that can be used to infer the probability of future observables y (Glahn 2004; Stephenson *et al.* 2005; Jolliffe and Stephenson 2005). To make inference about future observables one needs a probability model that can give the probability $p(y|x)$ of future observable quantities y when provided with model prediction data x , such as the model of Eqn. 7.

Unrealistic predictions may be due to biases linked to weakness of the method used for ensemble generation (Atger 2003). Hence, calibration by inflation of the ensemble spread is often performed in order to improve *forecast reliability* (e.g. Hamill and Colucci 1998; von Storch 1999). Reliability refers to the correspondence between the forecast probability of an event and the relative frequency of the event conditioned upon the forecast probability (Jolliffe and Stephenson 2003). Reliability is a measure of forecast uncertainty correctness and assesses the calibration of the forecasts. Forecast skill depends on the ability to discriminate between observable outcomes, which is known as *forecast resolution* (Jolliffe and Stephenson 2003). See Appendix B for additional information about forecast reliability and resolution. In summary, a good forecasting system should be able to produce reliable forecasts and also discriminate

between different observed situations.

The likelihood $p(x|y)$ is an essential ingredient in the Bayesian forecast assimilation updating procedure. It can easily be estimated by regression of past model predictions x on past observations y . However, as for the empirical model of section 3a, because of the large dimensionality of gridded data sets compared to the number of independent forecasts and the dependency between values at neighbouring grid points, multi-variate linear regression of a few MCA leading modes has been applied.

As described in Coelho *et al.* (2004), Coelho *et al.* (2003), Coelho (2005) and Stephenson *et al.* (2005), the Bayesian procedure used to perform forecast assimilation has three main ingredients: a) estimate the prior distribution $p(y)$; b) modelling of the likelihood function $p(x|y)$; and c) use of Bayes theorem to find the posterior distribution $p(y|x)$ from $p(x)$ and $p(x|y)$. For simplicity, it has been assumed that both prior and likelihood distributions are multi-variate normal (Gaussian), leading to a multi-variate normal posterior distribution. Analysis of skewness γ reveal that this assumption is generally acceptable for observed and predicted South American seasonal rainfall anomalies. Several regions of South America have observed and predicted seasonal rainfall skewness close to zero and therefore are not very far from following a normal distribution (Fig. 3).

The full equations of the multi-variate normal model used here to perform

Bayesian forecast assimilation are given in Coelho (2005) and Stephenson *et al* (2005). All results presented here were obtained with the cross-validation “leave one out” method (Wilks 1995, section 6.3.6). The matrix of predictions needed to perform forecast assimilation in the first experiment of section 4 (*Integrated forecasts*) was composed by model predictions of the three DEMETER coupled models of Table 1 and empirical model predictions produced by the model introduced in section 3a. For the second (*Integrated forecasts with empirical prior*) and third (*Coupled model integrated forecasts*) experiments of section 4 the matrix of predictions needed to perform forecast assimilation was composed by model predictions of the three DEMETER coupled models of Table 1.

For the first and third experiment of section 4, the prior distribution $p(y)$ was estimated using the mean and covariance of rainfall observations over the calibration period 1959-2001. For the second experiment of section 4 empirical predictions were used as estimates for the prior distribution $p(y)$. Because of the better quality of the forecasts obtained in the first experiment when compared to the other two experiments, only results of the first experiment are shown and discussed here. The first three leading modes of the MCA between observed and model predictions of South American rainfall anomalies were used in the forecast assimilation procedure of the first experiment. A large number of modes was retained and tested. Forecast assimilation with 3 modes gave the best cross-validated forecast results. These three modes account for 87.8% of

the squared covariance between observed and predicted rainfall. Figure 9 shows the SCF as a function of the number of modes. This figure reveals that the SCF drops monotonically until 3 modes. After 3 modes a very small amount of the squared covariance is accounted for by each additional MCA mode.

[Figure 9 about here.]

Figure 10 shows correlation maps (spatial patterns) and the expansion coefficients (time series) of the leading mode of the MCA between observed and model predicted South American rainfall anomalies. Correlation maps are obtained by correlating the prediction time series of expansion coefficients with the observed and predicted grid point values of each model. Correlations with magnitude greater than 0.3 are statistically significant at the 5% level using a two-sided Student's t-test. The leading mode accounts for 77.7% of the squared covariance between observed and predicted rainfall. The pattern of observed rainfall (Fig. 10a) shows a similar pattern to the rainfall pattern of the first MCA mode of Fig. 5b, which is related to ENSO. The correlation between the expansion coefficients (time series) of observed rainfall of the first MCA mode (solid line in Fig. 10f) and the expansion coefficients (time series) of observed rainfall of the first MCA mode of Fig. 5c (solid line) is 0.98. This correlation is statistically significant at the 1% level using a two-sided Student's t-test.

Figures 10b-e show correlation maps (spatial patterns) of the predictions produced by the four models investigated here (CNRM, ECMWF, UKMO and

EMP). The spatial structure of these patterns (Figs 10b-e) when compared to the observed pattern (Fig. 10a) provides an indication of the ability of these models to reproduce the observed rainfall. The magnitude of the correlations of Figs 10b-e gives an indication of the weights attributed to each model in the forecast assimilation procedure. The four models are able to reproduce the observed negative correlations over central northern South America, although the area of negative correlations in the three coupled models is much larger than observed. The pattern of positive correlation in northwestern South America, near Ecuador, is captured by the ECMWF, UKMO and EMP models, whereas CNRM fails to reproduce this feature. All four models are able to capture the sign of positive correlations in southern Brazil, Uruguay, Paraguay and northern Argentina, although the location of the maximum correlation does not perfectly match the observations. The second and the third MCA modes account for 6.2% and 3.9% of the squared covariance between observed and predicted rainfall and do not resemble any previously published mode of climate variability (not shown).

[Figure 10 about here.]

Finally, it is noteworthy that forecast assimilation has some advantages and some potential disadvantages:

Advantages of forecast assimilation

- produces well-calibrated probability forecasts

- able to deal with ensemble predictions
- able to deal with multi-model predictions
- preserves spatial structure present in the datasets
- allows spatial patterns to be shifted/corrected

Potential disadvantages of forecast assimilation

- need for data reduction to be able to estimate regression parameters
- relationships can change with time (stability)
- need to re-compute calibration equations (regression) each time the forecasting system changes

B. Brier score

The skill of forecast probabilities p_k for the event $p_k = Pr(y_k)$ can be assessed using the Brier score (Brier 1950) given by

$$BS = \frac{1}{n} \sum_{k=1}^n (p_k - o_k)^2, \quad (8)$$

where the index k denotes a numbering of the n forecast/observation pairs, $Pr(E)$ denotes the probability of the event E , and o_k is the corresponding binary observation (i.e., $o_k = 1$ when the event y_k has occurred and $o_k = 0$ when

the event y_k has not occurred).

The Brier score (BS) is analogous to the mean squared error, but instead of averaging squared differences between pairs of forecast and observed values, it averages the squared differences between pairs of forecast probabilities p_k and the subsequent binary observations o_k . The Brier score can take values in the range $0 \leq BS \leq 1$. The Brier score is negatively oriented, with perfect forecast exhibiting $BS = 0$. Therefore, the smaller the Brier score the better is the quality of the forecast.

Following Murphy (1973), the Brier score (Eqn. 8) can be expressed as the sum of three components:

$$BS = \underbrace{\frac{1}{n} \sum_{i=1}^I N_i (p_i - \bar{o}_i)^2}_{(reliability)} - \underbrace{\frac{1}{n} \sum_{i=1}^I N_i (\bar{o}_i - \bar{o})^2}_{(resolution)} + \underbrace{\bar{o}(1 - \bar{o})}_{(uncertainty)} \quad (9)$$

where N_i is the number of times each probability forecast p_i is used in the set of forecasts being verified. The total number of forecast/event pairs n is simply the sum of these counts: $n = \sum_{i=1}^I N_i$, where I is the number of discrete forecast values p_i . For each probability p_i , depicted by the I allowable forecast values, there is a relative frequency \bar{o}_i of the observed event. Since the observed event is dichotomous, a single conditional relative frequency defines the conditional distribution of observations given each forecast p_i . The subsample relative fre-

quency, or conditional average observation, is given by

$$\bar{o}_i = p(o_1|p_i) = \frac{1}{N_i} \sum_{k \in N_i} o_k, \quad (10)$$

where $o_k = 1$ if the event occurs for the k th forecast/event pair and $o_k = 0$ if it does not occur, and the summation is over only those values of k corresponding to occasions when the forecast p_i was issued. Similarly, the overall (unconditional) relative frequency, or sample climatology, of the observations is given by

$$\bar{o} = \frac{1}{n} \sum_{k=1}^n o_k. \quad (11)$$

More accurate forecasts are characterized by small values of Brier score. Therefore, the forecaster would aim for the reliability component of the Brier score to be as small as possible, and the resolution component to be as large (in the absolute sense) as possible. The uncertainty component depends only on the sample climatological relative frequency \bar{o} , and is not affected by the forecasts.

The reliability component summarizes the calibration, or conditional bias, of the forecasts. It consists of a weighted average of the squared differences between the forecast probabilities p_i and the relative frequencies of the forecast event in each subsample i . For perfectly reliable forecasts the subsample relative frequency \bar{o}_i is exactly equal to the forecast probability p_i in each subsample.

The relative frequency of the forecast event \bar{o}_i should be small when $p_i = 0$ is forecast, and should be large when $p_i = 1$ is forecast. When $p_i = 0.5$, \bar{o}_i should be near 0.5. For reliable, or well-calibrated forecasts, all the squared differences in the reliability component of the Brier score will be near zero, and their weighted average will be small.

The resolution component summarizes the ability of the forecasts to discern subsample relative frequencies forecasts \bar{o}_i from the observed overall sample climatology relative frequency \bar{o} . The forecast probabilities p_i do not appear explicitly in this term, yet it still depends on the forecasts through the sorting of the events making up the subsample relative frequency \bar{o}_i . The resolution component is a weighted average of the squared differences between \bar{o}_i and \bar{o} . Thus, if the forecasts sort the observations into subsamples having substantially different relative frequencies \bar{o}_i than the overall sample climatology \bar{o} , the resolution term will be large. This is a desirable situation, since the resolution component is subtracted in the Brier score decomposition equation. Conversely, if the forecasts sort the events into subsamples with very similar event relative frequencies \bar{o}_i , the squared differences in the summation of the resolution term will be small. In such a situation the forecasts resolve the event only weakly, and the resolution component will be small.

The uncertainty component depends only on the variability of the observations, and is not influenced by the forecasts. It has minima at zero when the climatological probability \bar{o} is either zero or one, and a maximum of 0.25 when

$\bar{o} = 0.5$. The uncertainty in the forecasting situation is small (close to zero) when the event being forecast almost always happens (\bar{o} close to 1) or when the event being forecast almost never happens (\bar{o} close to 0). In such situation, always forecasting the climatological probability \bar{o} will give generally good results. When the climatological probability is close to 0.5, there is substantially more uncertainty inherent in the forecasting situation, and the uncertainty component of the Brier score is commensurately larger.

References

- Anderson, J. L., H. van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Am. Meteorol. Soc.*, **80**, 1349–1361.
- Atger, F., 2003: Spatial and interannual variability of the reliability of ensemble-based probabilistic forecasts: Consequences for calibration. *Mon. Wea. Rev.*, **131**, 1509–1523.
- Barnston, A. G., M. H. Glantz, and Y. He, 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-98 El Niño episode and the 1998 La Niña onset. *Bull. Am. Meteorol. Soc.*, **80**, 217–243.
- Barnston, A. G., S. J. Mason, L. Goddard, D. G. DeWitt, and S. E. Zebiak, 2003: Multimodel ensembling in seasonal climate forecasting at IRI. *Bull. Am. Meteorol. Soc.*, **84**, 1783–1796.
- Barros, V. and G. E. Silvestri, 2002: The relation between sea surface temperature at the subtropical south-central Pacific and precipitation in southeastern South America. *J. Climate*, **15**, 251–267.
- Bayes, T., 1763: An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, **53**, 370–418.
- Berri, G. J. and P. L. Antico, 2005: Seasonal precipitation forecasts for the

- southeast of South America. Evaluating the first five years. *CLIVAR Exchanges*, **10**, 17; 25–27.
- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.
- Cavalcanti, I. F. A., J. A. Marengo, P. Satyamurty, C. A. Nobre, I. Trosnikov, J. P. Bonatti, A. O. Manzi, T. Tarasova, L. P. Pezzi, C. D’Almeida, G. Sampaio, C. C. Castro, M. B. Sanches, and H. Camargo, 2002: Global climatological features in a simulation using the CPTEC-COLA AGCM. *J. Climate*, **15**, 2965–2988.
- Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin, 2002: Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrometeorol.*, **3**, 249–266.
- Coelho, C. A. S., 2005: *Forecast calibration and combination: Bayesian assimilation of seasonal climate predictions*. Ph.D. thesis, Department of Meteorology, University of Reading, 178 pp.
- 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. *European Centre for Medium-Range Weather Forecasts Technical Memorandum number 426*, 16 pp.
- 2004: Forecast calibration and combination: A simple Bayesian approach for ENSO. *J. Climate*, **17**, 1504–1516.

- Coelho, C. A. S., D. B. Stephenson, F. J. Doblas-Reyes, and M. Balmaseda, 2005: From multi-model ensemble predictions to well-calibrated probability forecasts: Seasonal rainfall forecasts over South America 1959-2001. *CLIVAR Exchanges*, **10**, 14–20.
- Coelho, C. A. S., C. B. Uvo, and T. Ambrizzi, 2002: Exploring the impacts of the tropical Pacific SST on the precipitation patterns over South America during ENSO periods. *Theor. Appl. Climatol.*, **71**, 185–197.
- Diaz, A. F., C. D. Studzinski, and C. R. Mechoso, 1998: Relationships between precipitation anomalies in Uruguay and Southern Brazil and sea surface temperature in the Pacific and Atlantic oceans. *J. Climate*, **11**, 251–271.
- Folland, C. K., A. W. Colman, D. P. Rowell, and M. K. Davey, 2001: Predictability of Northeast Brazil rainfall and real-time forecast skill, 1987-98. *J. Climate*, **14**, 1937–1958.
- Glahn, B., 2004: Discussion of verification concepts in Forecast Verification: A Practitioner’s Guide in Atmospheric Science. *Weather and Forecasting*, **19**, 769–775.
- Greischar, L. and S. Hastenrath, 2000: The rainy seasons of the 1990s in Northeast Brazil: Real-time forecasts and verification. *J. Climate*, **13**, 3821–3826.
- Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: Basic concept. *Tellus*, In press.

- Hamill, T. and S. J. Colucci, 1998: Verification of Eta-RSM ensemble probabilistic precipitation forecasts. *Mon. Wea. Rev.*, **126**, 711–724.
- Jolliffe, I. N. and D. B. Stephenson, 2003: Forecast Verification: A practitioner's guide in atmospheric science. *Wiley and Sons*, First edition. 240 pp.
- Jolliffe, I. T. and D. B. Stephenson, 2005: Response of the book editors to Bob Glahn's Discussion of Verification Concepts in Forecast Verification: A Practitioner's Guide in Atmospheric Science. *Weather and Forecasting*, In press.
- Kiladis, G. N. and H. F. Diaz, 1989: Global climate anomalies associated with extremes in the Southern Oscillation. *J. Climate*, **2**, 1069–1090.
- Marengo, J., 1992: Interannual variability of surface climate in the Amazon basin. *Int. J. Climatol.*, **12**, 853–863.
- Marengo, J. A., I. F. A. Cavalcanti, P. Satyamurty, I. Trosnikov, C. A. Nobre, J. P. Bonatti, H. Camargo, G. Sampaio, M. B. Sanches, A. Manzi, C. A. C. Castro, C. D'Almeida, L. P. Pezzi, and L. Candido, 2003: Assessment of regional seasonal rainfall predictability using the CPTEC/COLA atmospheric GCM. *Clim. Dynamics*, **21**, 459–475.
- Martis, A., G. J. van Oldenborgh, and G. Burgers, 2002: Predicting rainfall in the Dutch Caribbean – More than El Niño. *Int. J. Climatol.*, **22**, 1219–1234.
- Mason, S. J. and G. M. Mimmack, 2002: Comparison of some statistical methods of probabilistic forecasting of ENSO. *J. Climate*, **15**, 8–29.

- Mechoso, C. R., S. W. Lyons, and J. A. Spahr, 1990: The impact of sea surface temperature anomalies on the rainfall over Northeast Brazil. *J. Climate*, **3**, 812–826.
- Moura, A. D. and S. Hastenrath, 2004: Climate prediction for Brazil's Nordeste: Performance of empirical and numerical modeling methods. *J. Climate*, **17**, 2667–2672.
- Moura, A. D. and J. Shukla, 1981: On the dynamics of droughts in northeast Brazil: Observations, theory and numerical experiments with general circulation model. *J. Atmos. Sci.*, **38**, 2653–2675.
- Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteorol.*, **12**, 595–600.
- Nobre, P. and J. Shukla, 1996: Variations of sea surface temperature, wind stress, and rainfall over the tropical atlantic and south america. *J. Climate*, **9**, 2464–2479.
- Paegle, J. N. and K. C. Mo, 2002: Linkages between summer rainfall variability over South America and sea surface temperature anomalies. *J. Climate*, **15**, 1389–1407.
- Palmer, T. N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Délecluse, M. Déqué, E. Díez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J.-F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonave, V. Marletto, A. P. Morse, B. Orfila,

- P. Rogel, J.-M. Terres, and M. C. Thomson, 2004: Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER). *Bull. Am. Meteorol. Soc.*, **85**, 853–872.
- Pezzi, L., V. Ubarana, and C. Repelli, 2000: Forecast and performance of a statistical regional model for seasonal climate prediction in Southern Brazil. *Rev. Bras. Geof.*, **18**, 129–146.
- Rajagopalan, B., U. Lall, and S. E. Zebiak, 2002: Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. *Mon. Wea. Rev.*, **130**, 1792–1811.
- Robertson, A. W., L. Upmanu, S. E. Zebiak, and L. Goddard, 2005: Improved combination of multiple atmospheric GCM ensembles for seasonal prediction. *Mon. Wea. Rev.*, In press.
- Ropelewski, C. F. and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626.
- 1989: Precipitation patterns associated with high index phase of Southern Oscillation. *J. Climate*, **2**, 268–284.
- Stephenson, D. B., C. A. S. Coelho, F. J. Doblas-Reyes, and M. Balmaseda, 2005: Forecast Assimilation: A unified framework for the combination of multi-model weather and climate predictions. *Tellus*, **57**, 253–264.

- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, **107**, 14291–14324.
- Uppala, S. M., K. P. W., A. J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Beljaars, J. van de Berg, N. Bidlot, S. Bormann, S. Caires, F. Chevallier, D. A., M. Dragosavac, M. Fisher, M. Fuentes, R. Jenne, M. A. P., J. Mahfouf, J. Morcrette, N. A. Rayner, R. W. Saunders, A. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, 2005: The ERA-40 Reanalysis. *Bull. Am. Meteorol. Soc.*, Submitted.
- Uvo, C. R. B., C. A. Repelli, S. E. Zebiak, and Y. Kushnir, 1998: On the relationships between tropical Pacific and Atlantic SST in northeast Brazil monthly precipitation. *J. Climate*, **13**, 287–293.
- van Oldenborgh, G. J., M. A. Balmaseda, L. Ferranti, T. N. Stockdale, and D. L. T. Anderson, 2005: Did the ECMWF seasonal forecast model outperform a statistical model over the last 15 years? *J. Climate*, In press.
- von Storch, H., 1999: On the use of inflation in statistical down-scaling. *J. Climate*, **12**, 3505–3506.

von Storch, H. and F. W. Zwiers, 1999: Statistical Analysis in Climate Research.

Cambridge University Press, ISBN 0521 450713. 484 pp.

Wallace, J. M. and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 785–812.

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

List of Tables

- 1 La Niña, neutral and El Niño years occurred during 1959-2001
as defined by the Climate Prediction Center (<http://www.cpc.noaa.gov/>). 50
- 2 Three European coupled ocean-atmosphere models used for pro-
ducing seasonal forecasts for South America. 51

	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:

Institution	Acronym	Country
Météo-France (Centre National de Recherches Météorologiques)	CNRM	France
European Centre for Medium-Range Weather Forecasts	ECMWF	International organisation
United Kingdom Met Office	UKMO	U.K.

Table 2:

List of Figures

- 1 a) Total annual mean rainfall. b) Total November-December-January mean rainfall. c) Annual total standard deviation. d) November-December-January total standard deviation. Units are in millimetres. Climatological reference is 1948-2001. . . . 56

- 2 November-December-January a) La Niña and b) El Niño composites for those years listed in Table 1. Composites are given by the quantity $c = \left(\frac{\bar{y}_*}{\bar{y}} - 1 \right) 100\%$, where \bar{y}_* is the mean November-December-January rainfall of La Niña and El Niño years of Table 1 and \bar{y} is the 1948-2001 mean November-December-January rainfall. Regions with positive anomalies (i.e. $\bar{y}_* > \bar{y}$) have the quantity $c > 0$. Regions with negative anomalies (i.e. $\bar{y}_* < \bar{y}$) have the quantity $c < 0$ 57

- 3 November-December-January 1959-2001 Yule-Kendall skewness statistics γ of rainfall anomalies. a) Observations, b) CNRM, c) ECMWF and d) UKMO coupled model predictions. 58

- 4 Squared covariance fraction as a function of the number of modes
for the MCA between observed May-June-July sea surface tem-
perature and observed November-December-January South Amer-
ica rainfall anomalies for the period 1959-2001. The first 6
modes account for 84.8% of the covariance between observed
sea surface temperature and observed rainfall anomalies. 59

- 5 First MCA mode between MJJ sea surface temperature anoma-
lies and NDJ South America rainfall anomalies for the period
1959-2001. The squared covariance fraction (SCF), which in-
dicates the percentage of the total squared covariance between
MJJ sea surface temperature anomalies and NDJ South Amer-
ica rainfall anomalies explained by this mode, is 51.5%. a) sea
surface temperature (SST) correlation pattern. b) Rainfall cor-
relation pattern. c) Expansion coefficients (time series) of SST
(dashed line) and rainfall (solid line). The correlation r between
these two time series is indicated in panel c. 60

- 6 Austral summer rainfall anomalies (mm) for 1982/83, 1988/89
and 1998/99. Observations (first row), empirical forecasts (sec-
ond row), coupled multi-model ensemble forecasts (third row)
and integrated forecasts (fourth row). 61

7	Brier scores for the binary event defined by negative seasonal mean rainfall anomalies and its reliability and resolution components for empirical (first row), ensemble (second row) and integrated (third row) forecasts. Brier scores are for austral summer forecasts for the period 1959-2001. The reliability and resolution components were estimated using 10 equally spaced probability bins from 0 to 1.	62
8	Mean austral summer 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 lines on the top of the white bars indicate the 95% confidence interval for the mean ACC of empirical forecasts, which were obtained using a bootstrap resampling procedure (Wilks 1995, section 5.3.2).	63
9	Squared covariance fraction as a function of the number of modes for the MCA between observed and predicted austral summer South American rainfall anomalies for the period 1959-2001. . .	64

- 10 Correlation patterns of the first MCA mode between observed and predicted austral summer South America rainfall anomalies for the period 1959-2001. The SCF is 77.7%. a) Observation. b) CNRM. c) ECMWF. d) UKMO. e) EMP f) Expansion coefficients (time series) of observed rainfall (solid line) and predicted rainfall of the four models (dashed line). The correlation r between these two time series is indicated in panel f. 65

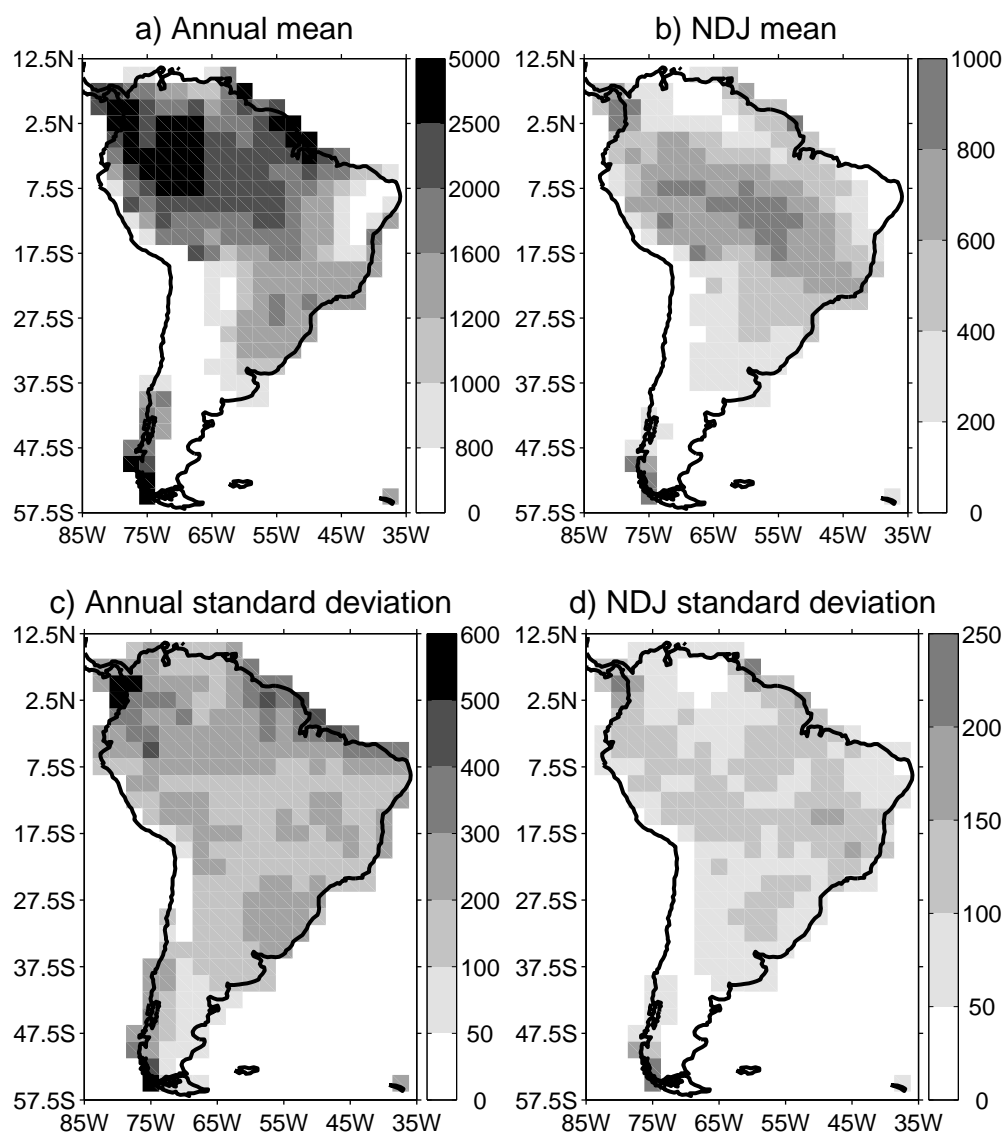


Figure 1:

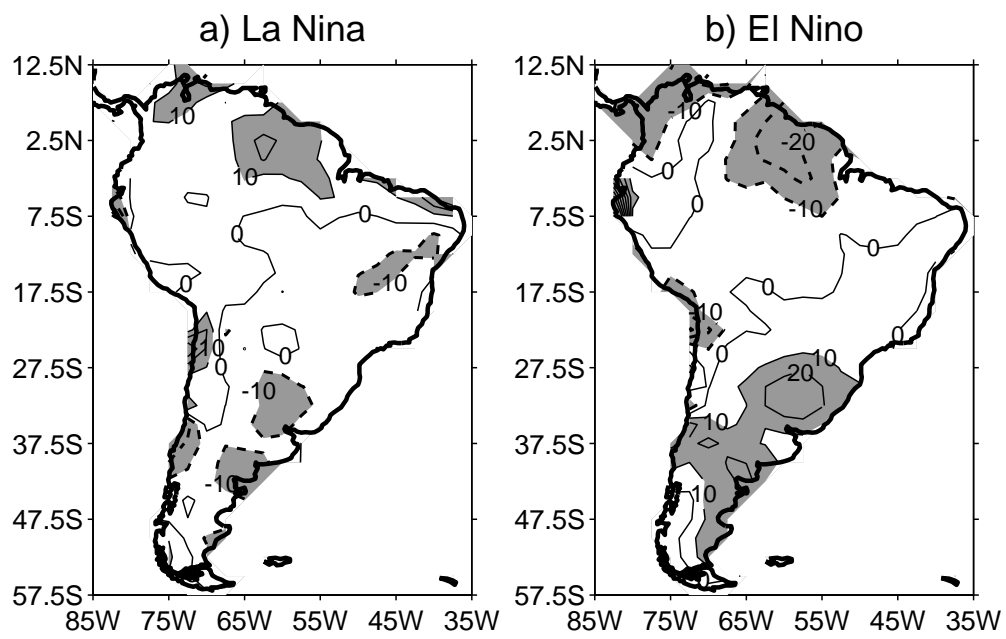


Figure 2:

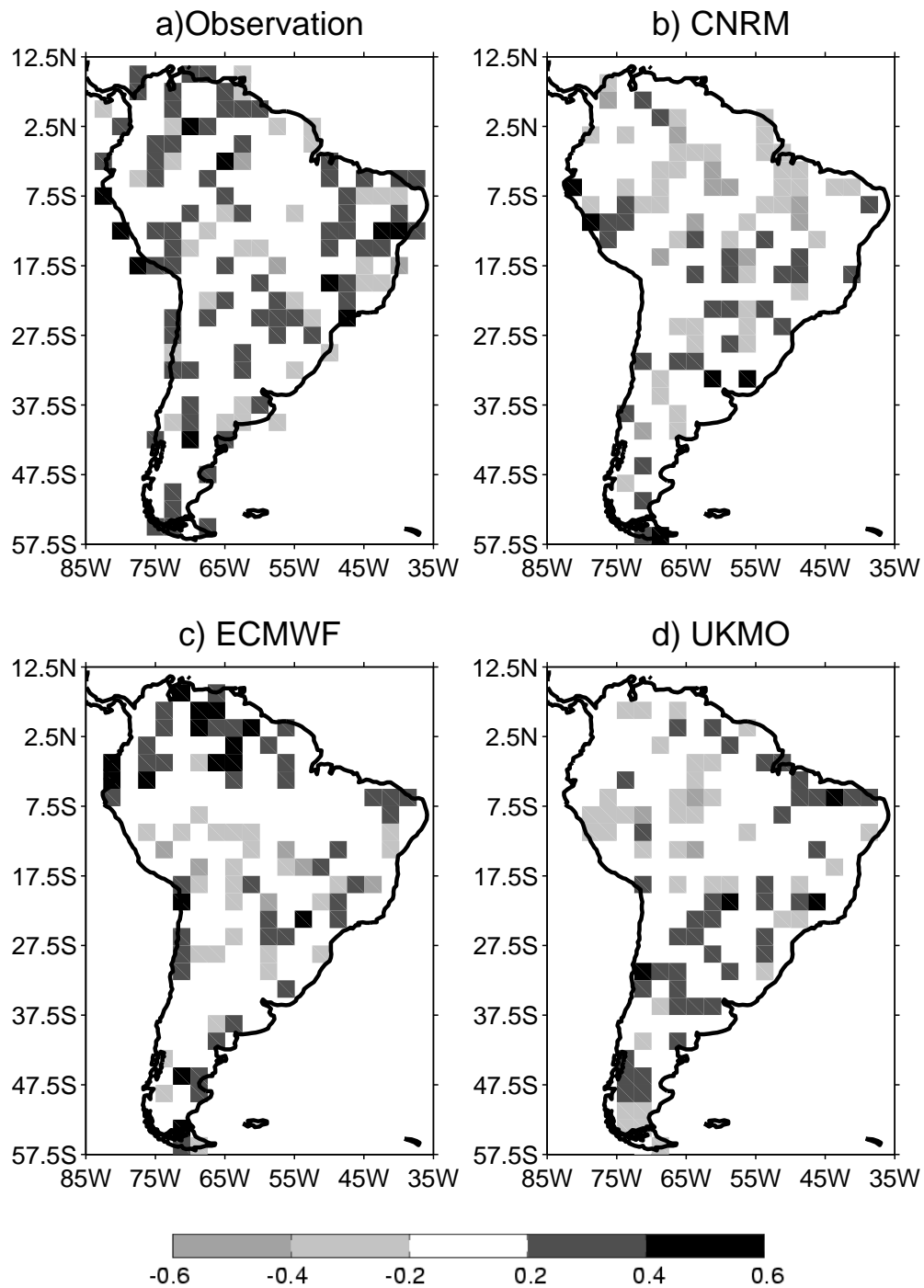


Figure 3:

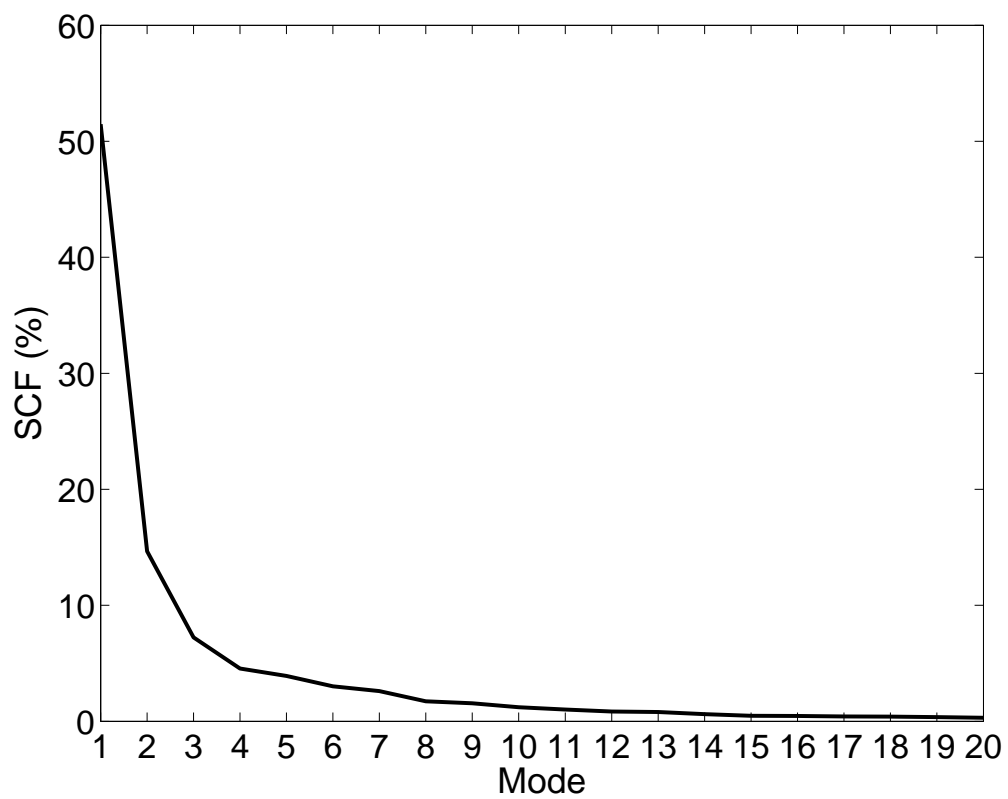


Figure 4:

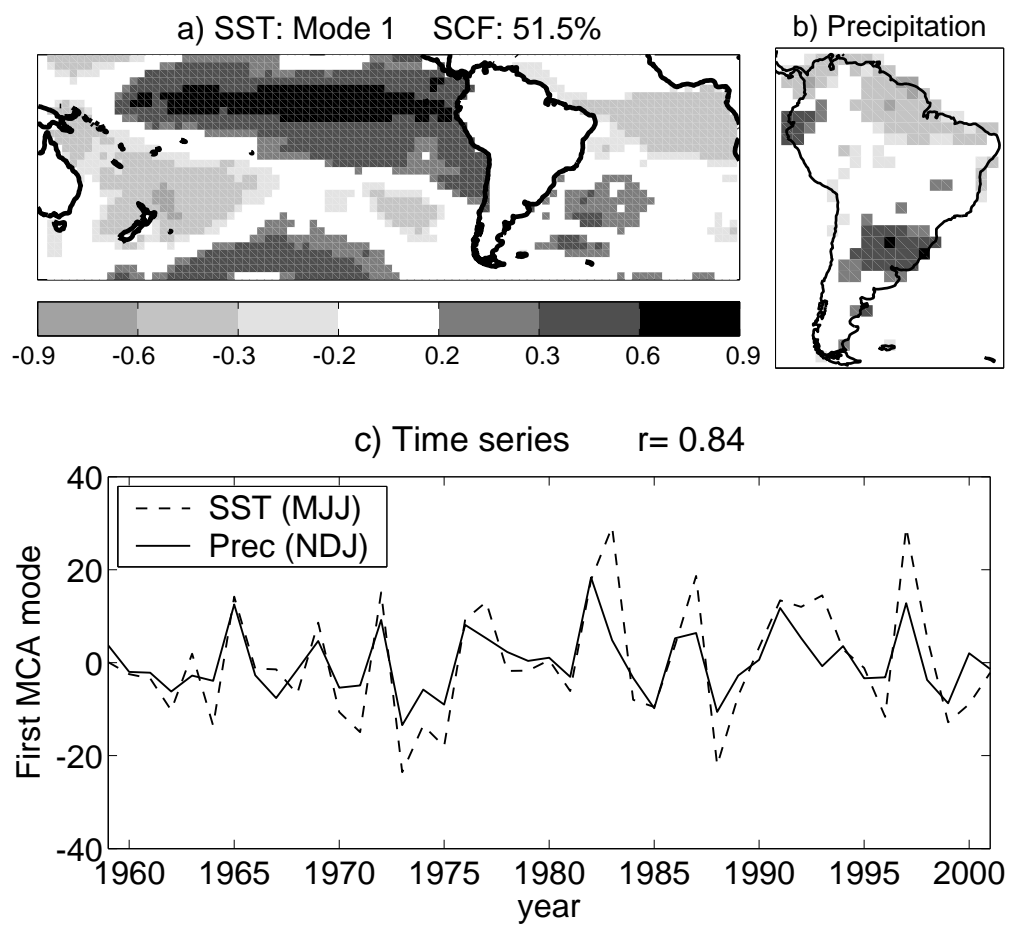


Figure 5:

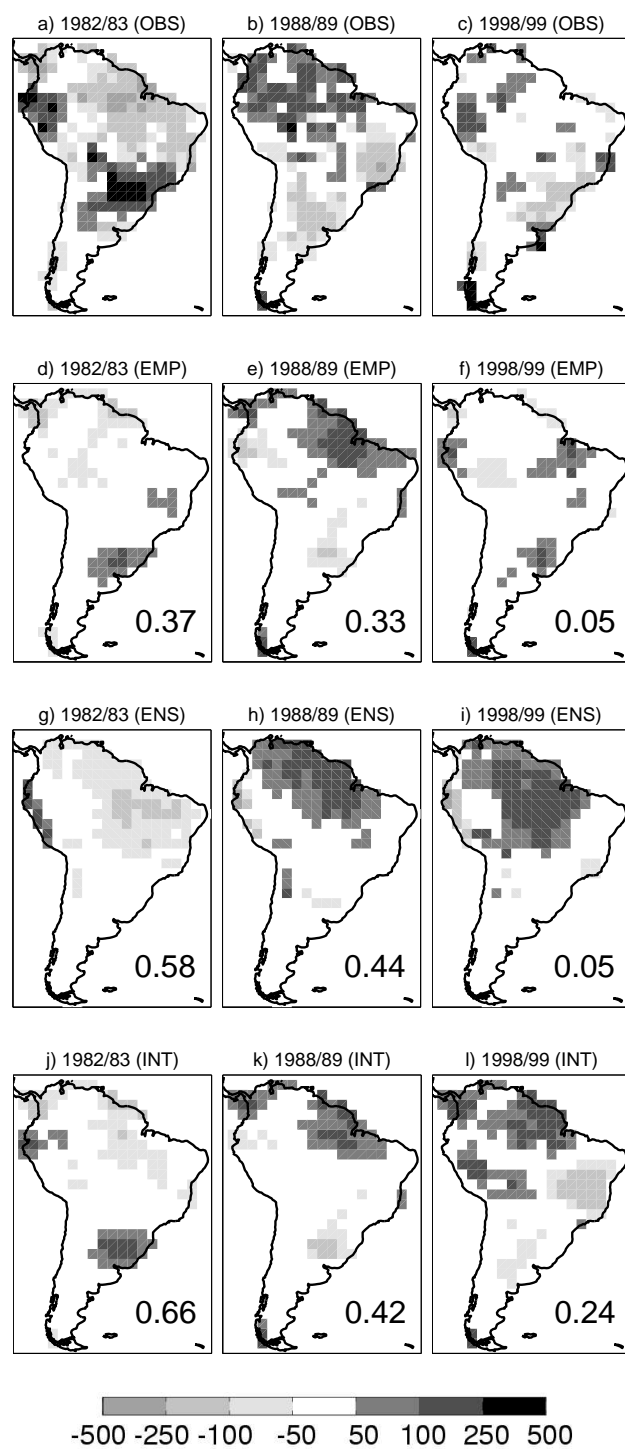


Figure 6:

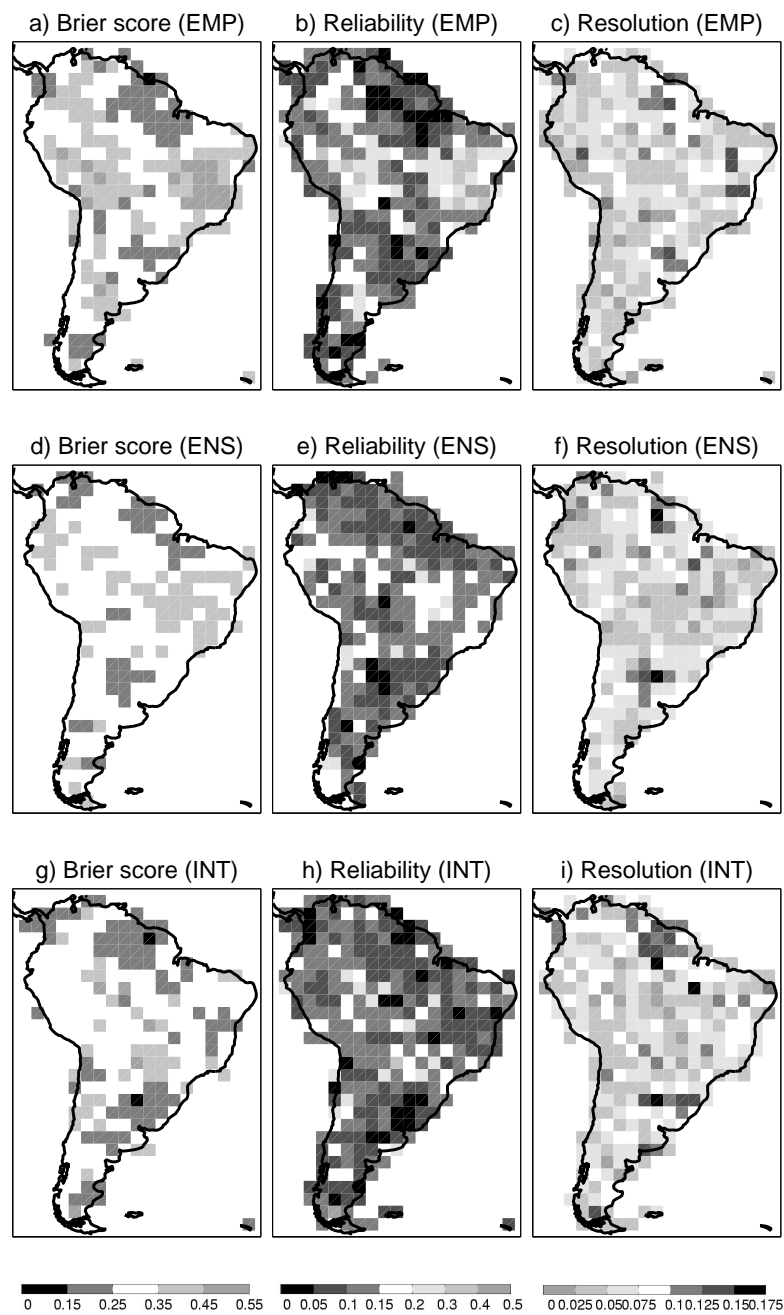


Figure 7:

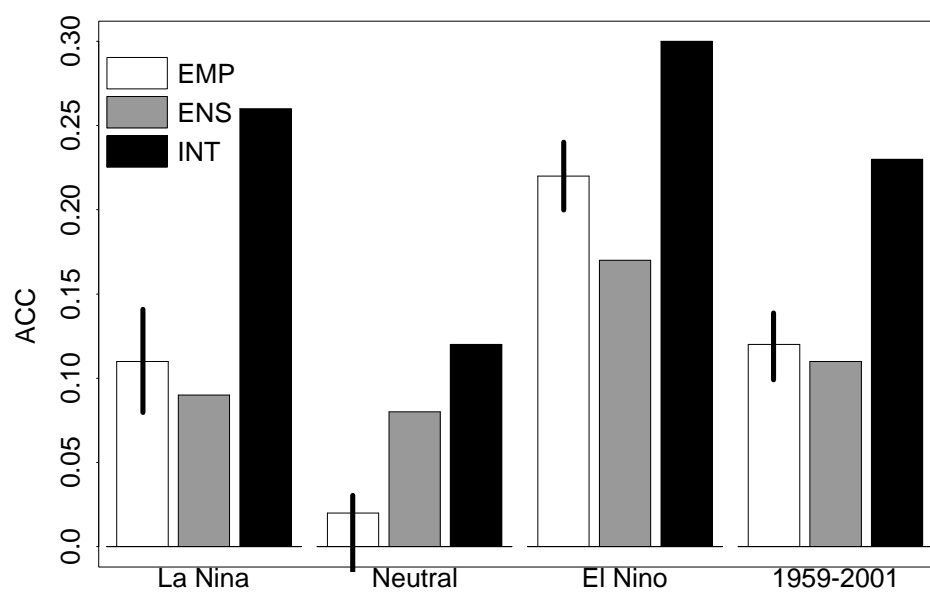


Figure 8:

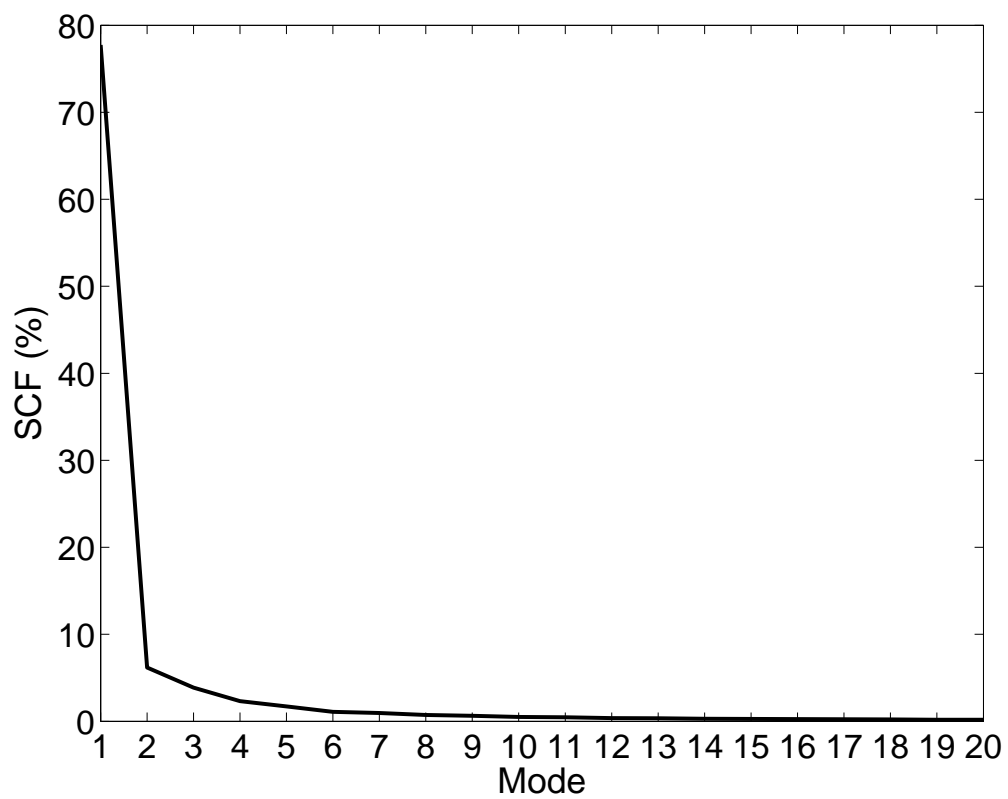


Figure 9:

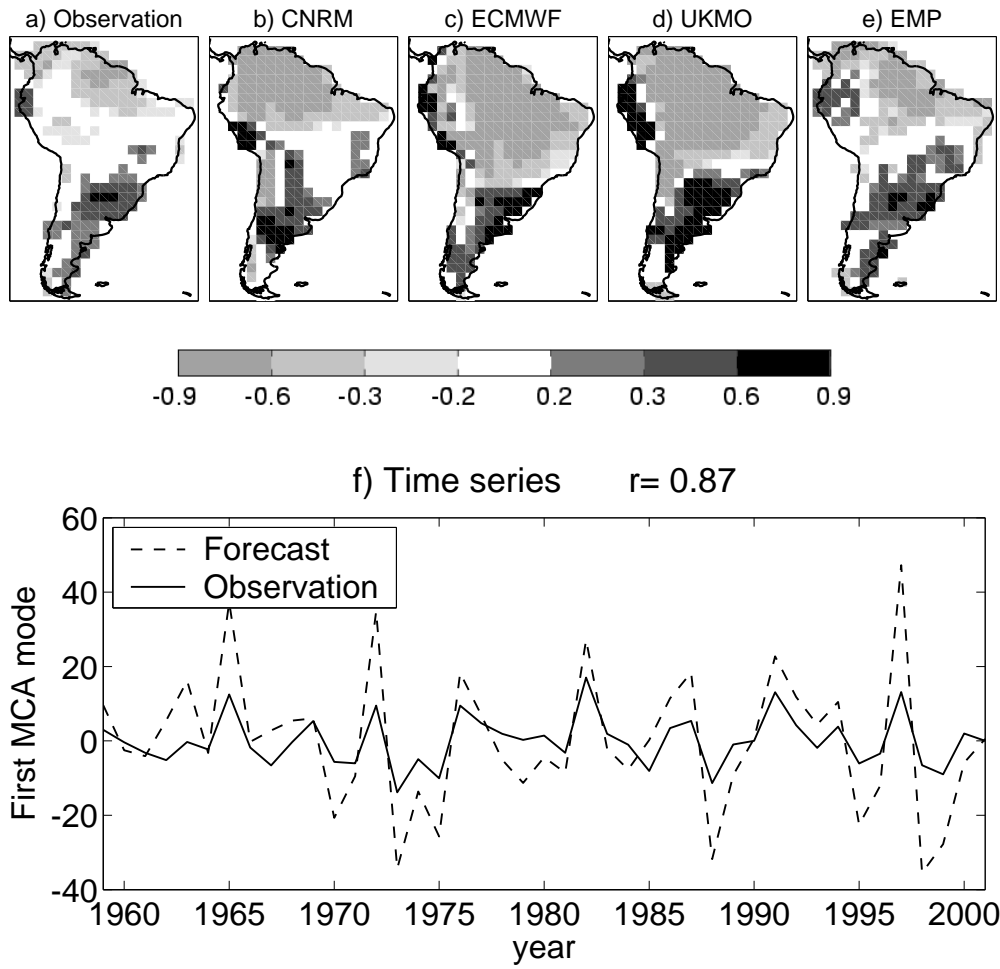


Figure 10: