6.13 IMPROVED SKILL OF ENSO COUPLED MODEL PROBABILITY FORECASTS BY BAYESIAN COMBINATION WITH EMPIRICAL FORECASTS

Caio A. S. Coelho*, S. Pezzulli, M. Balmaseda, F. J. Doblas-Reyes, and D. B. Stephenson
Department of Meteorology, University of Reading and ECMWF – Reading, U.K.

1. INTRODUCTION

ENSO forecasts are currently produced using either physically derived numerical climate models or using empirical (statistical) relationships based on historical data. The comparative skill of these two approaches is a subject of much debate. Given these two distinct approaches to forecasting, it is natural to ask whether combining them together may produce a forecast with more skill than either forecast considered separately.

Previous studies (e.g. Thompson 1977, Krishnamurti *et al.* 1999, Pavan and Doblas-Reyes 2000) have linearly combined different forecasts. Ad hoc optimal weights are determined to minimise the mean square error of the combined forecast. The idea of this study is to treat ENSO forecasts in probabilistic terms, with particular attention directed to the estimation of prediction uncertainty. A Bayesian methodology is used for this probabilistic combination (see Coelho *et al.* 2002).

This research was carried out using Reynolds *et al.* (2002) optimum interpolation version 2 SST climatological dataset (1950-2001) and 12 years (1987-1998) of ECMWF-DEMETER (Palmer and co-authors 2002) Niño-3 SST hindcasts. Only the ECMWF raw (i.e. not bias corrected) coupled model forecasts issued in August to forecast the next December (5-month lead time) have been used.

The idea is to use *all* available information in order to produce more skillful Niño-3 index probability forecasts. To avoid artificial skill, all results shown here were obtained using the cross-validation ``leave one out'' method.

2. RESULTS AND DISCUSSION

Figure 1 shows December Niño-3 index SST forecasts for the period 1987-1998 obtained from a) an empirical model defined as a linear regression between December Niño-3 historical values and the preceding July historical values, b) the ECMWF coupled model with initial conditions of July and c) a Bayesian combination of the two previous forecasts as described in Coelho *et al.* (2002). Forecasts are represented by the thick black line, the thin line shows observed values and the December climatological mean of 25ºC is represented by the dotted line. The 95% prediction interval (P.I.) is also shown (grey area

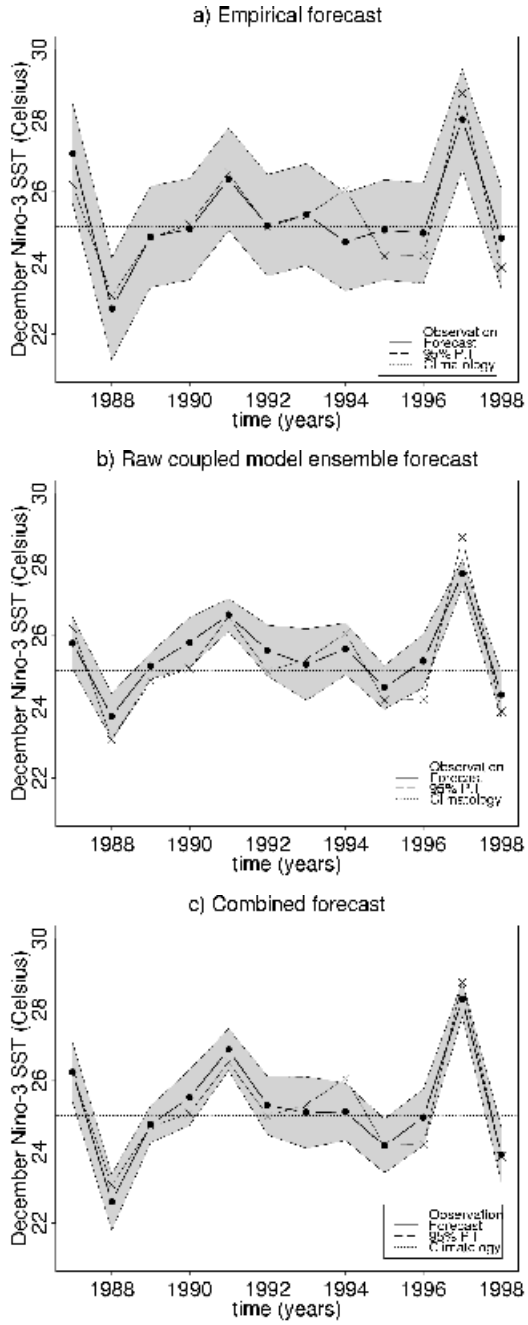

Figure 1. December 1987-1998 Niño-3 index forecasts (ºC). a) Empirical forecast, b) raw coupled model ensemble forecast and c) combined forecast.

surrounded by dashed lines). The 95% prediction interval is defined by the forecast value plus or minus 1.96 the forecast standard deviation. Good quality forecasts, that is, accurate and reliable

**Corresponding author address:* Caio Augusto dos Santos Coelho, Dept. of Meteorology, University of Reading. Earley Gate, PO Box 243, Reading RG6 6BB, U.K. Phone: +44 (0) 118 931 8954 Extn 4280, e.mail: c.a.d.s.coelho@reading.ac.uk

forecasts, should have 95% of observations falling inside this interval.

The empirical forecasts (Fig. 1a) are surprisingly accurate. Within the 12 forecasted years the model has only forecast the sign of the Niño-3 index incorrectly in 1994.

Figure 1b shows ECMWF raw coupled model ensemble forecasts for the same period. The ensemble system tends to underestimate the Niño-3 index and the size of the 95% P.I. is unrealistically much smaller than the size of the 95% P.I. of the empirical forecast. Many observations are out of the 95% P.I., indicating that the coupled model uncertainty is unrealistically underestimated, and is not able to cover the model forecast error limit.

The Bayesian method was then used to combine these two forecasts. Figure 1c shows the combined forecast. Comparison of this forecast with the empirical forecast (Fig. 1a) and the raw coupled model ensemble forecast (Fig. 1b) shows that the combined mean forecasts are in closer agreement with the observations. The 95% P.I.'s are also reduced compared to those of the empirical forecasts indicating less uncertainty due to combination with the coupled model forecasts. Unlike the coupled model forecasts, only one forecast (1994) falls outside the 95% P.I., indicating that the P.I. is more plausible.

Table 1 gives some deterministic cross-validated verification scores and a measure of forecast uncertainty for the forecasts of December Niño-3 index for the period 1987-1998 shown in Fig. 1. Also presented are scores for the climatological forecast obtained using the historical Niño-3 index December mean value of 25ºC and the historical December standard deviation of 1.23ºC.

Table 1: Verification scores, skill score and mean uncertainty. MSE and MAE are the mean squared error and mean absolute error, respectively. SS is the mean absolute error skill score defined as SS=1-MAE/MAE_c, where MAE_c is the climatological MAE. Values in brackets indicate the percentage improvement compared to the raw coupled model ensemble forecast skill score. Uncertainty is given by the mean predicted standard deviation of the forecasts within the period 1987-1998.

Forecast	MSE [ºC]^2	MAE [ºC]	SS [%]	Uncertainty [ºC]
Climatology	2.12	1.08	0	1.23
Empirical	0.44	0.50	53 (+2)	0.72
Raw ens.	0.37	0.53	51	0.33
Combined	0.20	0.35	68 (+17)	0.39

The combined forecast has the smallest values of Mean Squared Error (MSE) [0.20ºC^2] and Mean Absolute Error (MAE) [0.35ºC] of all the forecasts. It also shows an impressive improvement of 17% in skill when compared to the coupled model only forecasts. Additionally, it provides a much better and more realistic uncertainty estimate (0.39ºC) when compared with the other forecasts in the sense that it provided a reliable 95% prediction interval, in which only one year (1994) fell outside (Fig. 1c).

3. CONCLUSIONS

The empirical and the raw coupled model ensemble forecasts alone show similar skill score of around 50% (mean absolute error skill score relative to climatology). However, the combined forecast has a much better skill of 68% than the separate forecasts. The combined forecast also provides more accurate and realistic uncertainty estimate, which is unrealistically underestimated by the ensemble system. These results indicate that in order to obtain better quality probability forecasts of the Niño-3 index both empirical and coupled model ensemble forecasts should be combined together. Both empirical and raw coupled model ensemble forecasts contain mutually useful information. In other words, neither forecast is sufficient for the other forecast and so increased forecast skill can be obtained by combining both types of forecast.

4. ACKNOWLEDGMENTS

We wish to thank Dr. D. L. T. Anderson, head of the seasonal forecast group at ECMWF and Dr. T. N. Palmer, the DEMETER (EVK2-1999-00197) project principal investigator, 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.

5. REFERENCES

Coelho, C. A. S., Pezzulli, S., Balmaseda, M., Doblas-Reyes, F. J. and Stephenson, D. B., 2002: A Bayesian approach for combining empirical and coupled model ensemble predictions to give more skillful probability forecasts of ENSO. *J. Clim.,* Submitted**.**

Krishnamurti and co-authors, 1999: Improving tropical precipitation forecasts from a multianalysis superensemble. *Science*, **285**, 1548- 1550.

Palmer, T. N. and co-authors, 2002: Development of a European Ensemble System for Seasonal to Inter-annual Prediction (DEMETER). *Bull. Amer. Meteorol. Soc.*, Submitted.

Pavan, V. and Doblas-Reyes, F. J, 2000: Multi-model seasonal hindcasts over the Euro-Atlantic: skill scores and dynamic features. *Clim. Dyn.*,**16**, 616-625.

Reynolds, R. W. and co-authors, 2002: An improved in situ and satellite SST analysis for climate. *J. Clim.*, **15**(13), 1609-1625.

Thompson, P. D., 1977: How to improve accuracy by combining independent forecasts. *Mon. Wea. Rev.,* **105**, 228-229.