A Diagnostic Comparison of CFSv1 and CFSv2 Predictions of Nino3.4 SST

Anthony G. Barnston and Michael K. Tippett

Skills and other diagnoses are conducted to compare the relative performances of CFSv1 and CFSv2 in prediction of the ENSO-associated Nino3.4 SST index (5°N-5°S, 120°-170°W). The common 1982-2009 hindcast period is used, one-month averaging periods are used, and the 8 forecast lead times are examined.

CFSv2 slightly outperforms CFSv1, in terms of temporal correlation skill, for moderate to long lead times that traverse the northern spring ENSO predictability barrier, such as a forecast for July made in January or February. However, for forecasts during times of the year considered less challenging, such as a forecast for January made in late northern summer, CFSv1 has higher correlations than CFSv2. Inspection of the data reveals a moderate cold bias in CFSv2 forecasts for Nino3.4 SST during 1982-1998—a bias that suddenly becomes slightly positive during 1999-2009. Work by others has related this time-conditional bias to changes in the observing system in the late 1990s that affected the ocean reanalysis (CFSR) from which initial conditions for CFSv2 are taken. *A posteriori* treatment of these biases with two separate corrections improves the CFSv2 correlation scores such that they exceed those of CFSv1 at most times of year and at most lead times, often by substantial margins. The time-dependent bias in CFSv2 hides also its potential superiority over CFSv1 in terms of other basic performance measures such as RMSE and, to a lesser extent, standard deviation ratio. Such verifications are shown for CFSv1 against CFSv2 with and without the dual bias corrections.

CFSv2 is also seen to be better than CFSv1 in freedom from a recently introduced measure of forecast quality called target period “slippage”—a condition whereby skills are greater when verifying against a target period occurring earlier than the intended period. The presence of slippage cannot be diagnosed from the usual skill measures, which only compare forecasts with their verifying observations. Slippage, a systematic error, is indistinguishable from random error when forecasts at different leads are evaluated independently. Forecast systems that suffer from slippage tend to be late in predicting the initiation and termination of ENSO events. Slippage can be corrected in post-processing to the extent it is systematic. CFSv1 is found to benefit from such correction while CFSv2 is not found to require it.

A conclusion is that the analysis of the merits and weaknesses of CFSv2, and guidance for improvements in a future CFSv3 version, would be more easily achieved if problems in the CFSR data were first addressed and eliminated, allowing for diagnoses of CFSv2 using unbiased initial conditions.