



An Empirical Benchmark for Decadal Forecasts of Global Surface Temperature Anomalies

Matthew Newman

CIRES/University of Colorado and NOAA/ESRL/PSD, Boulder, United States (matt.newman@noaa.gov)

The suitability of an empirical multivariate red noise (AR1) model, or linear inverse model (LIM), as a benchmark for decadal surface temperature forecast skill is demonstrated. Constructed from the observed simultaneous and one-year lag covariability statistics of annually-averaged sea surface temperature (SST) and surface (2m) land temperature global anomalies during 1901-2009, the LIM has hindcast skill for leads 2-5 and 6-9 years comparable to and sometimes even better than skill of the CMIP5 model hindcasts initialized annually over the period 1960-2000, and has skill far better than damped persistence (e.g., a local univariate AR1 process). Over the entire post-1901 record, the LIM skill pattern is similar but has reduced amplitude. Pronounced similarity in geographical variations of skill between LIM and CMIP5 hindcasts suggests similarity in their sources of skill as well, supporting additional evaluation of LIM predictability. For forecast leads above 1-2 years, LIM skill almost entirely results from three non-orthogonal patterns, one corresponding to the secular trend and two more, each with about ten year decorrelation time scales but no trend, that represent most of the predictable portions of the Atlantic Multidecadal Oscillation (AMO) and Pacific Decadal Oscillation (PDO) indices, respectively. As found in previous studies, the AMO-related pattern also contributes to multidecadal variations in global mean temperature, and the PDO-related pattern has maximum amplitude in the west Pacific and represents the residual after both interannual and decadal ENSO variability are removed from the PDO time series. These results suggest that current coupled model decadal forecasts may not yet have much skill beyond that captured by multivariate red noise.