CMIP5 multi-model hindcasts for the mid-1970s shift and early 2000s hiatus and predictions for 2016–2035

Gerald A. Meehl1 and Haiyan Teng1

1National Center for Atmospheric Research, Boulder, Colorado, USA

Abstract Compared to uninitialized climate change projections, a multi-model ensemble from the CMIP5 10 year decadal prediction experiments produces more warming during the mid-1970s climate shift and less warming in the early 2000s hiatus in both the tropical Indo-Pacific region and globally averaged surface air temperature (TAS) in closer agreement with observations. Assuming bias in TAS has stabilized in the 10 year predictions, after bias adjustment, TAS anomalies for the 2016–2035 period in the 30 year predictions initialized in 2006 are about 16% less than the uninitialized projections. One contributing factor for the improved climate simulation is the bias adjustment, which corrects the models’ systematic errors and higher-than-observed decadal warming trend. Another important factor is the initialization with observations which constrains the ocean such that the starting points of the initialized simulations are close to the observed initial states.

1. Introduction

While skill in the Coupled Model Intercomparison Project phase 5 (CMIP5) near-term climate predictions is, to a large extent, attributed to changes in atmospheric composition [Doblas-Reyes et al., 2013], ocean initialization (initialization here and elsewhere refers to model simulations initialized with observations and bias adjusted to account for model error; further details are given in the supporting information) adds skill to surface temperature predictions in regions such as the North Atlantic subpolar region [Robson et al., 2012; Yeager et al., 2012; Kirtman et al., 2013]. The tropical Pacific generally has lower initial value predictability [Branstator and Teng, 2010] and skill scores [Doblas-Reyes et al., 2013; Kirtman et al., 2013]. But a previous study [based on one CMIP5 model with two different initialization methods, Meehl and Teng, 2012] showed that the hindcasts produced a more realistic mid-1970s climate shift as well as the early 2000s hiatus compared to the uninitialized climate projection experiments. The mid-1970s climate shift [Trenberth and Hurrell, 1994; Meehl et al., 2009] refers to the relatively rapid warming over a decade or so in the tropical Pacific. It is associated with a warming response to increasing greenhouse gases (GHGs) and an internally generated transition of the Interdecadal Pacific Oscillation (IPO) [Oscillation, Mochizuki et al., 2010] from negative to positive [Meehl et al., 2009; Meehl and Arblaster, 2011, 2012]. The early 2000s hiatus refers to the decade-long global warming plateau [Easterling and Wehner, 2009], which likely has had contributions from persistent tropical Pacific cooling, resulting from natural variability associated with the negative phase of the IPO partially offsetting warming from increasing greenhouse gases [Meehl et al., 2011, 2013; Guemas et al., 2013; Kosaka and Xie, 2013; England et al., 2014]. Additionally, aerosols from a collection of small to medium-sized volcanic eruptions [Santer et al., 2014] along with solar forcing variations could also have contributed [Fiato et al., 2013].

We revisit these two climate events with an analysis of the multi-model data set from CMIP5 [Taylor et al., 2012]. The motivation is to examine whether the initialized runs can improve simulations of the mid-1970s shift and the early 2000s hiatus compared with the uninitialized. The second motivation is to assess whether the second to the third decades of the 30 year predictions from the few available start dates add any information about the relative warming during the canonical near-term period assessed in the IPCC AR5 of 2016–2035 compared to the uninitialized projections.

2. Models, Experiments, and Analysis

We use 16 models for analysis of the 10 year hindcasts and 14 models for the 30 year hindcasts/predictions (Table 1 in the supporting information). A hindcast proceeds from imposed initial conditions for a period of either
10 or 30 years as described in the CMIP5 experiment design [Taylor et al., 2012]. The initialized simulations are then compared to the corresponding uninitialized CMIP5 simulations (including the historical runs during 1960–2005 and RCP4.5 during 2006–2035) from those same models.

Other factors may have contributed to the early-2000s hiatus. For example, there may be climate effects from a collection of moderately sized volcanoes noted above that were not included in the early-2000s period of RCP4.5, but there is evidence that post-2005 sulfate aerosol concentrations are comparable to those in RCP4.5 [Klimont et al., 2013]. However, a caveat that must accompany any analysis of the post-2005 period is that the actual observed external forcing may not be exactly what was included in the RCP4.5 simulations.

The variable we present here is annual mean surface air temperature (TAS). In order to remove model systematic errors, we bias adjust the output using the National Centers for Environmental Prediction/National Center for Atmospheric Research reanalysis as the observations [Kalnay et al., 1996; use of other observational data sets produces qualitatively similar results]. For both the observations and model outputs, annual mean TAS is calculated from the monthly data and then interpolated to the spatial resolution of a T42 grid, which corresponds to approximately a 2.8° latitude/longitude grid spacing. At each grid point, climatological annual mean differences for the 10 year hindcasts are computed with regard to the reanalysis data and composited by year following the start year of all the hindcast simulations. Then, these averaged time-evolving annual mean biases are subtracted using cross-validation from each hindcast/prediction [CLIVAR, 2011]. Though this method may not be optimum for some models (see more detailed description of the bias adjustment technique in the supporting information), it is chosen here because it is more feasible for multi-model analysis.

The full field bias adjustment method may lead to an artificially enhanced skill by removing time-dependent climatological errors. In order to separate the effect of bias adjustment from that of initialization, we apply the same bias adjustment procedure to outputs from the uninitialized runs (the results are referred to as “bias-adjusted uninitialized”). The difference between the initialized and the bias-adjusted uninitialized corresponds to the effect of initialization.

With only three 30 year hindcast/prediction start years in the CMIP5 experimental design, there are insufficient samples to construct a time-evolving systematic error between years 11 and 30. We assume, as in Meehl and Teng [2012], that from years 11 to year 30 the systematic error is nearly stationary and can be approximated as a climatological drift adjustment by the systematic error in year 10 (which is supported by the fact that grid point values of annual mean TAS drift on the order of 1 to 2°C early in the hindcast period, but the change in the drift is generally less than a couple tenths of a °C from years 9 to 10, Figure S1, supporting information). A significant caveat that accompanies this assumption is that the bias adjustment calculations may introduce further errors when comparing model results to observations [Goddard et al., 2013; Khairin et al., 2012].
3. Results

As in Meehl and Teng [2012], we compare runs starting from 1976 (except for CCSM4 that started from year 1975) and 2005 (except for FGOALS-g2, IPSL-CM5A-LR, MIROC4h, and MRI-CGCM3 that started from 2006) with the uninitialized during the same period. We compute each model’s ensemble mean first, followed by the multi-model ensemble average. Figure 1a shows time series of globally averaged TAS anomalies relative to years 1986–2005 for observations and for the two case study periods. For the mid-1970s shift, the anomaly during the 1978–1982 prediction period (5 year average of the prediction for years 3–7) from the initialized is 0.05°C and 0.07°C warmer than the uninitialized and the bias-adjusted uninitialized, respectively. For the 2000s hiatus, the anomaly for years 2007–2011 (prediction for years 3–7) from the initialized is 0.09°C and 0.08°C cooler than the uninitialized and the bias-adjusted uninitialized, respectively. All these differences are significant at the 95% level computed from a two-sided Student’s t test. For both events, the initialized runs outperform both the uninitialized and the bias-adjusted uninitialized, suggesting that initialization is responsible for the improvement.

However, the impact of initialization on globally averaged TAS does not exceed the 10 year prediction range. This conclusion is supported by two 30 year predictions starting from years 1981 and 2006 (Figure 1b). For

**Figure 2.** Spatial distribution of 5-year mean (years 3–7 of the prediction) surface air temperature changes (°C) for (left) the mid-1970s shift and (right) the early 2000s hiatus period relative to a prior 15 year climatology, from (top) the reanalysis, (second row) CMIP5 16-model ensemble mean 10 year hindcasts, (third row) uninitialized simulations, and (bottom) bias-adjusted uninitialized simulations (BA-uninitialized). The 10 year hindcast runs start from 1976 (except for CCSM4 from 1975) and 2005 (except for FGOALS-g2, IPSL-CM5A-LR, MIROC4h, and MRI-CGCM3 from 2006). The 1961–1975 and 1990–2004 climatologies are calculated based on observations for Figures 2a, 2b, 2d, 2e, 2f, and 2h and the uninitialized for Figures 2c and 2g. The red/blue stippling indicates where the multi-model ensemble mean is ±2 standard deviations warmer/colder than the observations as in Smith et al. [2012]. The two numbers at the lower left corner of the panels indicate pattern correlation and root mean square error (RMSE), respectively, compared with the observations for the domain displayed, and the number at the top right corner corresponds to globally averaged TAS anomaly.
both start years, annual mean anomalies from the initialized become indistinguishable from those of the bias-adjusted uninitialized after about 5 years. Although throughout the 30 year prediction period anomalies from the initialized are significantly lower than the uninitialized, the reduced warming after year 10 is mainly caused by the posteriori bias adjustment instead of initialization.

For both climate shift events, the uninitialized simulations produce rather spatially uniform warming over most of the tropical Indo-Pacific ocean area (50°E–90°W, 45°S–45°N). Five-year (corresponding to years 3–7 of the predictions) averaged TAS anomalies relative to a prior 15 year climatology from the uninitialized are presented in Figures 2c and 2g. For comparison, the corresponding observations are shown in Figures 2a and 2e. The two numbers at the lower left corner of the panels indicate pattern correlation and root mean square error (RMSE), respectively, compared with the observations for the domain displayed, and the number at the top right corner corresponds to globally averaged TAS anomaly. While the RMSE values are similar across all panels in Figure 2, the pattern correlation coefficients from both the initialized and bias-adjusted uninitialized are larger than the uninitialized, with the initialized values being the largest (larger pattern correlations indicate better simulations). In order to test whether the pattern correlation coefficient is significant compared to those from random patterns, we carried out a Monte Carlo test by randomly picking 10,000 detrended TAS annual means from the uninitialized runs during the period of 1960–2006 from each of the CMIP5 decadal prediction models. Then, the pattern correlations of the multi-model ensemble mean anomalies among all random pairs are calculated. The 99th percentile of the pattern correlations from 10,000 random draws is 0.52. Based on this test, the uninitialized runs do not have a significant pattern correlation compared to the observations for both events (Figures 2c and 2g), while both the initialized simulations (Figures 2b and 2h) and the bias-adjusted uninitialized runs (Figures 2d and 2h) have statistically significant pattern correlation values for both events.

The benefit of bias adjustment is also evident in the 30 year predictions. A comparison of TAS anomalies during 1990–2009 (years 10–29 predictions) relative to the period of 1960–1979 from the observations, initialized, uninitialized, and bias-adjusted uninitialized shows a positive IPO pattern in the Indo-Pacific ocean domain in observations, but the uninitialized is dominated by rather uniform warming (Figure S2 left). Patterns from the initialized and the bias-adjusted uninitialized are similar, and both resemble the observed pattern. We repeat the analysis with the 30 year predictions starting from year 2006 (Figure S2, right, no observations to validate). Again, the TAS anomaly pattern for years 10–29 of the initialized is close to the bias-adjusted uninitialized. Thus, for the 20–30 year prediction range, improvement in the initialized runs is mainly due to bias adjustment.

As can be seen in Figure 1b, bias adjustment, along with initialization during the cooler hiatus period, produces global warming for the 2016–2035 period compared to the 1986–2005 reference period of +0.62 ± 0.33°C (2σ), about 16% less than +0.74 ± 0.24°C for the uninitialized model projections, and close to +0.65 ± 0.27°C for the bias-adjusted uninitialized projections. This reduction in the magnitude of global warming in initialized predictions compared to uninitialized projections is qualitatively consistent with initialized predictions from a smaller set of models starting in 2011 [Smith et al., 2012].

In addition to the benefit of constraining the ocean initial state to the observations such that the model starting point is close to the observations, the reduced warming during the early 2000s hiatus period is partly caused by the posterior bias adjustment. In Figure 3, we compare the probability distribution function (PDF) of decadal warming trends in globally averaged TAS from the observations and the CMIP5 models. The decadal trend starting from year 1960 to 2002 in the observations averaged over all values is 0.15°C/decade,
compared to 0.21°C/decade from the uninitialized climate projections. This is consistent with previous studies showing that the model-simulated global warming trend is somewhat higher than the observed [Allen et al., 2000; Stott and Kettleborough, 2002; Stott and Jones, 2012; Smith et al., 2012]. In contrast, after bias adjustment, the averaged decadal warming trend is 0.03°C/decade and 0.05°C/decade from the initialized and bias-adjusted uninitialized, respectively, averaged over all values. Both trends are lower than the observed trend, suggesting that the bias adjustment may over-correct the decadal warming trend. In addition to these averages computed over all values, bias adjustment shifts the maxima of the trends to smaller values closer to the observations in Figure 3, with the biggest improvement occurring in the initialized.

Initialization with observations is responsible for improvements in the 10 year hindcasts/predictions of both climate shift events over and above bias adjustment, especially when globally averaged TAS is concerned (Figures 1a and 2). The initialized runs tend to produce warm and cold anomalies associated with the mid-1970s shift and the early 2000s hiatus, respectively, relative to the bias-adjusted uninitialized, and this is not limited to the two start years that we present in Figure 1a. Similar results are obtained when the analysis is extended to the adjacent start years.

In order to understand why the initialized runs tend to produce more warming during the mid-1970s and less warming during the early 2000s relative to the uninitialized, we treat the initialized run as a perturbed initial value problem with respect to the uninitialized runs (supporting information). The impact of initialization on the year 1 TAS forecast averaged over start years 1975–1980 and 2000–2005 across all models is shown in Figure S3. During the first period, the net effect of initialization is to produce a positive IPO-like pattern compared to the uninitialized. During the second period, the anomalies over the ocean exhibit a similar geographic pattern as in the first period but with the sign approximately reversed. The globally averaged TAS anomalies during the two periods are 0.04°C and −0.10°C, respectively, which are about the same order as the difference between the initialized and bias-adjusted uninitialized as shown in Figures 1a and 2. Because the uninitialized runs tend to produce a higher-than-observed long-term trend, by constraining the initial conditions with observations, the magnitude of both negative anomalies during the early decades and positive anomalies during the later decades relative to the long-term climatology is reduced. This is equivalent to introducing warm (cold) anomalies to the uninitialized both early and late in the 1960–2011 study period.

4. Summary

The CMIP5 multi-model near-term hindcasts/predictions produce a more realistic mid-1970s climate shift and early 2000s hiatus. The improvement is partly attributed to the posterior bias adjustment, which reduces the simulated higher-than-observed decadal warming trend. This is especially true for the 20–30 year range of the 30 year hindcasts/predictions. As a partial consequence, for the 30 year prediction starting from year 2006, there is a reduction of 16% to global warming for the 2016–2035 period compared to the uninitialized projections.

In addition, initialization also is an important factor in producing more realistic simulations relative to the uninitialized runs. Its net effect (after bias adjustment) is to introduce warm and cold TAS anomalies to the uninitialized for the two cases, respectively. This effect is caused by the constraint of the initial conditions rather than posterior bias adjustment. Further analyses are necessary to study the ocean processes that may possibly contribute to the predictions of the two climate shift events and whether the cooling during the early 2000s hiatus is associated with increased deep ocean heat uptake as postulated by previous studies. The biggest challenge for improving decadal hindcasts and predictions is to reduce model error. This will alleviate the need for the large bias adjustments necessary for the current generation of models.

References


