Scrutinizing Forced and Unforced Variability in CMIP5

Timothy DelSole, Xiaoqin Yan, Emerson LaJoie, and Laurie Trenary

George Mason University and Center for Ocean-Land-Atmosphere Studies, Fairfax, VA

Accuracy of Model Projections During the Hiatus Period

Data:
Surface air temperature ('tas') of the HadCRU4 dataset and pre-industrial control/ historical/rcp45 simulations of 14 models from CMIP5 project are interpolated onto common grid of 5°x5°. Mask is constructed based on data availability of HadCRU4.

Hypothesis:

ΔO	=	F	Δa	+	Δu
(10-yr mean obs)		(Response pattern)	(amplitude)		(internal variability)

Detection: $|\Delta\hat{\mathbf{a}}| \geq \sqrt{\frac{4}{3}}\sqrt{\left(\mathbf{F}^T\hat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{F}\right)^{-1}}\, t_{N_U-1}(\alpha/2)$

Attribution: $\Delta\hat{\mathbf{a}}_{model} \pm \sqrt{\frac{4}{3}(1+1/E)}\sqrt{\left(\mathbf{F}^T\hat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{F}\right)^{-1}}\, t_{N_U-1}(\alpha/2)$

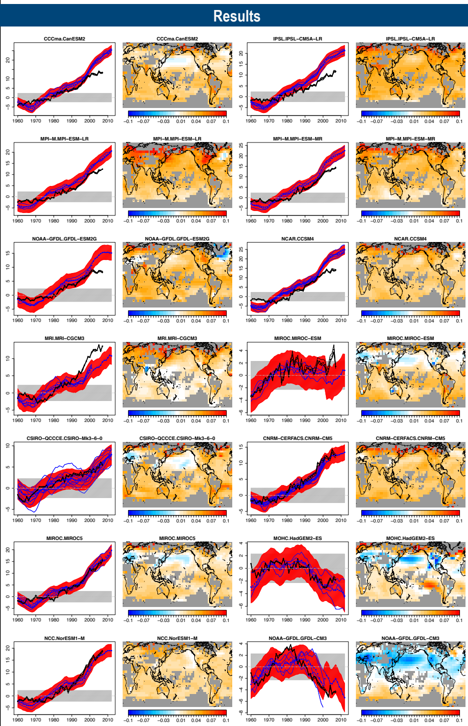
Δ denotes differences relative to climatology of 1961-1990.
E and α are ensemble size and 5% significance level.
$\hat{\boldsymbol{\Sigma}}_U$ denotes covariance matrix of internal variability.

Comments:
1) The forced response pattern of each model was obtained by applying discriminate analysis (Jia and DelSole, 2012) based on 12-eof truncations to each model's 10-year running mean 'tas' of the 500-year pre-industrial control run and the historical runs.
2) Detection/attribution analysis was applied to tas anomalies of both HadCRU4 and each model relative to the 1961-1990 climatology. RCP4.5 used to extend historical runs.
3) The confidence interval accounts for the fact that the climatology of 1961-1990 has 1/3 the variance of a 10-year mean.
4) Years shown on the figure of curves are the center year of corresponding 10-year mean (e.g.,1995 for mean of 1991-2000).
5) Consistency is achieved if observations are within the 95% confidence interval (red shading band).

Conclusions:
8 model simulations (blue) are inconsistent with HadCRU4 observations (black), 6 model simulations (blue) are consistent with HadCRU4 observations (black). The dominant forced response patterns differ significantly among some models.

References:
Jia, L. and T. DelSole, 2012: Optimal determination of time-varying climate change signals. *J. Climate,* **25**, 7122–7137.

Results



Changes in Internal Variability Due to 20th Century Climate Changes

This study investigates the possibility that internal variability responds to anthropogenic forcing on annual time scales. Current detection and attribution methods assume that internal variability does NOT change; our study tests if this assumption is valid. Previous studies assessed this question using univariate metrics such as the "residual consistency check" (Allen and Tett,1999). Unfortunately, this metric tests consistency in an aggregate sense, based on a suitably weighted total sum variability, and as such it allows errors in some components to be compensated by errors in other components, giving a potentially misleading impression of overall consistency. We examine consistency using more powerful discriminant analysis (DA) techniques, which effectively tests equality of the entire covariance matrix of variability.

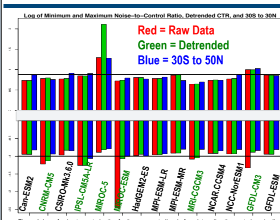


Figure 1. Log of noise-to-control ratios for three case studies; before detrending the control run (red), after detrending the control run (green), and masking out the poles to analyze only the domain from 30S to 50N (blue). Black lines indicate significance at the 1% level. We find that most of the differences in internal availability happen in polar regions.

We apply DA to annual mean, near-surface (2m) temperature from 14 CMIP5 climate models that simulated at least 500 pre-industrial control years and had a 3-member historical ensemble. The data was interpolated onto a common 5x5 grid and climatology was removed. DA finds the linear combination of variables that maximize (and minimize) the ratio of variances. The resulting ratio is called the **noise-to-control ratio** and given by

$$\lambda = \frac{\mathbf{q}^T\Sigma^{20C}\mathbf{q}}{\mathbf{q}^T\Sigma^{ctr}\mathbf{q}}$$

Internal variability in the 20th century is estimated by subtracting out the ensemble mean. Optimization is performed on the leading 30 EOFs.

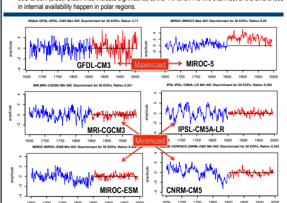


Figure 2. Variates for the optimized noise-to-control ratios. Control period is in blue. One ensemble member from the historical period is in red. For the maximized cases notice the variance appears to increase in the 20th century. Conversely, in the minimized cases the variance decreases.

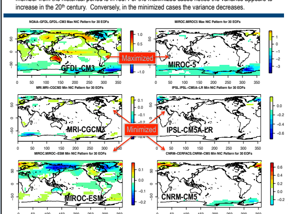


Figure 3. The spatial pattern that corresponds to the variates in Fig. 2.

The upper panel in Fig. 1 shows the log of the maximized ratios for three separate applications of DA and the horizontal black line provides the 99% confidence level. In this figure, a significant *maximized* ratio means that an INCREASE in 20th century internal variability was detected. Similarly, the lower panel shows the results for the log of the minimized ratios and the horizontal black line there indicates the 1% confidence level. Bars that stay between these two black lines indicate no changes in 20th century internal variability are detected. We found that 2 models had significant maximized ratios and 4 models had significant minimized ratios. However, we noted trends in the control runs, so, we detrended all the control runs and repeated our analysis. The green bars indicate detrended data; detrending substantially reduced the change in variability for two models.

We find that CNRM, IPSL-LR, MIROC-ESM, and MRI experienced a decrease in internal variability, while MIROC5 and GFDL-CM3 experienced increases in internal variability. The time series for the significant components are shown in fig. 2, while the corresponding spatial patterns are shown in fig. 3. These results suggest the following conclusions:

- GFDL-CM3 experiences an increase in internal variability that is widespread throughout the globe.
- The decrease in 20th century internal variability tends to be concentrated in regions of sea ice. We conjecture that 20th century losses in sea ice reduce annual mean variability because, without sea ice, the overlying atmosphere is exposed to the ocean's surface for longer durations of the year and the ocean's surface temperature would have a moderating effect on the near-surface temperature of the atmosphere.
- MRI and IPSL-LR possess a low frequency oscillation in their respective control runs that does not persist in the historical period.
- After masking out the polar regions outside 30S and 50N (see blue bars in fig. 1), we find that only MIROC5 yields a significant change in variability. However, the change in MIROC5 appears to be an artifact of a change in mean in the historical runs.
- Some models breach the significance level only marginally, so we hesitate to identify them as exhibiting changes in internal variability because of the subjective nature of truncating EOFs.

Summary

Our results show that some models exhibit a change in internal variability of annual mean near surface temperature in response to human influences. Much of this response is in polar regions and a plausible consequence of the loss of sea ice. In addition, changes in internal variability impact the assumptions underlying detection and attribution studies. We plan to continue our analysis to better address the sensitivity of our results to the choice of EOF truncation and the problem of overfitting when applying discriminant analysis. We are also working on applying our analysis to the 21st century to determine if changes in internal variability continue in the future.

Relation Between AMO and AMOC

MOTIVATION: Research suggests that the Atlantic Multi-decadal Oscillation (AMO) is the most predictable mode of temperature variability. This predictability is often attributed to variations in the Atlantic Meridional Overturning Circulation (AMOC). Here we ask the following questions:

1. Is there a relation between the AMOC and AMO?
2. Is the relation consistent across climate models?
3. Is the maximum streamfunction the best variable for assessing the AMOC-AMO relation?

Methodology

DATA: Analysis is performed on 450 years of annual anomalies of sea surface temperature (SST) and Atlantic mass overturning circulation (AMOC) data from the following pre-industrial control runs:

CMIP5 I.D.	Label	Atmos. Model (lon x lat)	Ocean Model(lat x depth)
CanESM2	CCC	CanAM4 (128x64)	OGCM4 (192x41)
MPI-ESM-MR	MPI	ECHAM6 (192x96)	MPIOM (180x41)
MRI-CGCM3	MRI	MRI-AGCM3.3 (320x160)	MRI.COM3 (154x52)
CCSM4	NCAR	CAM4 (288x192)	POP2 (395x61)
NorESM1-M	NCC	CAM4-Oslo (144x96)	NorESM-Ocean (166x70)

An index for the Atlantic Multi-decadal Oscillation (AMO) is defined as the area-weighted annual SSTA over the north Atlantic from 0-60°N. The maximum strength of the AMOC is defined as the maximum in the AMOC between 30°N-60°N, and denoted as Ψ_{max}. Control runs are split into two parts: a training part (225 years) for building the model and a verification part (225 years) for testing the model.

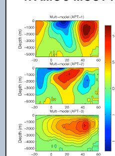
OPTIMIZATION: The relation between the AMO and AMOC is diagnosed by finding the linear combination of AMOC principal components (PCs) that are most highly correlated with AMO in an integral sense (Jia and Del- Sole, 2011). This integral is called Average Predictability Time (APT):

$$APT = \max_{\beta_1,\ldots,\beta_M} \int_{-\infty}^{\infty} \mathrm{cor}^2\left[AMO, \beta_1\psi_1 + \cdots + \beta_M\psi_M\right] d\tau$$

ψ_i i'th principal component
τ time lag

Results

1. AMOC MOST RELATED TO AMO



- Only ***3*** APT are found to be significant across all models
- No robust relation was found on decadal-to-multidecadal time scales

Figure 1: The leading three components of the AMOC that are most related to the AMO, in a multi-model sense. Structures are for the training period.

3. AMOC and Ψ_{max}

- APT patterns have larger correlation with AMO compared to Ψ_{max}
- Regressing out APT from Ψ_{max} removes significant correlations with AMO

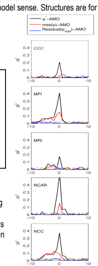


Figure 3: Squared correlation for 3 leading APT-AMO (black), ψ_{max}-AMO (red), and residual ψ_{max} -AMO (blue). Residual ψ_{max} is recovered after the 3 leading APT has been regressed out of ψ_{max}. Negative lags indicate APT and ψ_{max} lead AMO.

2. MODEL COMPARISION

Differences in cross model AMOC-AMO relation are quantified by fitting an empirical model:

$$AMO(t) = \beta_0 AMO(t-1) + \beta_1\psi_1^* + \beta_2\psi_2^* + \epsilon,$$

where $\psi_i^* = ith$ APT timeseries, β = regression coefficients, and ϵ = noise.

An F-test is used to test equality of all models

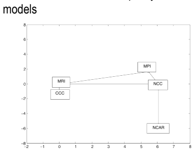


Figure 2: Diagram showing the clustering of the empirical models for 5 CMIP5 climate models. Clustering is based on F-Statistic for pairwise comparison of empirical models. Lines connecting models denote statistically equivalent models at α=0.05. The distance between unconnected models is proportional to F.

- AMOC-AMO relations are statistically indistinguishable for MRI/MPI/NCC
- CCC and NCAR are significantly different from all other models except MRI and NCC

Future Work

- Extend optimization to include sea level pressure over the North Atlantic
- Construct simple dynamical model to describe AMO-AMOC mechanism
- Attribute differences in AMO variability to differences in regression model and AMOC variability