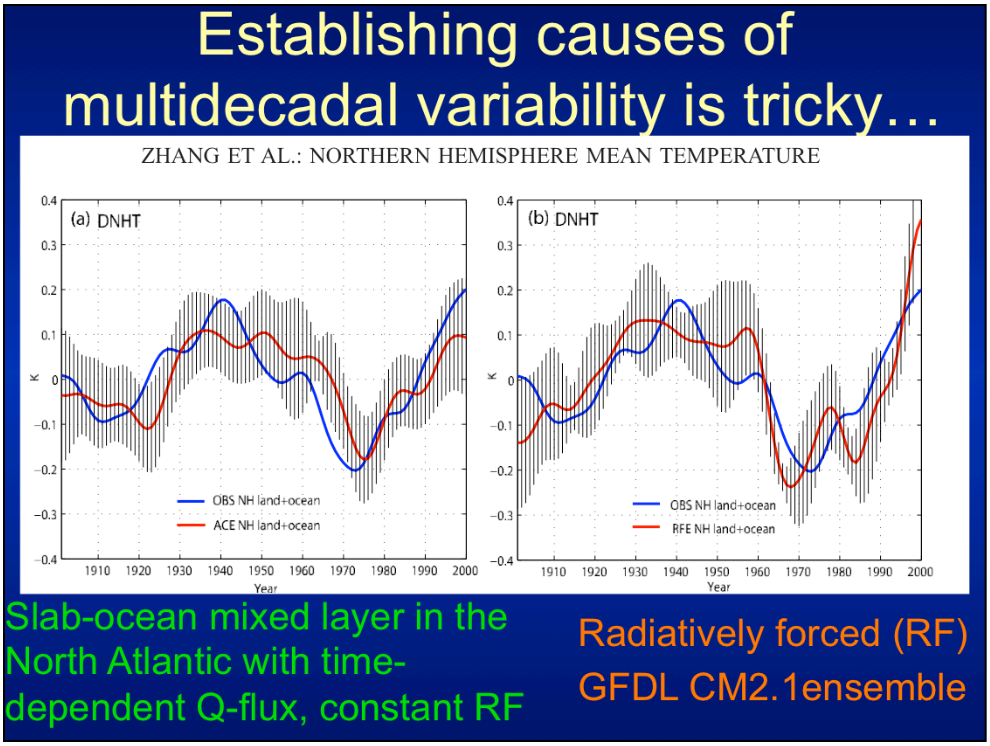
GFDL seminar
Princeton, NJ

October 20, 2016

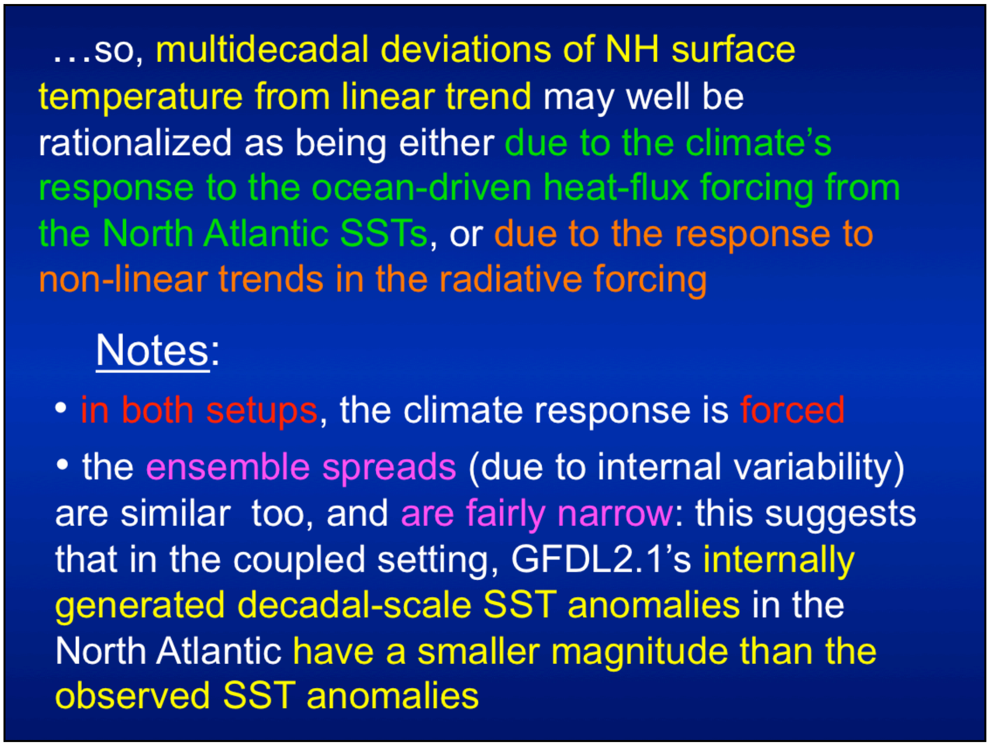
On semi-empirical decomposition of multidecadal climate variability into forced and internally generated components

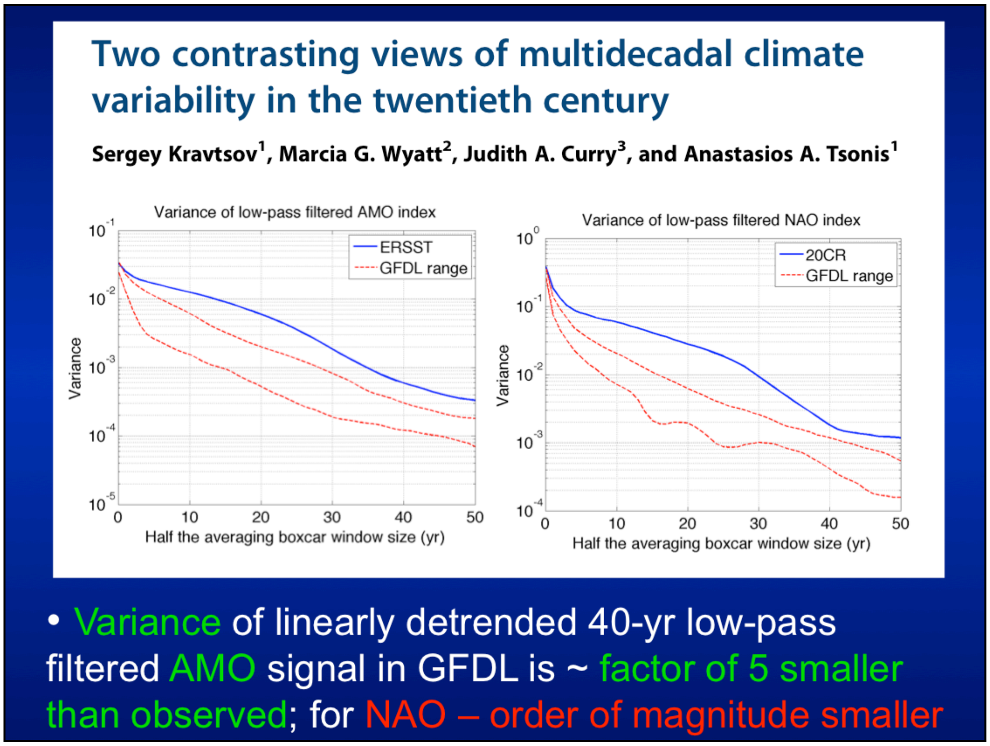
Sergey Kravtsov
University of Wisconsin-Milwaukee
Department of Mathematical Sciences
Atmospheric Science Group

This work was supported by the NSF grants ATM-1236620 and OCE-1243158.
https://people.uwm.edu/kravtsov/

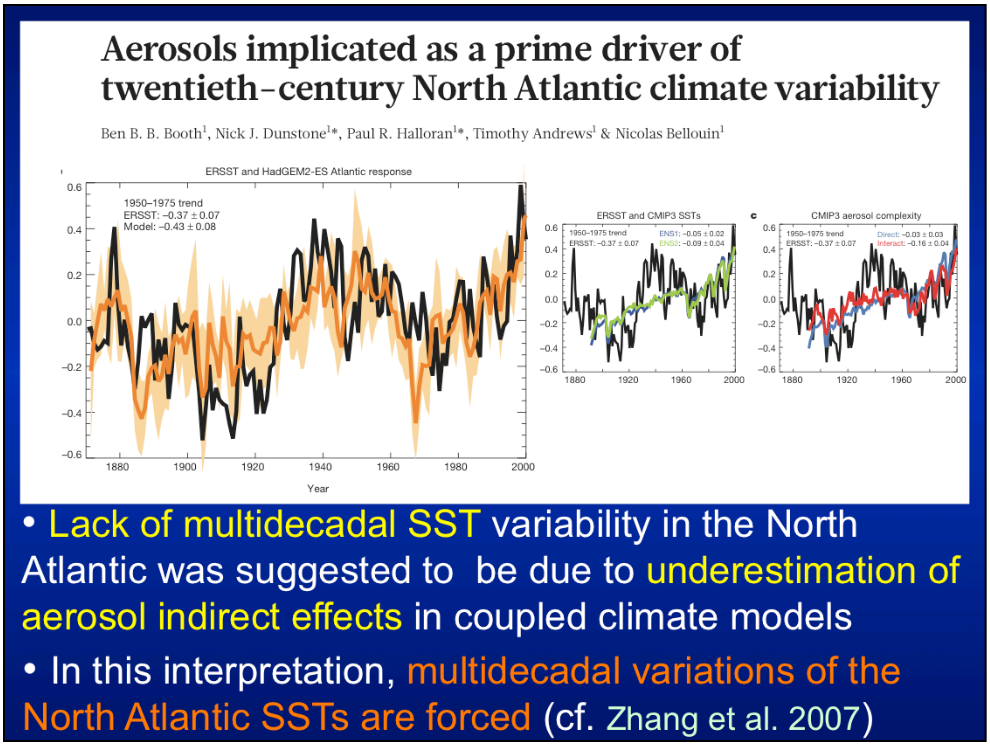


…since the internally generated SST anomalies (e.g. due to variations in AMOC) and non-uniform (in time) radiative forcing may both be responsible for the observed non-uniformities in the NH warming!


…so, multidecadal deviations of NH surface temperature from linear trend may well be rationalized as being either due to the climate's response to the ocean-driven heat-flux forcing from the North Atlantic SSTs, or due to the response to non-linear trends in the radiative forcing
Notes:
• in both setups, the climate response is forced
• the ensemble spreads (due to internal variability) are similar too, and are fairly narrow: this suggests that in the coupled setting, GFDL2.1's internally generated decadal-scale SST anomalies in the North Atlantic have a smaller magnitude than the observed SST anomalies



The multidecadal deviations from the linear trend in GFDL2.1 simulated North Atlantic SSTs are indeed much smaller than observed, but this is even more so for NAO (which, incidentally, has essentially no forced component in the CMIP5 model simulations). Note that the linearly detrended NHT variance is similar to the observed, consistent with Zhang et al. (2007).


Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability
Ben B. B. Booth[1], Nick J. Dunstone[1]*, Paul R. Halloran[1]*, Timothy Andrews[1] & Nicolas Bellouin[1]
ERSST and HadGEM2-ES Atlantic response
1950–1975 trend
ERSST: −0.37 ± 0.07
Model: −0.43 ± 0.08
Year
ERSST and CMIP3 SSTs
ENS1: −0.05 ± 0.02
ENS2: −0.09 ± 0.04
CMIP3 aerosol complexity
Direct: −0.03 ± 0.03
Interact: −0.16 ± 0.04
• Lack of multidecadal SST variability in the North Atlantic was suggested to be due to underestimation of aerosol indirect effects in coupled climate models
• In this interpretation, multidecadal variations of the North Atlantic SSTs are forced (cf. Zhang et al. 2007)

Aerosols?

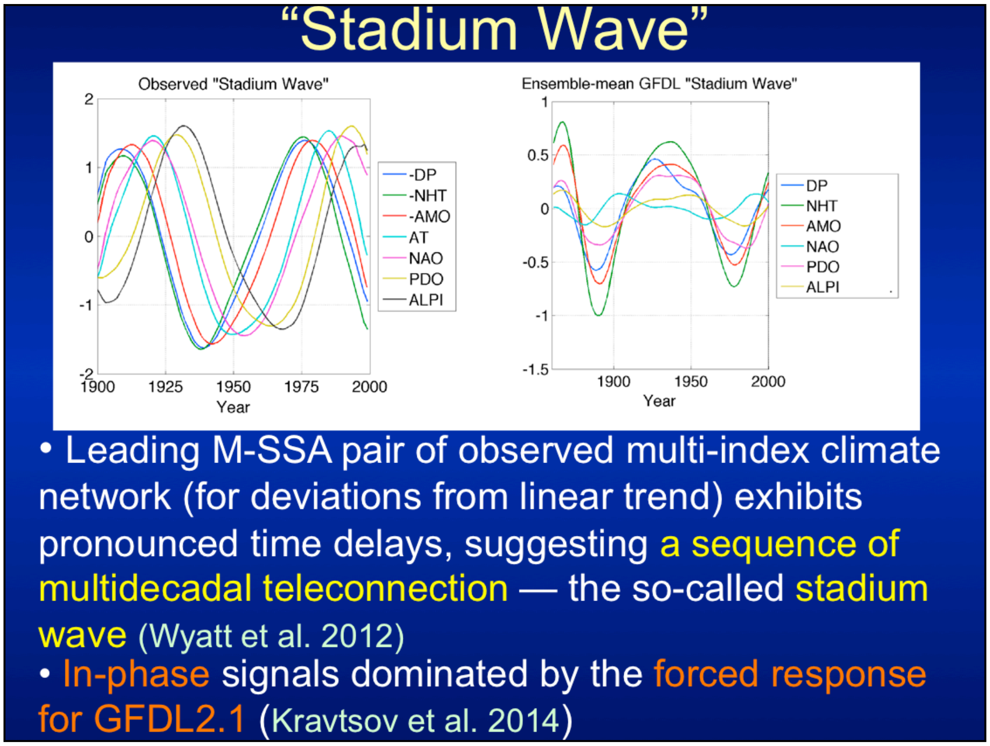
Have Aerosols Caused the Observed Atlantic Multidecadal Variability?

RONG ZHANG,* THOMAS L. DELWORTH,* ROWAN SUTTON,+ DANIEL L. R. HODSON,+ KEITH W. DIXON,* ISAAC M. HELD,* YOCHANAN KUSHNIR,# JOHN MARSHALL,@ YI MING,* RYM MSADEK,* JON ROBSON,+ ANTHONY J. ROSATI,* MINGFANG TING,# AND GABRIEL A. VECCHI*

- major discrepancies between HadGEM2-ES simulations and observations in terms of the 3-D structure of multidecadal upper-ocean temperature and salinity in the North Atlantic, as well as in various fields outside of North Atlantic (we'll see some of that later in the talk)

- Still, if observed multidecadal deviations of North Atlantic SSTs from linear trend are internally generated, why is their magnitude so much larger than that in CMIP5 coupled runs?

Well, maybe not, but what would then be the explanation for the insufficient simulated amplitude of multidecadal variability?...



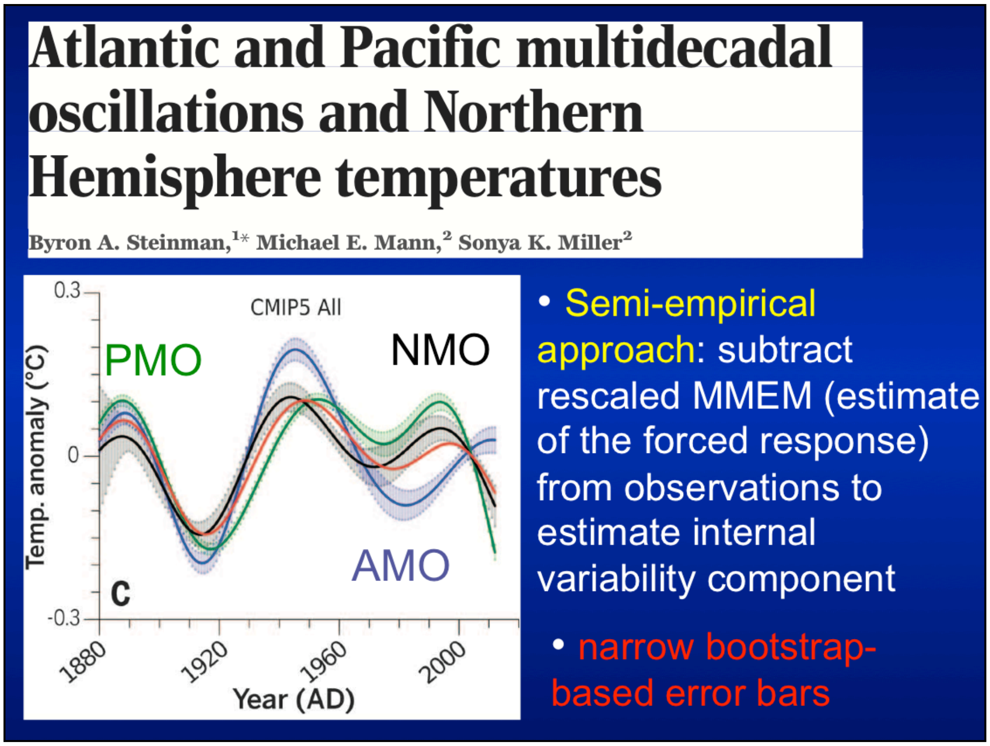
The spatiotemporal structures of dominant multidecadal climate variability over a multi-index climate network in observations and GFDL2.1 model are also very different. This multidecadal signal in GFDL is predominantly *forced*. The run-to-run uncertainties in the simulated phases of the stadium-wave components corresponding to different indices (not shown) are small compared to the phase shifts between the observed components of the stadium wave. So the observed and GFDL2.1 simulated multidecadal variability in the 20th century (in deviations from the linear trend, which trend is, btw, similar in GFDL model and observations) differ in magnitude and spatiotemporal structure.

Forced climate model runs can be used to estimate the response of the climate system to forcing

• Linear detrending is not meant to isolate forced and internal components of variability (as opposed to claims in, e.g., Mann et al. 2014)

• One can use ensemble simulations using single (SM) or multiple (MM) climate models to estimate the climate's forced response over the 20th century (Kravtsov and Spannagle 2008; Knight 2009; Terray 2012, Steinman et al. 2015a)

• SMEM (ensemble mean) time series of a given climatic quantity would approximate an individual model's forced response. MMEM would characterize the average forced response of the MM ensemble



Steinman et al. used semi-empirical approach to isolate internal variability in regional surface temperature time series. It was used to attribute recent pause in the NH warming to the downswing in the SST over the Pacific. Narrow bootstrap error bars are misleading, as they make an impression of the consistency in the forced response "shape" among different CMIP5 models, which is not at all the case (cf Booth et al. 2012). In reality, each bootstrap sample averages among ~2/3 of the 170 individual simulations considered, and thus contains signatures of the forced responses of most of the 40 individual models. So bootstrap based error bars for MMEM are narrow because each MMEM subsample averages over the same set of individual forced signals.

Comment on "Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures"

S. Kravtsov,[1*] M. G. Wyatt,[2] J. A. Curry,[3] A. A. Tsonis[1]

$$x_m^{(t)};\ m = 1, \dots, M;\ t = 1, \dots, T \quad (1)$$

$$\left[x^{(t)}\right] = \frac{1}{M}\sum_{m=1}^{M} x_m^{(t)} \quad (2)$$

$$x_m^{(t)} = f_m^{(t)} + \epsilon_m^{(t)} \quad (3)$$

$$f_m^{(t)} = \left[x^{(t)}\right], \quad (4a)$$

$$f_m^{(t)} = a_m\left[x^{(t)}\right]. \quad (4b)$$

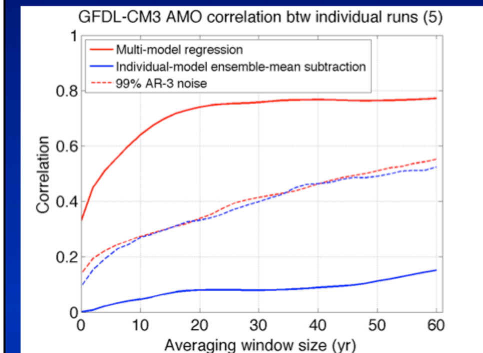
$$\left[\epsilon^{(t)}\right]=0;\ t = 1, \dots, T, \quad (5)$$

$$\overline{[\epsilon]^2} \ll \left[\overline{\epsilon^2}\right]/M \quad (6)$$

- Small variance of ensemble-mean of internal residuals in Steinman et al. (2015a) does not imply statistical independence of residuals, but is an artifact of constraint (6)

This result also holds if you exclude single simulation or blocks of simulations from the computation of MMEM estimate, creating out-of sample versions of MMEM. In this case, the variance of the ensemble-mean residual time series would scale as 1/M^2 and not as 1/M no matter of whether the residuals are actually independent or not.

SMEM vs. MMEM



$$C = \frac{2}{M(M-1)} \sum_{m>l} C_{ml}\, H(C_{ml}),$$

- Individual model residuals after MMEM subtraction are correlated!
- Rescaled MMEM is not a good measure of the forced signal in individual model runs of CMIP5 ensemble, and the resulting residual "internal variability" is dominated by the differences between MMEM and a given model's true forced signal
- 5-yr low-pass filtered SMEM is much more accurate

We aim to:

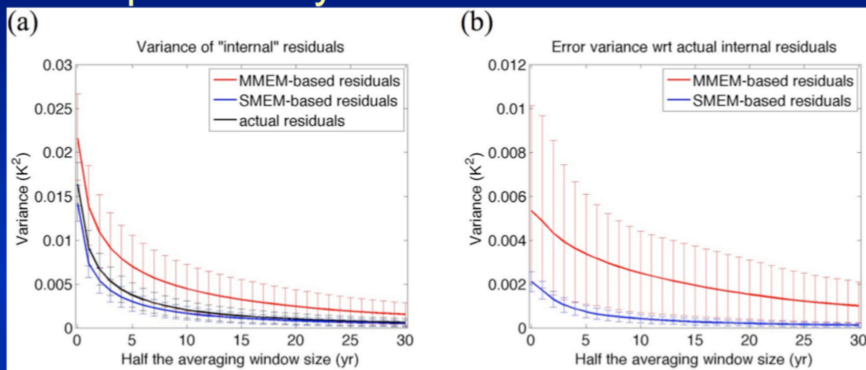
- Modify Steinman et al.'s approach and develop a framework to properly estimate forced signal and its uncertainty from CMIP5 multi-model ensemble
- Combine these forced-signal estimates with individual model simulations as well as observations to obtain estimates of internal climate variability
- Compare characteristics of the simulated and "observed" (semi-empirical) internal variability

Methodology: Models

- Analyze CMIP5 historical runs for models with four or more realizations (18 models, 116 simulations, table two slides down)
- Use 5-yr low-pass filtered SMEM as an initial estimate of each model's forced signal, compute 'internal' residuals
- Fit low-order ARMA models to these residuals and produce multiple synthetic versions of internal variability foe each model
- Add synthetic residuals to estimated forced signals to produce synthetic CMIP5 "ensembles"
- Use the synthetic ensembles to correct for the biases in the initial estimates of the forced signals and internal variability; these biases can be computed since the true forced signals in the synthetic samples are known by construction

This gives us 116 bias corrected time series of the internal variability as simulated by CMIP5 models

Biases and uncertainties of synthetic AMO decomposition by SMEM and MMEM methods



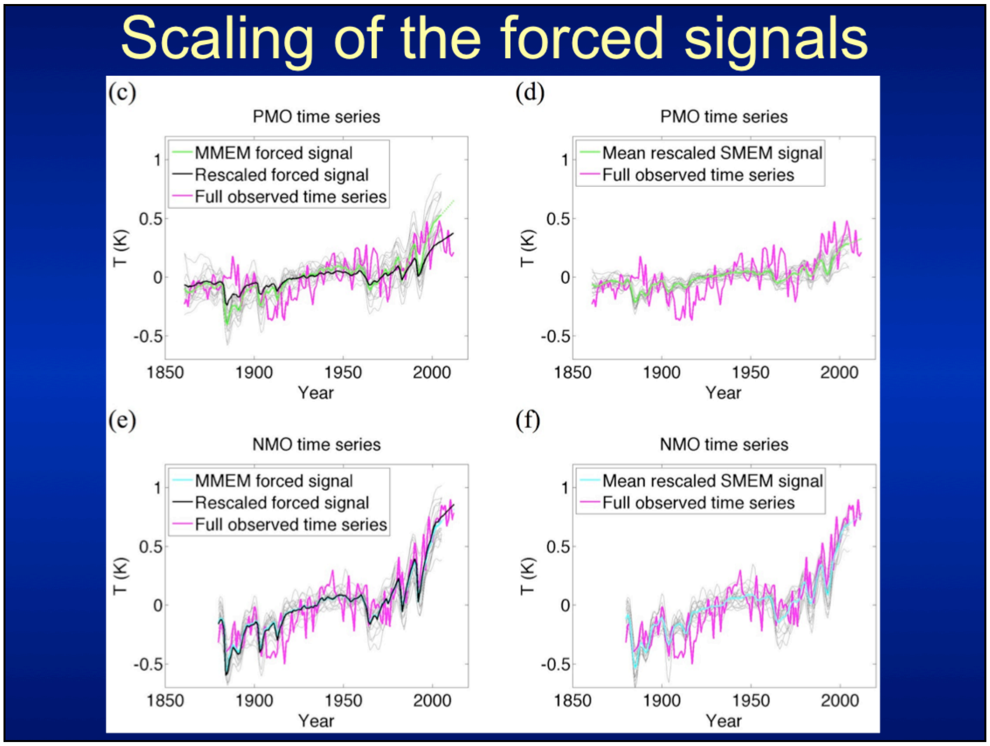
- SMEM method slightly underestimates the variance of internal variability due to aliasing some of this variance into the estimated forced signal. Error variance is small.
- MMEM method grossly overestimates the internal variance by misidentifying the difference btw the MMEM signal and true forced signal as the internal variability. Error variance is large.

This is an example of how all of this works. SMEM is clearly a much better estimate of a given model’s forced signal compared to the scaled MMEM signal. The SMEM-based estimated variance of the internal variability has to be inflated to compensate for the method’s variance bias.

Model runs used (20th century)

#	Model acronym	Number of realizations	Scaling wrt MMEM			Scaling wrt observed		
			AMO (0.89)	PMO (0.58)	NMO (1.05)	AMO (0.8)	PMO (0.57)	NMO (1.02)
1.	*CCSM4*	*6*	*1.18*	*1.50*	*1.31*	*0.68*	*0.47*	*0.78*
2.	*CNRM-CM5*	*10*	*0.68*	*0.86*	*0.84*	*1.15*	*0.76*	*1.22*
3.	**CSIRO-Mk3-6-0***	**10**	**0.68**	**0.64**	**0.72**	**1.10**	**0.57**	**1.21**
4.	*CanESM2*	*5*	*0.97*	*1.02*	*1.00*	*0.81*	*0.53*	*0.95*
5.	GFDL-CM2p1	10	1.26	1.41	1.31	0.61	0.48	0.78
6.	GFDL-CM3*	5	0.67	0.60	0.79	0.80	0.21	0.91
7.	GISS-E2-Hp1	6	1.0	0.95	1.00	0.82	0.70	1.04
8.	*GISS-E2-Hp2*	*5*	*0.83*	*0.82*	*0.84*	*1.03*	*0.72*	*1.21*
9.	GISS-E2-Hp3	6	1.23	1.17	1.15	0.72	0.61	0.92
10.	GISS-E2-Rp1	6	1.03	0.82	0.93	0.8	0.70	1.11
11.	GISS-E2-Rp2	6	0.76	0.67	0.73	1.1	0.71	1.30
12.	GISS-E2-Rp3	6	1.53	0.93	1.11	0.48	0.64	0.93
13.	GISS-E2-Rp4	6	1.46	1.42	1.42	0.56	0.50	0.71
14.	*HadCM3*	*10*	*0.66*	*1.09*	*0.90*	*0.83*	*0.57*	*1.11*
15.	**HadGEM2-ES***	**5**	**0.86**	**0.34**	**0.64**	**0.84**	**0.33**	**1.21**
16.	**IPSL-CM5A-LR**	**6**	**1.66**	**1.72**	**1.44**	**0.48**	**0.42**	**0.72**
17.	**MIROC5***	**4**	**0.92**	**0.67**	**0.77**	**0.93**	**0.64**	**1.13**
18.	MRI-CGCM3*	4	0.75	0.73	0.73	0.77	0.70	1.20

We analyzed model 20th century CMIP5 runs for models with 4 or more realizations available, and several climate indices: AMO, PMO, NMO of Steinman et al. (2015a), as well as SLP based indices not shown here (NAO, ALPI).

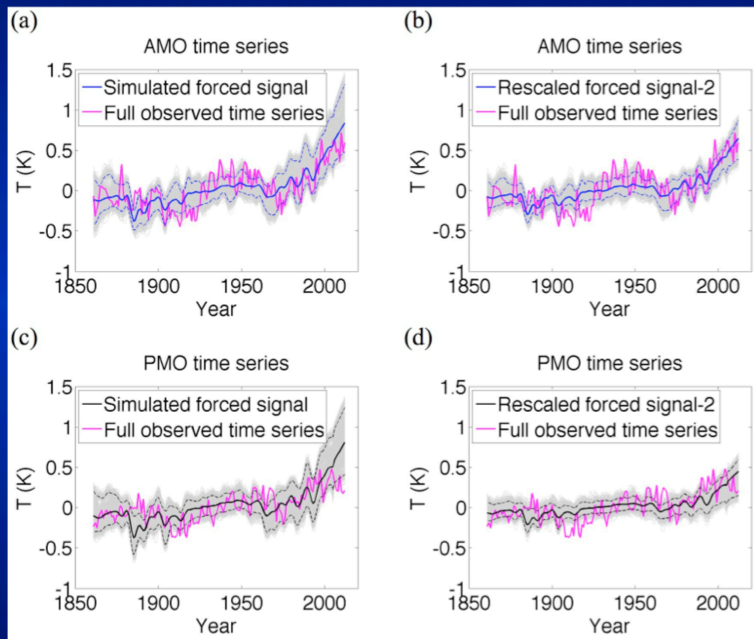


This figure illustrates scaling of the estimated forced signals in Steinman et al. (2015) [left] and in our procedure (right). Steinman et al. scale MMEM and we scale individual models' SMEMs. The ensemble mean over the latter is similar (but not identical) to the scaled MMEM. This is meant to correct for different climate sensitivity of different models; this procedure naturally removes some of the spread in the estimates of the forced response among individual models. On average, the models tend to warm too fast in the Pacific region, moderately too fast in the Atlantic (not shown), and do about right over the entire Northern Hemisphere (which means they have to warm slower than observed over land). This is consistent with sensitivity factors (regression slopes) in the previous slide's table.

Methodology: Observations

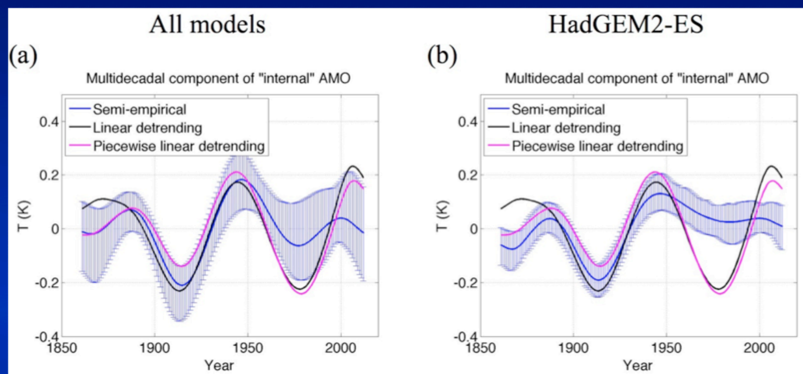
- In constructing the synthetic CMIP5 ensembles, first rescale the initial estimates of the forced signal in individual models (5-yr LPF SMEM) to best fit the observed time series considered. This is meant to correct for different climate sensitivities of different models.
- Estimate the forced signal uncertainty by computing 100 versions of SMEMs in the 100 synthetic CMIP5 ensembles (the total of 18x100=1800 estimates of the forced signal). These are now our estimates of the forced signal in observations (due to prior rescaling!)
- Subtracting the estimated forced signals from the observed time series gives us 1800 estimates of internal variability in observations

Estimates of the forced signal

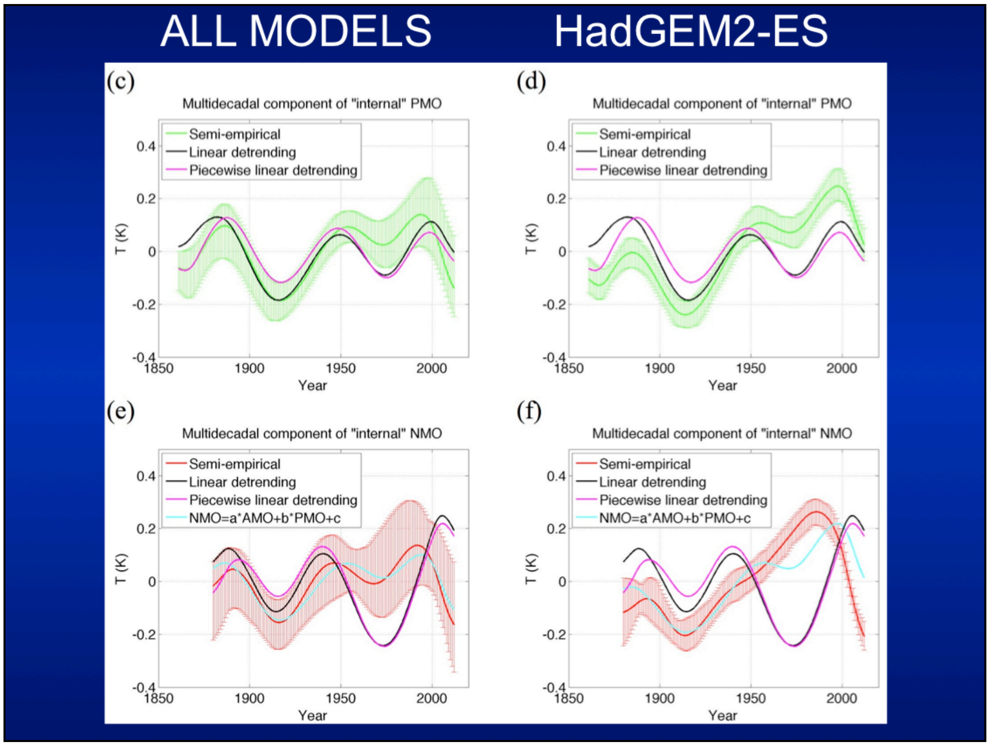


Left – no rescaling; right – rescaled signals. Linear growth of uncertainty at the end of the record is due to linear extrapolation of model time series from 2005 through 2012.

Semi-empirical internal variability



- Note large uncertainties (cf. Steinman et al.)
- Part of the uncertainty is due to different forced signals of different models (so the total spread is due to overlapping spreads of forced signal from different models)
- Example: HadGEM2-ES forced signal matches the full observed AMO time series well in the 2/2 of the 20th century, so its estimated observed internal variability is small there



Important difference from Steinman et al.: the estimated internal cooling of the PMO and NMO during the climate hiatus is of about the same magnitude in the estimate based on all models (or larger for NMO in HadGEM2-ES). If this is so, the Pacific is not easily interpreted to have caused the hiatus, contrary to Steinman et al. claims.

Causes of GW hiatus

- Steinman et al.: “internal” AMO flat, PMO drops strongly, NMO in between, hence PMO drives the NMO’s internal downswing, which counteracts forced warming
- Our results: the NMO’s “internal” drop is steeper than PMO’s, hence NMO decrease cannot be solely due to PMO decrease, and hemispheric-scale dynamics must be in play to cause the hiatus

Caveat: it all depends on rescaling!

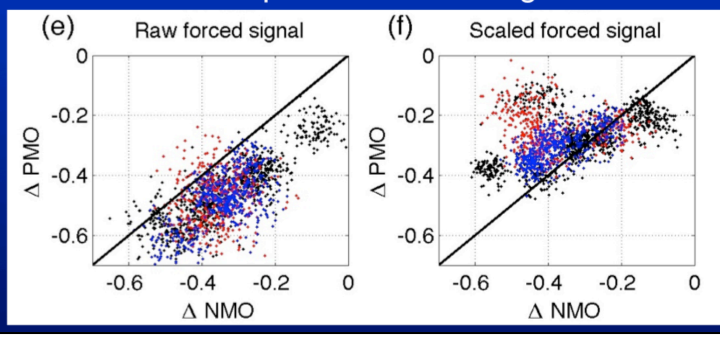
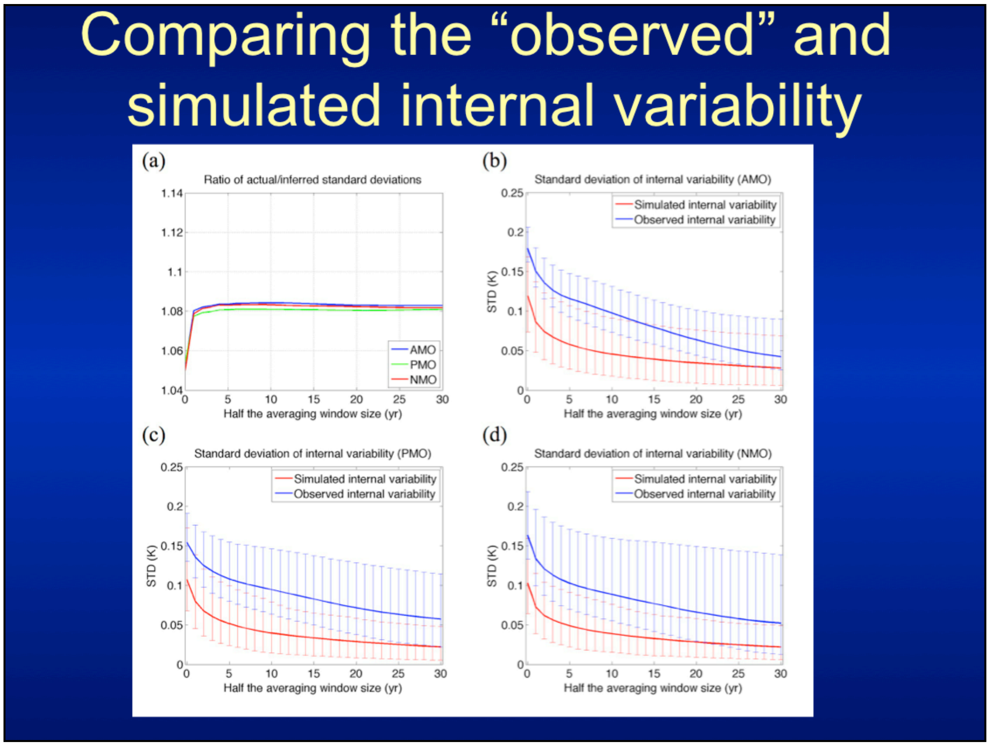
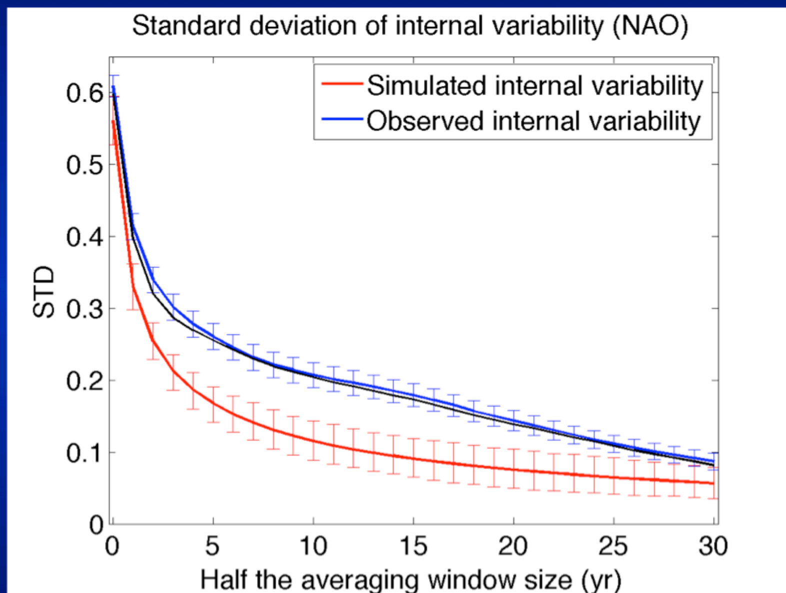


Fig.: In the models, Pacific warms faster than NH; if we rescale to match observations, NMO drop doesn’t change much, but PMO forced warming gets scaled down a lot, leading to the I”nternal” NMO drop being larger than PMO drop.

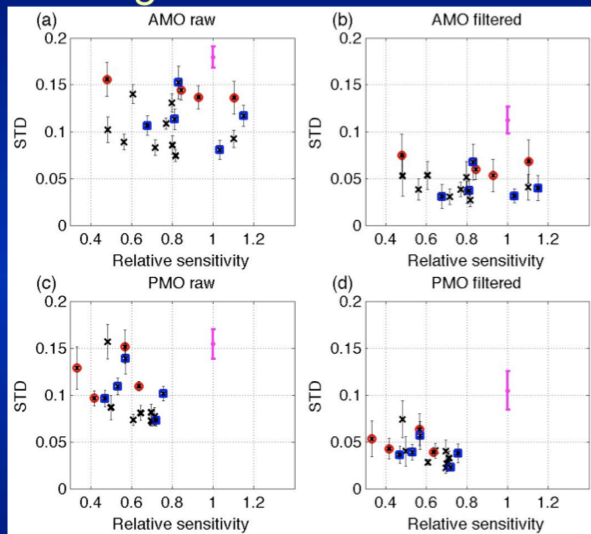


Even inflated internal variability in CMIP5 simulations is significantly smaller than observed estimates!

Same story for NAO and ALPI

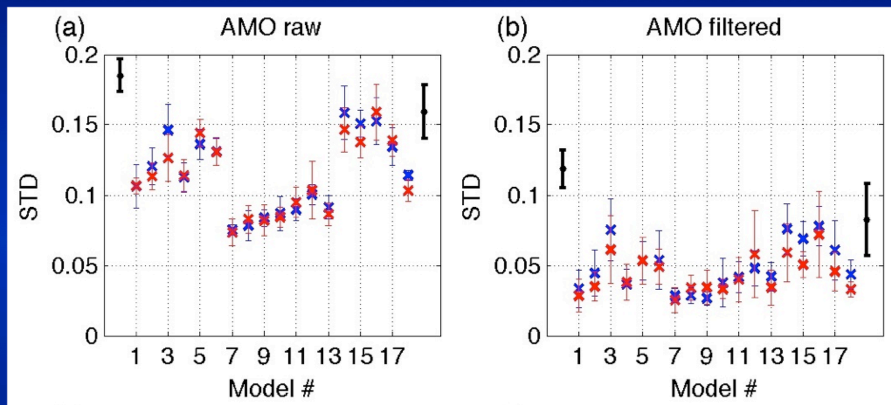


Any connection btw model climate sensitivity and magnitude of internal variability?



… Doesn't look like it: forced response processes and feedbacks seem to be different from those for internal variability

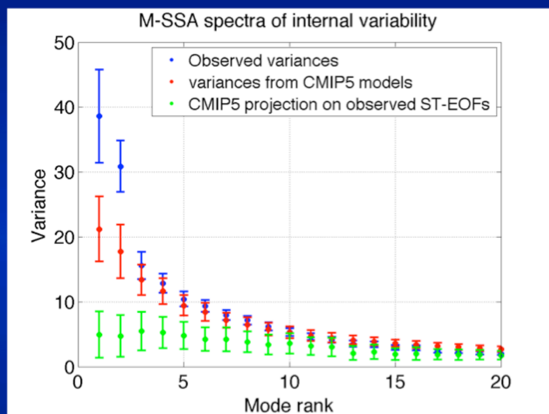
Any evidence of suppression of internal variability in the second half of the century?



Nope! There is no evidence of forced suppression of internal variability in models.

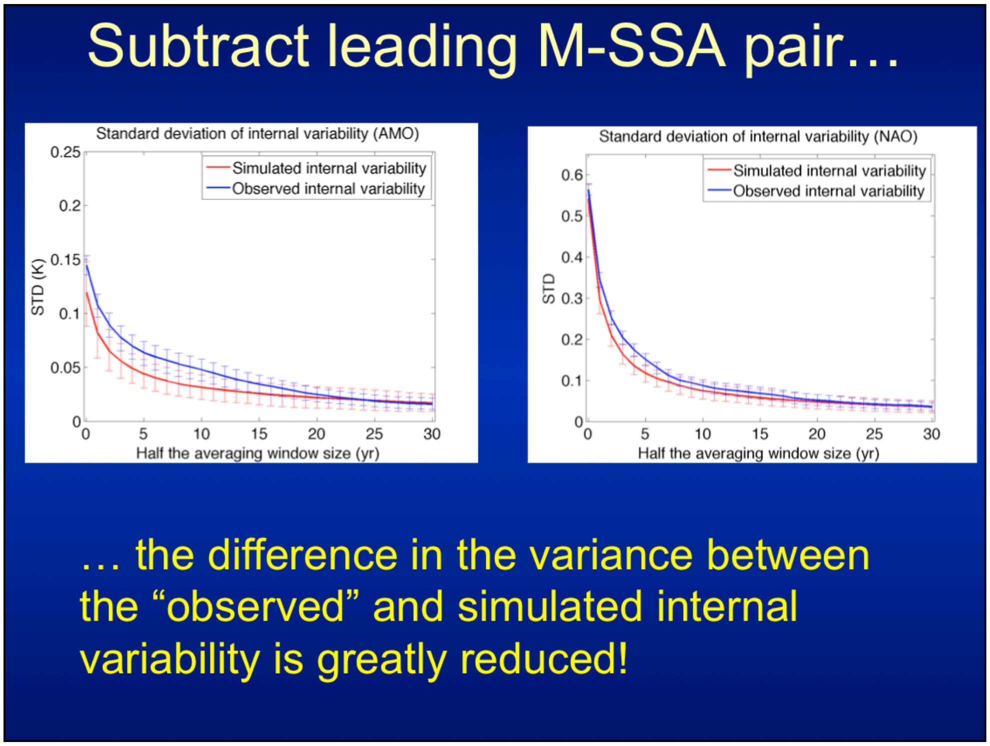
- Both of these properties justify our treatment of forced signal and internal variability as being independent of each other

M-SSA of internal variability network (AMO, PMO, NMO, NAO, ALPI)



- Dominant leading pair in observations
- Monotonic decay of variance in models
- Leading pair dominates the difference btw models and observations
- Different spatiotemporal structure of the observed and simulated variability

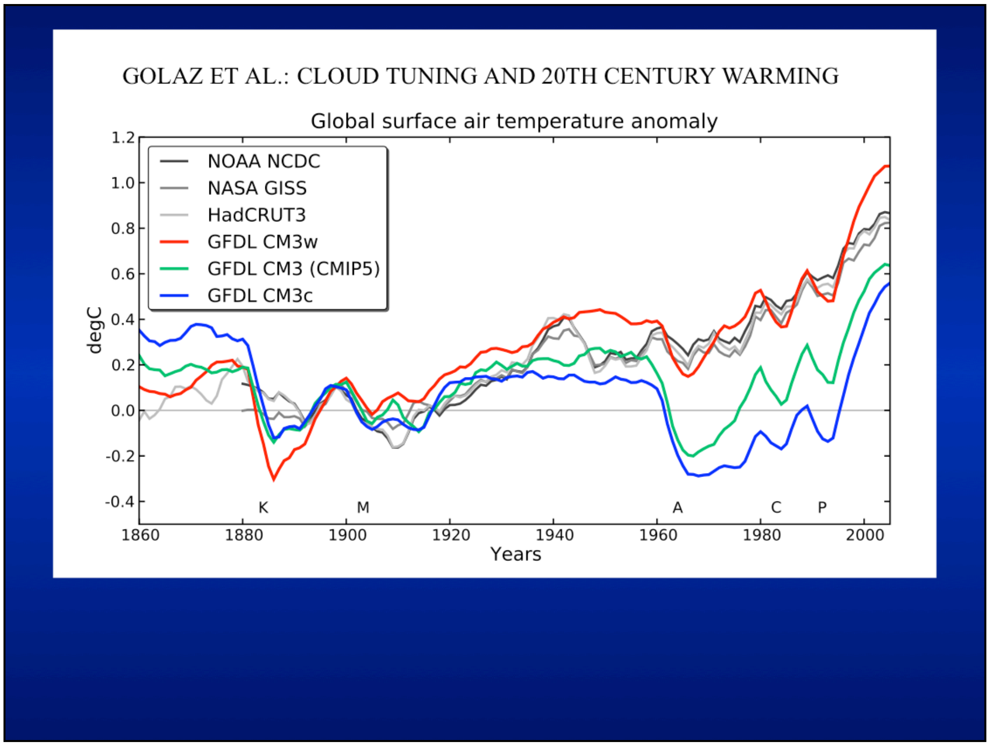
Blue – observed, red – simulated, green – variance of projections of the simulated trajectory matrices onto observed T-EOFs. Dominant leading pair in observation, with the variance much larger than that of the dominant pair in models. Model's projections onto observed T-EOFs are tiny: different spatiotemporal structure of the simulated variability compared to the observed.



So comparing semi-empirical internal variability in models and observation recovers the results based on comparing the deviations from the linear trends in observations and model simulations: lack of multidecadal variance and different spatiotemporal structure of variability in the models relative to observations.

Summary

- We used Monte-Carlo approach to estimate forced signals from multi-model ensemble of CMIP5 historical simulations
- These forced signals were subtracted from individual model runs and, after rescaling, from observed time series to derive the internally generated component of the observed and simulated climate variability
- Internal climate variability in models has a smaller amplitude and different spatiotemporal structure wrt the observed variability
- The differences between models and observations are dominated by a low-dimensional multidecadal mode of the observed climate variability, which is apparently absent from models



GOLAZ ET AL.: CLOUD TUNING AND 20TH CENTURY WARMING
Global surface air temperature anomaly
NOAA NCDC
NASA GISS
HadCRUT3
GFDL CM3w
GFDL CM3 (CMIP5)
GFDL CM3c
degC
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
K
M
A
C
P
1860
1880
1900
1920
1940
1960
1980
2000
Years