

# **Do Climate Models Underestimate the Sensitivity of Northern Hemisphere Sea Ice Cover?**

Michael Winton  
Geophysical Fluid Dynamics Laboratory/NOAA, Princeton, New Jersey

Submitted to *Journal of Climate*

12 October 2010

---

*Corresponding author address:* Michael Winton, NOAA/GFDL, Princeton University Forrestal Campus,  
201 Forrestal Rd., Princeton, NJ 08540.  
E-mail: Michael.Winton@noaa.gov

## Abstract

The sensitivity of Northern Hemisphere sea ice cover to global temperature change is examined in a group of climate models and in the satellite era observations. The models are found to have well defined, distinguishable sensitivities in climate change experiments. The satellite era observations show a larger sensitivity - a larger decline per degree warming - than any of the models. To evaluate the role of natural variability in this discrepancy, the sensitivity PDF is constructed based upon the observed trends and natural variability of multi-decadal ice cover and global temperature trends in a long control run of the GFDL CM2.1 climate model. This comparison shows that the model sensitivities range from about one to more than two pseudo-standard deviations of the variability smaller than observations indicate. The impact of natural Atlantic multi-decadal temperature trends (as simulated by the GFDL model) on the sensitivity distribution is examined and found to be minimal.

## 1. Introduction

A multi-decadal decrease in Northern Hemisphere sea ice cover has been detected with satellite and earlier observations (Comiso et al 2008). The anthropogenic signal in the decline emerged in the early 1990s (Min et al 2008). The decline is qualitatively consistent with climate model simulations of the post-war period when anthropogenic forcings are included (Vinnikov et al 1999; Vinnikov et al 2006; Stroeve et al 2007). Vinnikov et al (2006) found that in a group of eleven IPCC AR4 models, four had annual NH ice extent decline rates larger than observed over the 1972-2004 period. However, using a longer observed dataset (1953-2006), Stroeve et al (2007) found that the observed March and September fractional rates of decline were triple the respective model means and, for September, larger than the decline rate in all of the individual model runs. The disparities were smaller over the period of satellite observations. They note that the disagreement between models and observations could indicate a substantial natural variability component to the observed decline or an underestimation of the sea ice sensitivity in the models. This study attempts to clarify this ambiguity by quantifying the role of natural variability in multi-decadal trends and assessing the likelihood that it can account for the difference between simulations and observations.

To sharpen the capacity of our analysis to verify or falsify the simulations with observations, three changes are made in the following analysis relative to the Stroeve et al (2007) study:

- 1) Only satellite observations are used. Although this restricts the comparison to the post 1979 period, it eliminates uncertainty associated with the sparse and heterogeneous observations of the earlier period and their blending with the satellite observations. Observational uncertainty is not completely eliminated by restricting to satellite observations. There are two algorithms used for converting the satellite observation to sea ice extent: NASA Team and Bootstrap. The trends of monthly anomalies 1979-2006 (available at [http://nsidc.org/data/smmr\\_ssmi\\_ancillary/area\\_extent.html#gsfc](http://nsidc.org/data/smmr_ssmi_ancillary/area_extent.html#gsfc)) differ by about 5% with the NASA Team decline being larger. We use the NSIDC Sea Ice Index which is based on NASA Team including the preliminary data through 2009.
  
- 2) Some studies have focused on the dramatic September ice cover decline (Boe et al 2009a; Wang and Overland 2009, Zhang 2010). Here, annual average sea ice extent is used in preference to September or other monthly values. Observations and models show that Arctic sea ice anomalies typically persist for only a few months (Blanchard-Wrigglesworth et al, submitted). Additionally, September sea ice cover, the focus of many ice sensitivity studies, is particularly variable and its variability is expected to increase with thinning of the ice (Holland et al 2006; Goosse et al 2009; Eisenman 2010). Therefore considerable variation that is not related to long term trends can be reduced by using annual averages.
  
- 3) In order to factor out the potential for uncertainties in global sensitivity or forcing to impact the results, we evaluate the ice cover sensitivity to global warming

rather than the ice cover change itself. The Arctic is strongly coupled to the global climate through a large atmospheric heat transport. Gregory et al (2002) show for HadCM3 (and we will verify this for other models) that the NH annual ice cover decline is proportional to the global temperature change. Therefore an error in ice cover response may arise from an error in simulated global warming either due to problems with simulated global sensitivity or forcing. This possibility can be eliminated by comparing sensitivities - the ice cover decline per degree of global warming. A similar sensitivity-based approach has been employed by Zhang (2010) but using seasonal ice cover and Arctic temperatures.

Elaborating on the last point, the ice cover response  $\Delta I$ , to a radiative forcing  $F$ , can be written:

$$\Delta I = \frac{\Delta I}{\Delta T} \cdot \frac{\Delta T}{F} \cdot F \quad (1)$$

where  $\Delta T$  is the change in global temperature. This paper focuses on evaluating the first factor on the right side of (1). The second factor on the right is closely related to the transient climate response (TCR) – the global temperature change in a climate model at CO<sub>2</sub> doubling in a reference 1%/year CO<sub>2</sub> increase experiment. Gregory and Forster (2008) make an observation-based estimate of the TCR and its uncertainty. The uncertainty is dominated by natural variability of  $\Delta T$  which they estimate using a climate model. Since the NH sea ice covers only about 2% of the global surface, it is not a major contributor to this variation even after accounting for a several-fold regional amplification of temperature change. Gregory and Forster (2008) note that the forcing

has a smaller uncertainty over the post-1970 period than over the entire post-industrial period where changes in aerosol and solar forcing uncertainty are major factors.

To evaluate the ice-temperature sensitivity term in (1)  $\Delta I/\Delta T$ , we must choose a method for evaluating the relationship from noisy time series of I and T. This is done in section 2. After describing the model/observation discrepancy in the sensitivity, we attempt to clarify the individual roles of model bias and natural variability in section 3. The variability that is most likely to account for the discrepancy is multi-decadal, not interannual, and so must be evaluated with time series outside the period of interest. We lack sufficiently long and accurate observations of ice cover and global temperatures that are not contaminated by anthropogenic forcing that would allow us to make an observational estimate of natural multi-decadal variability. Following Gregory and Forster (2008), we rely on a model estimate of the natural variability. The natural multi-decadal variability of climate models is sparsely documented, probably due to the expense of generating historical ensemble members and long control runs. In section 3, we will make use of a 4000 year control run of the GFDL CM2.1 model to develop a variability measure for the model/observation discrepancy. Conclusions are presented in section 4.

## **2. The ice cover sensitivity to global warming**

To begin, we compare time series of global temperature and NH sea ice cover in observations and five climate models. The observed temperatures are global means from the GISTEMP combined land-ocean dataset and the sea ice cover is Northern Hemisphere extent from the NSIDC sea ice index dataset (Fetterer et al 2009). The

models shown are the members of the multi-model ensemble presented later with the more sensitive ice responses. Figure 1 shows time series with the models displayed from most to least sensitive, top to bottom. The most sensitive model is the new GFDL CM3 model developed for the IPCC fifth report. The other models are IPCC fourth report models. For all of the model runs, a short segment of a scenario run is appended to the historical run to bring the time series up to 2009. The GFDL CM3 run has a very close simulation of the annual ice cover and its decline but the accompanying global temperature trend is larger than observed. We are not concerned here with whether the individual model trends are within natural variability of the observed, only the relationship between the two trends in models and observations. The CCSM3 and HadGEM runs also have ice trends comparable to observed but temperature trends that are larger. MIROC and ECHAM5 have temperature trends close to observed but ice declines that are smaller. In summary, none of the runs has both a temperature trend as small as observed and an ice decline as large.

A clearer picture of this discrepancy comes from Fig. 2 where we scatter plot the temperature and ice cover trends for the model runs and observations. We might measure the sensitivity,  $\Delta I/\Delta T$ , as the ratio of the trends (the slope of the line from the origin to the trend pair). The slopes to the models trends are all smaller than for the observed indicating less sensitivity in all of the model runs. This trend ratio is a straightforward measure of the sensitivity but there are other possibilities to consider.

Ordinary least squares regression (OLS) is appropriate for estimating a relationship between a random and non-random variable but gives a biased estimate when both variables are random. OLS regression of ice cover on temperature and the inverse of the temperature-on-ice OLS regression give upper and lower bounds on the relationship (Table 1). The ratio of standard deviations is a neutral estimate equal to the geometric mean of the two OLS estimates. This estimate, however, treats all the variance of both variables as if it contributes to the relationship between them. We can clearly see significant ENSO variability in the global temperature that we do not expect to be informative on the forced relationship because of its equatorial origin and short timescale. The traditional way of filtering this is to take trends. The trend ratio then becomes a measure of the sensitivity. The trend accounts for a greater proportion of the ice time series ( $r^2=.83$ ) than of the global temperature time series ( $r^2=.71$ ). Consequently the trend ratio sensitivity is larger than the neutral estimate.

Another method can be constructed using the trends. Total least squares (TLS), also known as Deming regression in the two-dimensional case, minimizes the squared orthogonal distance between the data and the regression line. Since this distance incorporates units of both variables, the TLS relationship depends upon the choice of units. TLS is the maximum likelihood estimator of the relationship when the units are chosen to be the standard deviations of the respective normally-distributed error terms. We use the residuals of the time trends for our estimate of these standard deviations. The TLS estimate of the sensitivity is (Fuller 1987)

$$\left( \frac{\Delta I}{\Delta T} \right)_{TLS} (I(t), T(t)) = \frac{\sigma_I^2 - \lambda \sigma_T^2 + \sqrt{(\sigma_I^2 - \lambda \sigma_T^2)^2 + 4\lambda \rho_{I,T}^2 \sigma_I^2 \sigma_T^2}}{2\rho_{I,T} \sigma_I \sigma_T} \quad (2)$$

where

$$\lambda = \frac{\sigma_I^2 (1 - \rho_{I,t}^2)}{\sigma_T^2 (1 - \rho_{T,t}^2)} \quad (3)$$

is the ratio of the trend residual variances. The formula for TLS is more complex than that for the trend ratio and involves the correlation of the variables with each other in addition to their correlations with time. We can test TLS and the trend ratio by applying the two techniques to artificial time series constructed by adding known trends ( $0.5 \text{ K}(31 \text{ yr})^{-1}$  and  $-1.5 \cdot 10^{12} \text{ m}^2(31 \text{ yr})^{-1}$ ) to 100 31-year time series of global temperature and NH ice cover taken from the long control run of GFDL CM2.1. Table 2 shows the bias and standard deviation of the two methods. Both methods recover the specified sensitivity,  $-3 \cdot 10^{12} \text{ m}^2/\text{K}$ , with small biases but the TLS estimate has much lower variance, making it the better choice. Although the two methods give similar estimates when the observed satellite era data are used, this test on a larger set of model-generated time series shows that this similarity is not generally true.

Now we turn to the sensitivity of the models. The long-term relationship between ice cover and global temperature is quite linear for the large changes from 1900 to 2100 using forcing scenarios for the 21st century. The OLS sensitivities over this period for

eight models are listed in Table 3. The method used to evaluate the relationship does not matter here because the correlations are so high (all of the methods listed in Table 1 approach  $-\sigma_I/\sigma_T$  as  $\rho_{I,T}$  approaches one). Fig. 3 shows scatter plots of global annual mean temperature and Northern Hemisphere ice cover for two GFDL models which fall at either end of the model sensitivity range. Two scenario-forced experiments are shown: a strong forcing and a medium forcing. For the medium forcing, we distinguish the 21<sup>st</sup> century when the forcing is increasing from the subsequent two centuries when forcing is stabilized. The sensitivity, as represented by the OLS regression line, is not dependent upon the magnitude of the forcing or its rate of change. For each model the sensitivity is well defined – the relationships are linear – and the differences between the models are robust to forcing details. Note that in the CM3 experiment using the RCP8.5 (8.5 W/m<sup>2</sup> anthropogenic forcing), the annual ice goes to zero. The Arctic ice is eliminated in winter as well as in summer by the year 2100 in this experiment but the trajectory in ice/temperature space remains linear right to the elimination of the NH sea ice. Thus the ultimate magnitude of the forcing also does not disturb the linearity of the relationship.

The large correlations of temperature and ice cover in Table 3 show that other models have similarly linear relationships. Thus it makes sense to think of the sensitivity as a property of a climate model, and -- assuming the linear dynamics that characterize all the models is correct -- the climate system as well. While the sensitivity from observations using trend-based methods is about  $-3 \cdot 10^{12} \text{m}^2/\text{K}$  (Table 1), the model sensitivities listed in Table 3 range from -2.2 to about  $-0.8 \cdot 10^{12} \text{m}^2/\text{K}$ . As is the case with the TCR, the models agree that the ice-temperature sensitivity parameter is fairly constant in forced transient

simulations but they have substantial disagreement on the value of that constant.

However, unlike the TCR where the observational estimate falls in the middle of the model range (Gregory and Forster 2008), the model ice-temperature sensitivity values are all well less in magnitude than the observed value. Ridley et al (2007) explore variation of the ice-temperature sensitivity due to perturbing parameters in HadCM3 and find that the change in total heat transport into high latitudes strongly influences the range of sensitivities.

The discrepancy between observed and simulated sensitivities does not necessarily imply model error because the observed sensitivity is an apparent sensitivity that is influenced by both the true sensitivity of the climate system and natural variability. Over time the apparent sensitivity approaches the true sensitivity as the forced component of the temperature and ice cover responses rises above natural variability. We can observe this convergence in the model projections. Fig. 4 shows a 21st century time series of the model apparent sensitivities calculated with the TLS method using data from 1979 up to the particular date. Noting the convergence of ensemble members and comparing the 2100 values with those listed in Table 3, we see that the model sensitivities converge on their true values over the 21st century. The ensemble members approach the true value from both directions; there is no indication that the current apparent sensitivity is a biased estimator of the true sensitivity. At the beginning of the century the runs show considerable disagreement among the apparent values. The running TLS apparent sensitivity from observations is also plotted up to present. Currently, the observed sensitivity is larger in magnitude than that in any of the model runs. The observed value

has increased in magnitude over the last decade as global warming has slowed but the ice retreat has accelerated. Earlier in the decade there were ensemble members with sensitivities as large in magnitude as the observed but that is no longer the case as these ensemble members have converged toward their true values. The model sensitivities shown in Fig. 4 were also calculated using the trend ratio method for comparison to TLS (not shown). The root mean square difference between the 1979-2009 sensitivities and the 1979-2100 sensitivity was found to be 50% larger with the trend ratio method confirming the test result with CM2.1 generated time series (Table 2).

### **3. Natural variability of the sensitivity**

Our goal is to construct the PDF of true sensitivities given the apparent sensitivity of the observations in order to gauge the likelihood of the model true values. The few ensemble members available for the models are grossly inadequate for this purpose. Because of the lengthy integration needed to bring a simulation to the beginning of the satellite period, it is prohibitively costly to produce the sensitivity PDF from an ensemble of historical runs. Assuming that the natural variability is not altered by the climate change since pre-industrial times, we can use a long pre-industrial control run to generate potential influences of natural variability on apparent sensitivity.

Our method involves subtracting a potential natural trend from the observed data to obtain a residual containing a forced component and interannual variability. These artificial time series are specified as

$$T_i(t) = T_O(t) - \beta_{Ti}t \quad (4a)$$

$$I_i(t) = I_O(t) - \beta_{Ii}t \quad (4b)$$

where  $i$  ranges from 1 to 100, and  $\beta_{Ti}$  and  $\beta_{Ii}$  are the trends of temperature and ice cover in the  $i^{\text{th}}$  31-year segment of the CM2.1 control run. The “O” subscripts indicate the observed time series. For the trend ratio method we calculate the trends of (4a) and (4b) to estimate each potential forced component. The ratio of ice cover and temperature forced components is the trend ratio estimate of the true sensitivity. For the TLS estimate we feed the (4a) and (4b) time series into (2) and (3). The 100 potential natural trends are obtained by segmenting the final 3100 year section of the 4000 year control run of CM2.1 -- an early period of drift is removed.

The trends from these segments are plotted in Fig. 5 along with the observed trend pair. The natural trend standard deviations are  $0.14 \text{ K}(31 \text{ yr})^{-1}$  and  $0.59 \cdot 10^{12} \text{ m}^2(31 \text{ yr})^{-1}$  for temperature and ice cover respectively. Since the observed trends are  $0.51 \text{ K}(31 \text{ yr})^{-1}$  and  $-1.6 \cdot 10^{12} \text{ m}^2(31 \text{ yr})^{-1}$ , the observed changes are extremely unlikely to be due to natural variability alone. The natural trends are correlated ( $\rho = -0.61$ ) and their axis of co-variation aligns fairly well with the observed trend. Therefore the CM2.1 natural variability associates a larger ice cover loss with a degree of warming than does its forced response which is significantly smaller than observed (Table 3). The trend ratio true sensitivities are easy to visualize in Fig. 5 as the slope of the line connecting the observed trend pair with a natural trend pair. The alignment of the natural trends with the observed limits the variation of these slopes, reducing the uncertainty in the estimated true

sensitivity. The TLS estimate cannot easily be visualized but is subject to the same qualitative considerations. To obtain a smoother representation of the sensitivity distributions, we fit a bivariate normal distribution to the correlated trends and feed stochastically generated trend pairs based on this distribution into the two sensitivity formulae.

Fig. 6 shows the PDFs of sensitivity that are generated using this procedure -- both the scaled histograms generated from the actual natural trends and the smooth fits from 10 million correlated random pairs. The distributions are fairly symmetric and have means near the observed trend ratio and TLS means. We would like to characterize these distributions with standard deviations but a technical detail must be considered. It is known that the distribution of the ratio of correlated, normally distributed random variables does not have moments due to the potential for very large values when the denominator variable, global temperature change in this case, takes values very close to zero (Marsaglia 2006). This is true even when the denominator variable has a non-zero mean and prevents the trend ratio distribution from having a well defined standard deviation. Simulations with random numbers indicate that this is also the case for the TLS sensitivity estimate. However, it can also be shown that, as the mean of the denominator variable grows in magnitude, the distribution becomes symmetric and the tails thin. For denominator coefficients of variation (standard deviation to mean ratios) somewhat smaller than we have here, the distribution is well approximated by the normal distribution (Hayya et al 1975). At some point it makes sense to disregard rare large values and calculate pseudo-moments to characterize the central part of the distribution.

This bounding is done here by placing a floor of  $-8 \cdot 10^{12} \text{m}^2/\text{K}$  and a ceiling of  $1 \cdot 10^{12} \text{m}^2/\text{K}$  on the sensitivity values. The slight upticks at the ends of the distributions in Fig. 6 are a consequence of this range restriction. The range limiting procedure makes a slightly larger distortion to the trend ratio distribution than to the TLS distribution. We refer to the standard deviation of the bounded variable as a pseudo-standard deviation.

Both trend ratio and TLS methods give pseudo-standard deviations of about  $1 \cdot 10^{12} \text{m}^2/\text{K}$  with the TLS standard deviation slightly smaller. The models range from about one to more than two pseudo-standard deviations of the sensitivity natural variability smaller (in magnitude) than the observed value. Table 4 lists the two-sided p-values for the models determined from the smoothed distributions. These numbers can be interpreted as the fraction of natural 31-year trends that allow for a true sensitivity value as far from the observed apparent value as the true sensitivity of the particular model. About 1/3 of the variability gives true values as far from observed as CM3's; for CCSM3, about 1/5; for MIROC, about 1/10. The TLS distribution is slightly more constraining than the trend ratio distribution. Only one model is falsified by the customary 95% confidence interval. Although all of the models are significantly less sensitive than observations, generally the difference is not so extreme as to rule out natural variability as a cause. We emphasize that these results are based on the natural variability of a single model. To indicate their sensitivity to the distribution of trends, we increased and decreased the trend standard deviations by 20% and recalculated the model p-values with the broader and narrower sensitivity distributions produced by the random number method. With the larger

variability, no model was falsified by the 5% criterion, while with the smaller variability, all but the three most sensitive were falsified.

There are indications from observations and modeling that the North Atlantic Ocean is subject to natural multi-decadal variability related to its overturning circulation, and that this variability has impact upon the NH sea ice cover and global temperature (Polyakov et al 2003; Chylek et al 2009; Mahajan et al 2010). The observed index for the AMO has a significant upward trend over the satellite era (Enfield and Cid-Serrano 2010). It is difficult to separate the influence of natural variability and forcing on the observed Atlantic Multi-decadal Oscillation (AMO, Zhang et al, 2007) but the increase in North Atlantic temperatures since the mid-1970s may partly be due to natural variability. The situation we have, with all of the model sensitivities to one side of the observation, is what would be expected from a non-neutral state of the natural variability and reasonable agreement between the models. Setting aside the ambiguity in interpreting the observed index, we can ask if natural trends in the model's AMO contribute to the true sensitivity that is diagnosed by the procedure that we have used to generate the sensitivity distribution. The AMO index used here is simply the average N. Atlantic SST between the equator and 65N in the CM2.1 control run. Fig. 7 shows the diagnosed sensitivities using TLS scattered against AMO trends over the same period. The correlation is positive but weak ( $\rho=0.18$ ). AMO trends are correlated with both global temperature ( $\rho=0.64$ ) and NH sea ice cover trends ( $\rho=-0.53$ ). An AMO trend makes it a little more likely to diagnose a smaller value of the true sensitivity because a growing AMO contributes slightly to a large apparent sensitivity. But the relationship is too weak to be

very useful. The mean AMO trend of the points in Fig. 7 that are in the model range ( $> -2.2 \cdot 10^{12} \text{m}^2/\text{K}$ ) is less than  $\frac{1}{4}$  of an AMO trend standard deviation. The natural variability of the apparent sensitivity is not well characterized by the AMO because the AMO variability impacts both global temperature and sea ice in a proportion that moves the temperature/ice state, roughly, along the major axis of the natural variability (Fig. 5). Because a natural AMO trend has small impact on the sensitivity and the natural component of the observed trend is uncertain, we make no adjustment to our estimate of the true sensitivity distribution.

## 4. Conclusions

In this paper we have investigated the sensitivity of northern hemisphere ice cover to global warming in observations and models. The interpretation of this sensitivity is more straightforward than that of the ice cover response itself. We have used the robust simulated proportionality between the ice cover change and global temperature change to factor out influences such as global sensitivity and forcing that primarily affect the latter.

The sensitivity is also better constrained by observations than the ice cover response. This can be shown using the multi-decadal variability of ice cover, sensitivity, and global temperature from the GFDL CM2.1 control run along with the values of these quantities in the observed record -- our central estimates of the forced component. From these we can form a noise-to-signal ratio as the ratio of the standard deviation of the variability to the observed changes. These coefficients of variation are .38, .31, and .27 for the ice cover change, the ice-temperature sensitivity, and the temperature change respectively.

The sensitivity is considerably better constrained than the ice response and nearly as well constrained as the temperature response – remarkable considering that the northern ice covers only about 2% of the globe. The reason for the tight constraint on the sensitivity is the similarity in the ratio of ice and temperature changes in the natural variability and observations. Even though the observed changes are too large to be solely due to natural variability, a natural component does little to disturb the relationship.

Currently, the sensitivity distribution is approximately symmetric with the model true sensitivities ranging from about 1 to more than 2 pseudo-standard deviations less sensitive than the observed apparent sensitivity. Only the least sensitive model of the ensemble used in this study is falsified using the 95% confidence interval. It is interesting to contrast this situation with that for observational estimate of the TCR. Gregory and Forster (2008) found that the 95% confidence interval for the TCR estimated using 1970-2006 observations was very similar to the range of TCRs in climate models. Here we find that the models occupy only the less sensitive portion of the ice/temperature sensitivity confidence interval. Using the IPCC recommended language all but one of the six models used in this study have ice/temperature sensitivities that are unlikely (<33% probability), four are very unlikely (<10% probability) and one is extremely unlikely (<5% probability). The IPCC also recommends characterizing the level of scientific understanding behind a result based on the amount of evidence and the level of agreement between evidence. Since only one model has been used, the level of understanding for these results is low. Clearly, a next step toward answering the title question is to estimate the sensitivity PDF using the natural variability of other climate

models. Climate models are known to have differences in the magnitudes of their variability and the alignment of GFDL CM2.1 natural trends and the observed trends might also be fortuitous.

Assuming the fidelity of the GFDL CM2.1 natural variability, the AMO does not introduce significant uncertainty in the sensitivity since it affects ice cover and temperature in a proportion similar to general natural variability and the observed sensitivity. To ascertain the natural variability component of the observed trends it would be useful to understand the nature of the variability in the perpendicular direction which associates ice cover increases with decreases in global temperature, for example. This is left to future work.

Although most models are not strictly ruled out by the analysis here, substantial natural variability is necessary to reconcile even the most sensitive model with observations. The observational constraint will tighten slowly with time but in the interim it is useful to explore the possibility that the models are not sufficiently sensitive. This has been the theme of several analyses of IPCC AR4 models since the Stroeve et al (2007) study (Bitz et al 2010; Boe et al 2009b). The results here support the importance of this work while holding onto the possibility that, at least for some of the models, the model/observations discrepancy may be due solely to natural variability.

## **Acknowledgments**

The author thanks Isaac Held and Rong Zhang for helpful reviews. The author acknowledges the international modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modeling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy.

## References

- Blanchard-Wrigglesworth, E., K. Armour, C. M. Bitz, and E. deWeaver, 2010: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations, submitted.
- Bitz, C.M., J. K. Ridley, M. M. Holland, and H. Cattle, 2010: 20th and 21st century Arctic Climate in Global Climate Models, in press in Arctic Climate Change - The ACSYS Decade and Beyond, edited by P. Lemke.
- Boe, J., A. Hall, and X. Qu, 2009a: September sea-ice cover in the Arctic Ocean projected to vanish by 2100, *Nature Geosciences*, doi:10.1038/NGEO467.
- Boé, J., A. Hall, and X. Qu, 2009b: Current GCMs' Unrealistic Negative Feedback in the Arctic. *J. Climate*, **22**, 4682–4695.
- Comiso, J. C., C. L. Parkinson, R. Gersten, and L. Stock, 2008: Accelerated decline in the Arctic sea ice cover, *Geophys. Res. Lett.*, **35**, L01703, doi:10.1029/2007GL031972.
- Chylek, P., C.K. Folland, G. Lesins, M.K. Dubey, and M. Wang, 2009: Arctic air temperature change amplification and the Atlantic Multidecadal Oscillation, *Geophys. Res. Lett.*, **36**, doi:10.1029/2009GL038777.

Eisenman, I., 2010: Geographic muting of changes in the Arctic sea ice cover, *Geophys. Res. Lett.*, 37, L16501, doi:10.1029/2010GL043741.

Enfield, D.B, and L. Cid-Serrano, 2010: Secular and multidecadal warmings in the North Atlantic and their relationships with major hurricane activity, *Int. J. of Climatology*, **30**, 174-184.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie. 2002, updated 2009. *Sea Ice Index*. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.

Fuller, W.A., 1987: *Measurement Error Models*. John Wiley & Sons, 440 pp.

Goosse, H., O. Arzel., C. M. Bitz, A. de Montety, M. Vancoppenolle, 2009: Increased variability of the Arctic summer ice extent in a warmer climate, *Geophys. Res. Lett.*, 36, doi:10.1029/2009GL040546

Gregory, JM; Forster, PM, 2008: Transient climate response estimated from radiative forcing and observed temperature change, *J. Geophys. Res.-Atmos.*, 113(D23), . doi:10.1029/2008JD010405

Gregory, J.M., P.A. Stott, D.J. Cresswell, N.A. Rayner, C. Gordon, and D.M.H. Sexton, 2002: Recent and future changes in Arctic sea ice simulated by the HadCM3 AOGCM, *Geophys. Res. Lett.*, 29, doi:10.1029/2001GL014575.

Hayya, J., D. Armstrong and N. Gressis ,1975: A Note on the Ratio of Two Normally Distributed Variables, *Management Science* 21 (11), 1338-1341.

doi:10.1287/mnsc.21.11.1338.

Holland, M.M., C.M. Bitz, and L.B. Tremblay, 2006: Future abrupt reductions in the summer Arctic sea ice, *Geophys. Res. Lett.*, 33, L23503, doi:10.1029/2006GLO28024.

Mahajan, S, R. Zhang, and T.L. Delworth, 2010: Impact of the Atlantic Meridional Overturning Circulation (AMOC) Variability on Arctic Surface Air Temperature and Sea-Ice Variability, *J. Climate*, submitted.

Marsaglia, G., 2006: Ratios of normal variables, *Journal of Statistical Software*, 16(4).

Min, S.-K., X. Zhang, F. W. Zwiers, and T. Agnew, 2008: Human influence on Arctic sea ice detectable from early 1990s onwards, *Geophys. Res. Lett.*, 35, L21701, doi:10.1029/2008GL035725.

Polyakov, I. V., G.V. Alekseev, R.V. Bekryaev, U.S. Bhatt, R. Colony, M.A. Johnson, V.P. Karklin, D. Walsh, and A.V. Yulin, 2003: Long-term ice variability in arctic marginal seas, *J. Climate*, 16 (12), 2078–2085.

Ridley, J., J. Lowe, C. Brierley, and G. Harris, 2007: Uncertainty in the sensitivity of Arctic sea ice to global warming in a perturbed parameter climate model ensemble, *Geophys. Res. Lett.*, 34, L19704, doi:10.1029/2007GL031209.

Stroeve, J., M. M. Holland, W. Meier, T. Scambos, and M. Serreze (2007), Arctic sea ice decline: Faster than forecast, *Geophys. Res. Lett.*, 34, L09501, doi:10.1029/2007GL029703.

Vinnikov, K.Y., A. Robock, R.J. Stouffer, J.E. Walsh, C.L. Parkinson, D.J. Cavalieri, J.F.B. Mitchell, D. Garrett, and V.F. Zakharov, 1999: Global warming and Northern Hemisphere sea ice extent. *Science*, 286, 1934-1937.

Vinnikov, K.Y., D.J. Cavalieri, and C.L. Parkinson, 2006: A Model Assessment of Satellite Observed Trends in Polar Sea Ice Extents, *Geophys. Res. Lett.*, 33, L05704, doi:10.1029/2005GL025282.

Wang, M., and J.E. Overland, 2009: A sea ice free summer Arctic within 30 years? *Geophys. Res. Lett.*, 36, doi:10.1029/2009/GL037820.

Zhang, R., T.L. Delworth, and I.M. Held, 2007: Can the Atlantic Ocean drive the observed multidecadal variability in Northern Hemisphere mean temperature? *Geophysical Research Letters*, 34, L02709, doi:10.1029/2006GL028683.

Zhang, X., 2010: Sensitivity of arctic summer sea ice coverage to global warming forcing: toward reducing uncertainty in arctic climate change projections, *Tellus*, 62A, 220-227.

## List of Tables

Table 1. Ice cover (I)/global temperature (T) sensitivity estimation methods with values for 1979-2009 changes. Sample correlations and standard deviations of the subscripted variables are denoted by  $\rho$  and  $\sigma$ , respectively.

Table 2. Test of trend ratio and OLS on CM2.1 31-year data sections with trends added (0.5 K/31 years for temperature;  $-1.5 \cdot 10^{12} \text{m}^2/31$  years for ice cover).

Table 3. Model sensitivity (from OLS) and global temperature/NH ice cover correlation for climate model annual data.

Table 4. P-values for the climate model true sensitivities using trend ratio and TLS sensitivity estimates on combined CM2.1 natural trends and observations.

## List of Figures

Fig. 1. Global mean surface temperature anomalies from 1951-1980 (left) and NH ice cover right for observations (black) and climate model simulations. The observed temperature anomalies and sea ice cover are from GISS and NSIDC, respectively. For the simulations, a short section of a projection is concatenated onto the historical run to bring the simulation up to present. The projection used is SRES A2 for all models except GFDL CM3 which uses RCP8.5.

Fig. 2. Simulated (colors) and observed (black) trends in global surface temperature and NH ice cover. The trends are from the time series shown in Fig. 1.

Fig. 3. Global temperature and NH ice cover annual means for GFDL CM2.1 (blue) and CM3 (red) projection experiments. Results are shown from a medium forcing scenario 21<sup>st</sup> century (+'s) and post-2100 with stabilized forcing (o's). Strong forcing 21<sup>st</sup> century results are shown as dots. The medium forcing scenarios is SRES A1B for CM2.1 and RCP4.5 for CM3. The strong forcing scenarios are SRES A2 and RCP8.5, respectively.

Fig. 4. Total least squares sensitivity estimate over time in observations and for more sensitive models.

Fig. 5. Trends from 100 31-year sections of the CM2.1 preindustrial control run (red). Observed trends are plotted in black. A bivariate normal fit to the natural trends is also plotted (black contours).

Fig. 6. PDF of the true sensitivity using trend ratio (green) and TLS (blue) methods from histograms (asterisks) and from simulated trends drawn from correlated bivariate normal distribution (smooth curves). Natural trends are combined with observed temperature and ice time series using equations (4a) and (4b).

Fig. 7. TLS sensitivities and North Atlantic SST (AMO) trends using 31-year sections of the GFDL CM2.1 control run.

Table 1.

<b>Estimator</b>	<b>Expression</b>	<b>Value – <math>10^{12} \text{ m}^2/\text{K}</math></b>
OLS $\Delta I$ on $\Delta T$	$\rho_{I,T} \sigma_I / \sigma_T$	-2.35
$(\text{OLS } \Delta T \text{ on } \Delta I)^{-1}$	$\rho_{I,T}^{-1} \sigma_I / \sigma_T$	-3.44
Neutral Estimate	$\text{Sign}(\rho_{I,T}) (\sigma_I / \sigma_T)$	-2.85
Trend Ratio	$(\rho_{I,t} / \rho_{T,t}) (\sigma_I / \sigma_T)$	-3.09
Total Least Squares	Eqns. 2 and 3	-3.01

Table 2.

	<b>Bias</b>	<b>St. Dev. (<math>10^{12}\text{m}^2/\text{K}</math>)</b>
Trend ratio	+3%	1.6
TLS	-5%	0.85

Table 3.

Model	Sensitivity ( $10^{12}\text{m}^2/\text{K}$ )	Correlation
GFDL CM3	-2.2	.98
NCAR CCSM3	-1.9	.99
UKMO HadGEM1	-1.6	1.0
MIROC MEDRES	-1.5	.98
MPI ECHAM5	-1.3	.98
CNRM CM3	-1.3	.98
GFDL CM2.1	-1.2	.96
CCMA CGCM3	-.78	.98

Table 4.

<b>Model</b>	<b>Sensitivity (<math>10^{12}\text{m}^2/\text{K}</math>)</b>	<b>Trend Ratio P-value</b>	<b>TLS P-value</b>
GFDL CM3	-2.2	.34	.34
NCAR CCSM3	-1.9	.22	.20
UKMO HadGEM1	-1.6	.13	.12
MIROC MEDRES	-1.5	.11	.10
MPI ECHAM5	-1.3	.08	.06
CNRM CM3	-1.3	.08	.06
GFDL CM2.1	-1.2	.07	.05
CCCMA CGCM3	-.78	.04	.02

Fig. 1. Global mean surface temperature anomalies from 1951-1980 (left) and NH ice cover right for observations (black) and climate model simulations. The observed temperature anomalies and sea ice cover are from GISS and NSIDC, respectively. For the simulations, a short section of a projection is concatenated onto the historical run to bring the simulation up to present. The projection used is SRES A2 for all models except GFDL CM3 which uses RCP8.5.

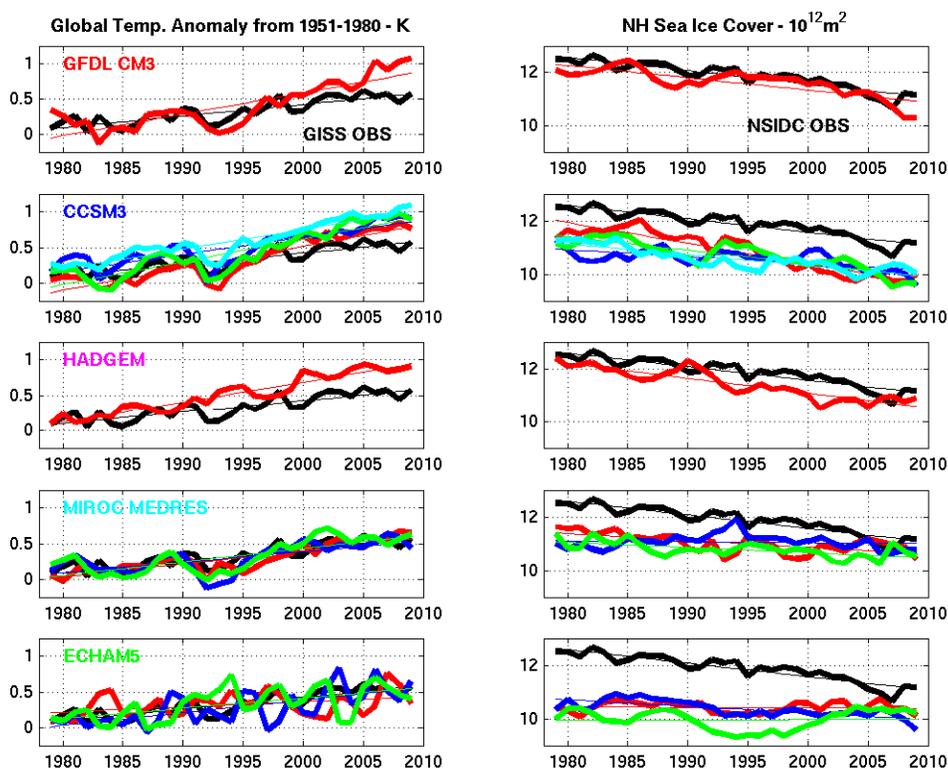


Fig. 2. Simulated (colors) and observed (black) trends in global surface temperature and NH ice cover. The trends are from the time series shown in Fig. 1.

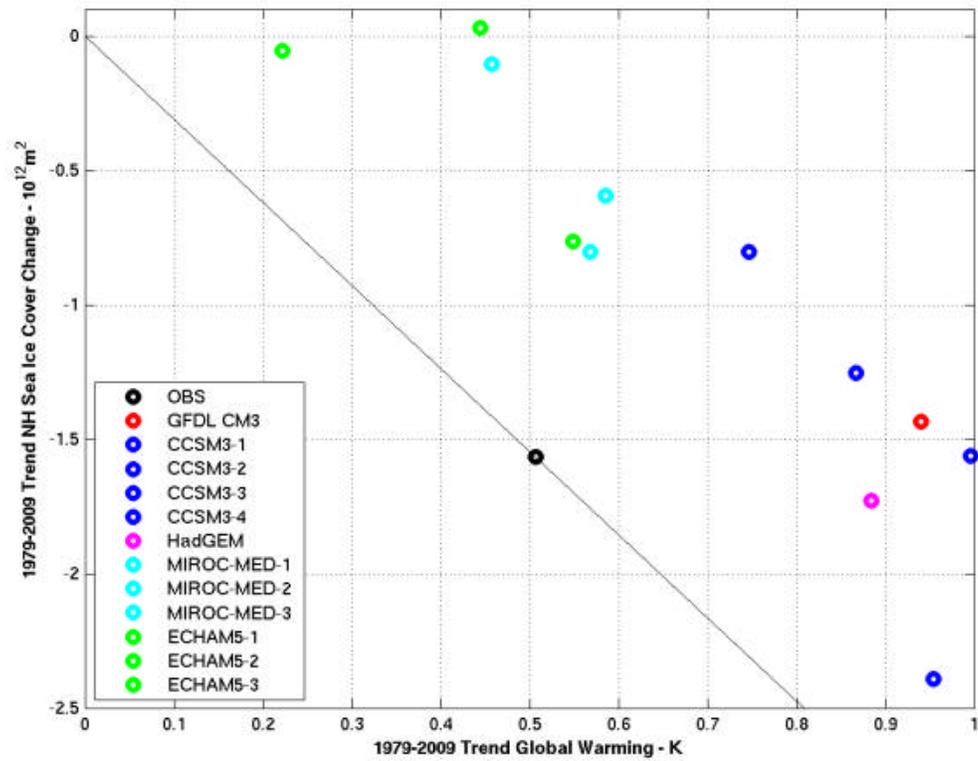


Fig. 3. Global temperature and NH ice cover annual means for GFDL CM2.1 (blue) and CM3 (red) projection experiments. Results are shown from a medium forcing scenario 21<sup>st</sup> century (+'s) and post-2100 with stabilized forcing (o's). Strong forcing 21<sup>st</sup> century results are shown as dots. The medium forcing scenarios is SRES A1B for CM2.1 and RCP4.5 for CM3. The strong forcing scenarios are SRES A2 and RCP8.5, respectively.

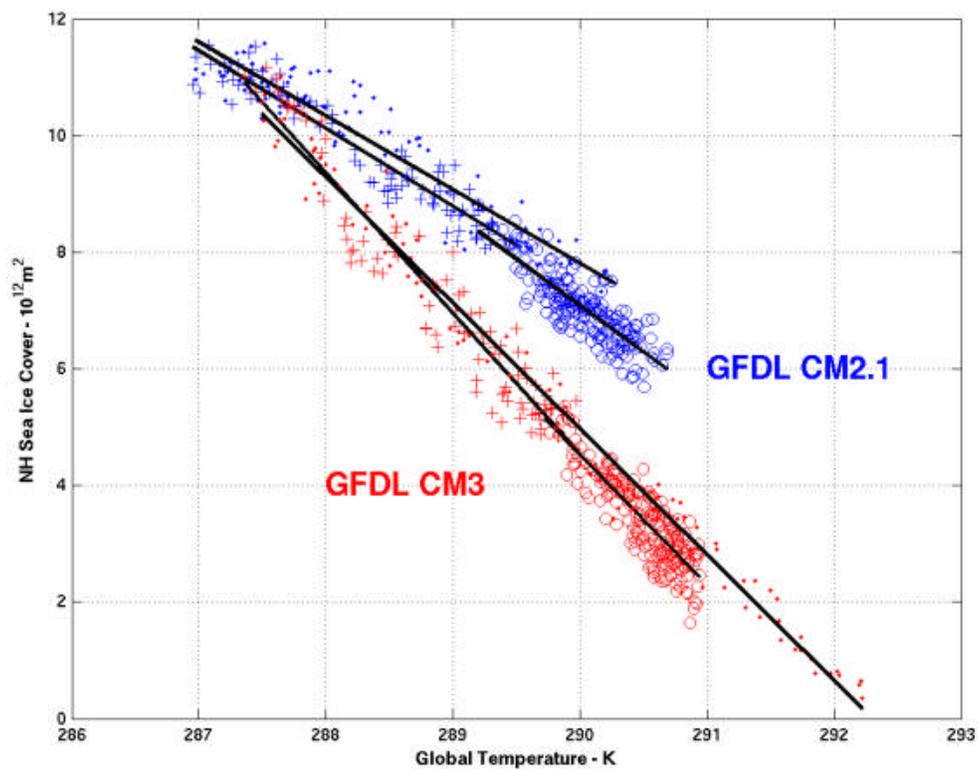


Fig. 4. Total least squares sensitivity estimate over time in observations and for more sensitive models.

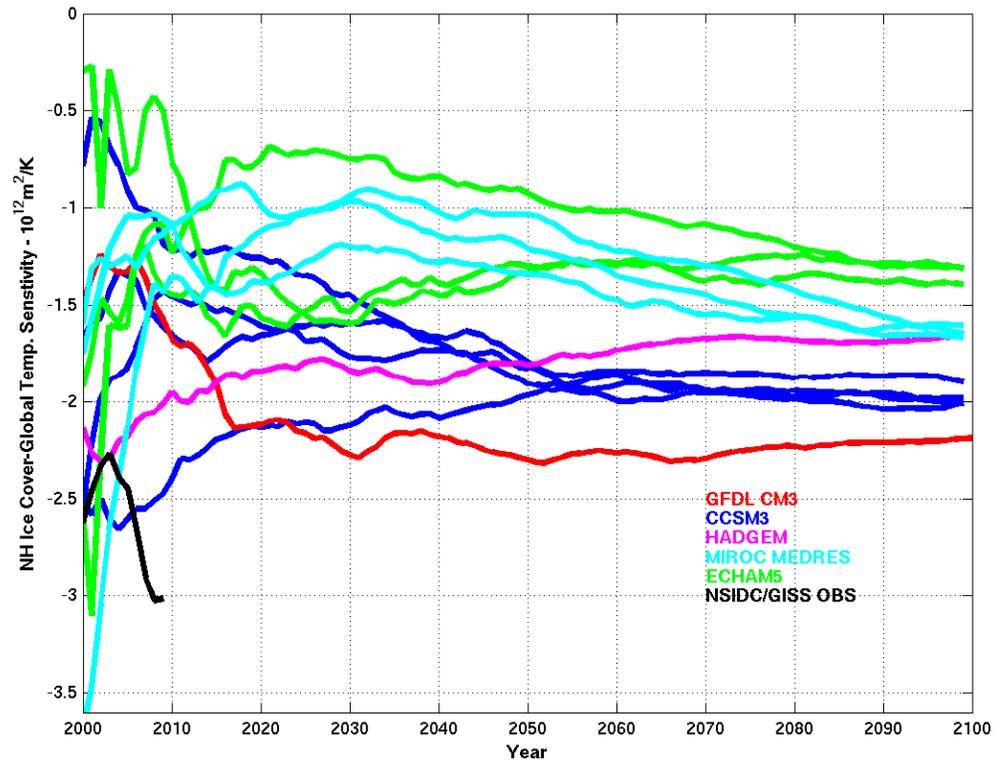


Fig. 5. Trends from 100 31-year sections of the CM2.1 preindustrial control run (red). Observed trends are plotted in black. A bivariate normal fit to the natural trends is also plotted (black contours).

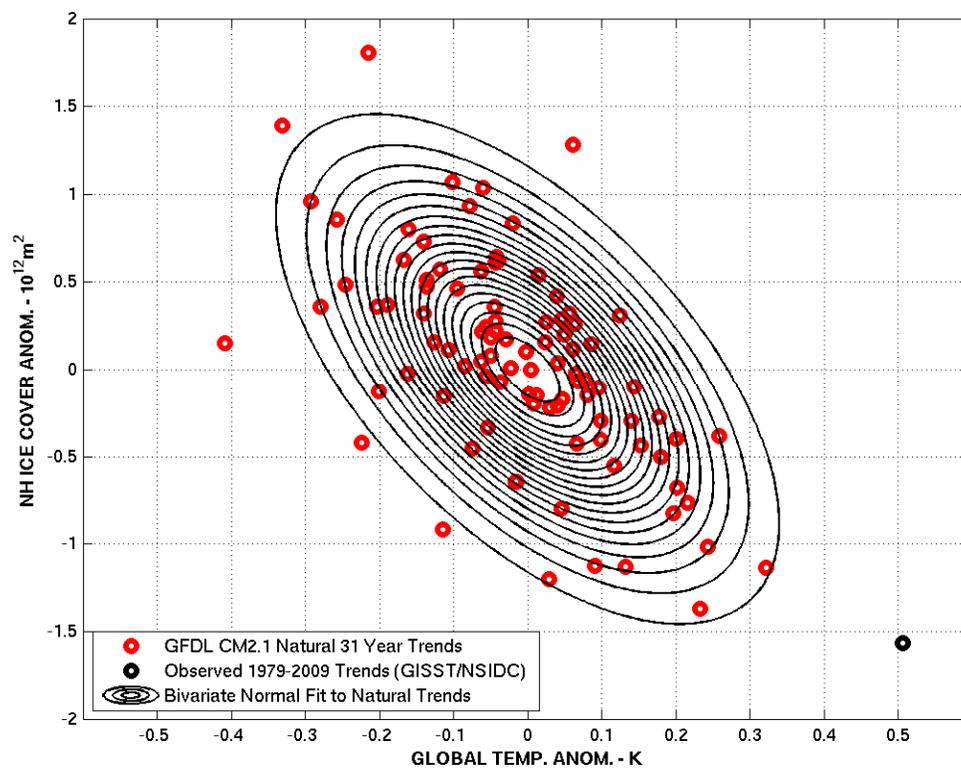


Fig. 6. PDF of the true sensitivity using trend ratio (green) and TLS (blue) methods from histograms (asterisks) and from simulated trends drawn from correlated bivariate normal distribution (smooth curves). Natural trends are combined with observed temperature and ice time series using equations (4a) and (4b).

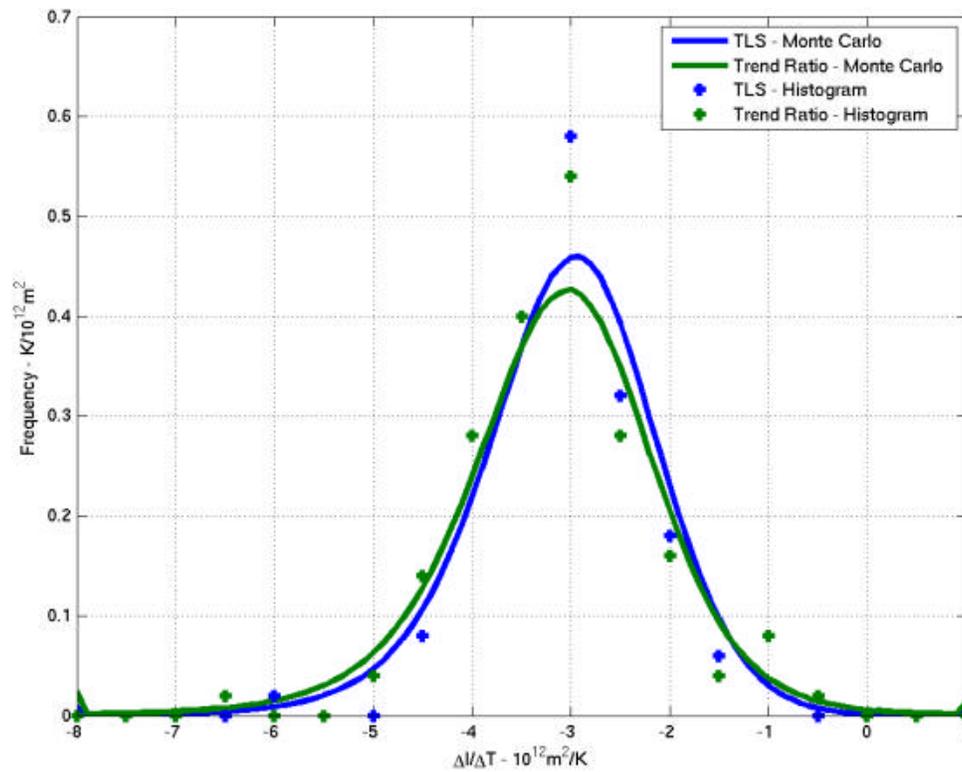


Fig. 7. TLS sensitivities and North Atlantic SST (AMO) trends using 31-year sections of the GFDL CM2.1 control run.

