Temperature and humidity biases in global climate models and their impact on climate feedbacks

V. O. John[1] and B. J. Soden[1]

Received 2 May 2007; revised 11 July 2007; accepted 13 August 2007; published 25 September 2007.

[1] A comparison of AIRS and reanalysis temperature and humidity profiles to those simulated from climate models reveals large biases. The model simulated temperatures are systematically colder by 1–4 K throughout the troposphere. On average, current models also simulate a large moist bias in the free troposphere (more than 100%) but a dry bias in the boundary layer (up to 25%). While the overall pattern of biases is fairly common from model to model, the magnitude of these biases is not. In particular, the free tropospheric cold and moist bias varies significantly from one model to the next. In contrast, the response of water vapor and tropospheric temperature to a surface warming is shown to be remarkably consistent across models and uncorrelated to the bias in the mean state. We further show that these biases, while significant, have little direct impact on the models' simulation of water vapor and lapse-rate feedbacks. **Citation:** John, V. O., and B. J. Soden (2007), Temperature and humidity biases in global climate models and their impact on climate feedbacks, *Geophys. Res. Lett.*, *34*, L18704, doi:10.1029/2007GL030429.

1. Introduction

[2] Atmospheric water vapor is widely recognized to be a key climate variable. It is the dominant greenhouse gas and provides a key feedback for amplifying the sensitivity of the climate to external forcings [*Bony et al.*, 2006; *Held and Soden*, 2000; *Soden and Held*, 2006]. Water vapor is also an important component of the hydrological cycle. Future increases in water vapor in response to a warming climate are fundamentally linked to the expected changes in moisture convergence, precipitation extremes, meridional energy transport and an overall weakening of the atmospheric circulation [*Held and Soden*, 2006].

[3] Naturally, there has been a considerable effort to assess the credibility of model simulations of atmospheric water vapor. Because of the highly variable nature of water vapor, global satellite observations are critical. In particular, a number of studies have compared satellite measurements of water vapor with those simulated by General Circulation Models (GCMs) [e.g., *Soden and Bretherton*, 1994; *Bates and Jackson*, 1997; *Allan et al.*, 2003; *Forster and Collins*, 2004; *Brogniez et al.*, 2005]. Many of these studies have identified systematic biases in the climatological distribution of water vapor simulated by various models.

[4] More recently, a study by *Pierce et al.* [2006] compared global specific humidity measurements from the AIRS instrument with that from GCMs. They found that these models display a distinct moist bias that is considerably larger than the estimated uncertainty in the AIRS data and suggest systematic deficiencies in the model simulations. However, the influence of water vapor on the climate response depends primarily on the projected perturbations to the water vapor field, rather than on the mean state. Thus the impact of these biases on the utility of the model for climate change studies remains unclear.

[5] In this study, we extend the results of *Pierce et al.* [2006] by comparing AIRS temperature and humidity retrievals to simulations from 16 different models using the archive of climate model simulations compiled for the WCRP CMIP3. We find specific humidity biases consistent with those identified by *Pierce et al.* [2006]. We also present temperature and relative humidity biases in the models. We then examine the extent to which these biases may impact the climate sensitivity of the models, based on our current knowledge of temperature and humidity feedbacks.

2. Data Sets

[6] The Atmospheric Infrared Sounder (AIRS) is a high spectral resolution radiometer with 2378 bands in the thermal infrared (3.7–15.4 m) and 4 bands in the visible (0.4–1.0 m). The AIRS was designed to provide vertical profiles of atmospheric temperature and humidity with higher accuracy and better resolution than previous satellite instruments. In this study, we used monthly mean temperature and water vapor profiles derived from AIRS version 4.0 Level 3 products. The expected accuracy of AIRS retrievals is 1 K RMS in 1 km layers below 100 hPa for air temperature and 10% RMS in 2 km layers below 100 hPa for water vapor concentration. Note that the retrievals include microwave measurements also; please refer to *Fetzer* [2006] and references therein for retrieval as well as validation details.

[7] In addition to AIRS observations, the model temperature and humidity fields are also compared with ECMWF (ERA-40) [*Uppala et al.*, 2005] and NCEP/NCAR [*Kalnay et al.*, 1996] reanalysis data to provide an independent validation source. The comparison between reanalyses and AIRS retrievals also provides an estimate of uncertainty which exists between the various observational data sets.

[8] Coupled ocean-atmosphere model outputs from the climate of the twentieth century (20C3M) of the CMIP3 archive are used here. Because the mass of water vapor is tightly coupled to temperature through the Clausius-Clapeyron (C-C) relation, the use of specified SST in GCM simulations heavily constrains the model response.

[1]Meteorology and Physical Oceanography, Rosenstiel School for Marine and Atmospheric Science, University of Miami, Miami, Florida, USA.

0094-8276/07/2007GL030429$05.00

Table 1. Vertically Integrated Global Mean Model Biases[a]

	$\int \Delta T[K]$		$\int \frac{\Delta q}{q}$ [%]		$\int \Delta RH$[%RH]		$\int \frac{\Delta q}{q}$[%] (*AIRS*)		
Model	AIRS	ECMWF	AIRS	ECMWF	AIRS	ECMWF	1000–200	850–200	1000–850
BCCR_CM2_0	−2.10	−2.09	−7.73	−11.88	12.07	2.76	−5.14	−4.59	−7.74
CNRM_CM3	−0.92	−0.91	4.83	00.67	11.82	2.51	7.05	9.10	−2.49
CSIRO_MK3_0	−1.16	−1.14	9.21	05.05	11.58	2.24	4.96	7.92	−8.83
GFDL_CM2_0	−1.85	−1.83	−3.62	−7.78	10.33	1.00	−4.82	−4.05	−8.41
GFDL_CM2_1	−1.15	−1.13	0.43	−3.73	09.53	0.19	−0.62	0.00	−3.49
GISS_MODEL_E_H	−0.19	−0.16	22.40	18.23	06.03	−3.10	20.84	24.64	3.14
GISS_MODEL_E_R	−0.90	−0.88	12.82	08.64	05.60	−3.53	11.59	14.90	−3.87
IAP_FGOALS1_0_G	−1.31	−1.29	12.78	08.62	13.23	3.93	9.89	13.69	−7.83
INMCM3_0	−1.13	−1.11	5.16	01.01	10.27	0.96	2.54	4.28	−5.58
MIROC3_2_MEDRES	−1.86	−1.84	5.78	01.63	14.19	4.84	1.89	3.36	−4.94
MPI_ECHAM5	−0.58	−0.57	13.85	09.70	13.31	4.00	13.18	16.84	−3.92
MRI_CGCM2_3_2A	−1.94	−1.92	−0.03	−4.19	10.03	0.69	−0.21	1.78	−9.50
NCAR_CCSM3_0	−0.24	−0.22	6.92	02.75	09.95	0.60	6.08	8.53	−5.36
NCAR_PCM1	−1.21	−1.19	7.44	03.28	11.96	2.62	7.15	10.07	−6.46
UKMO_HADCM3	−0.04	−0.03	6.74	02.58	09.84	0.54	7.06	8.58	−0.04
UKMO_HADGEM1	−0.90	−0.89	−5.18	−9.34	07.72	−1.60	−5.30	−4.89	−7.24
Mean Model	−0.86	−0.84	05.68	01.50	10.55	1.19	4.70	6.88	−4.75

[a]Columns from 2 to 7 presents vertically integrated global mean biases ($\int \Delta x$, where $\Delta x = x_{\text{model}} - x_{\text{obs}}$ and x is T, q or RH; note Δq is in percent) from 1000 to 100 hPa for T [K], q [%], and RH [%RH] for all 16 models and the mean model with respect to AIRS and ECMWF. Columns from 8 to 10 show $\int \frac{\Delta q}{q}$ with respect to AIRS for different layers of the atmosphere.

For this reason, we consider only fully coupled ocean-atmosphere model simulations in which the ocean temperature is a predicted quantity. The 16 coupled models used here are listed in Table 1.

3. Results and Discussion

[9] We evaluate model simulations of the vertical distribution of zonal-mean, annual-mean climatology of temperature and water vapor. Comparison of model data and reanalysis data is done for the same time period: 1990–1999. Since AIRS data are only available starting from August 2002, the comparison between models and AIRS data is only done for four years (August 2002–July 2006). However, previous studies [*Pierce et al.*, 2006] indicate that the shorter record of AIRS data does not give rise to significant sampling errors when considering the zonal, annual mean climatology.

[10] In this section we compare the ensemble mean of the 16 CMIP3 models with three different "observational" estimates: AIRS, NCEP, and ECMWF. We present latitude-height cross sections of differences between the ensemble-mean model and observations. Plots for individual models are given in Figures S1 and S2 in the auxiliary material.[1] Differences in temperature and relative humidity are shown in absolute values, i.e., $\Delta T = T_{\text{model}} - T_{\text{obs}}$ and $\Delta\text{RH} = \text{RH}_{\text{model}} - \text{RH}_{\text{obs}}$. The difference in q is expressed in percent, i.e., $\Delta q = 100 \times \frac{q_{\text{model}} - q_{\text{obs}}}{q_{\text{obs}}}$. Vertical integrals of global mean biases for each model are given in Table 1. They are calculated as $\int \Delta x = \frac{\int \Delta x \; dp}{\int dp}$, where x = T, q, or RH.

3.1. Temperature

[11] Latitude-height cross section of ΔT is shown in row 1 of Figure 1. On average, most models exhibit a cold bias of 1–2 K throughout the vast majority of the troposphere. The biases tend to be larger in the extratropics than the tropics and are most prominent in the Northern Hemisphere, perhaps reflecting a cold bias over land regions. Another notable feature is the cold bias near 200 hPa in both hemispheres outside the tropics. In this region, the temperature difference exceeds 6 K with respect to all 3 observational data sets. The prevalence of the upper-tropospheric cold bias across models and observational data sets suggests a common source of error in the models. *Gates et al.* [1999] also noted this cold bias in AMIP-I models. They attributed the bias to errors in the model's advection schemes. Warm biases in the boundary layer over southern polar region might be due to retrieval errors in AIRS data as a result of surface contamination.

[12] This general pattern of cold bias is fairly robust across models (see Figure S1). A measure of the robustness is provided in rows 3 and 4 of Figure 1 which show the number of models in which the simulated variable (e.g. T or q) is larger than the observed value. Values at either extreme of this range (all models or few models), indicate a robust positive or negative bias respectively. While all models exhibit some degree of cold bias in the troposphere, the magnitude of the bias varies significantly from model to model. Illustration of the inter-model variability is provided in Table 1 which lists the globally-averaged, vertically-integrated bias from the surface to 100 mb. In this gross measure, all models show a cold bias ranging from −0.04 to −2.10 K with the multi-model ensemble mean being near the center of this range (Table 1). Below we will show that these biases in the base climatology have little impact on the response of tropospheric temperatures to increasing greenhouse gases.

3.2. Specific Humidity

[13] The fractional error for q is also presented in Figure 1 (row 2) for the ensemble mean of all models and the number of models with a positive bias (row 4). Because the absorption of longwave radiation by water vapor is propor-

[1]Auxiliary materials are available in the HTML. doi:10.1029/2007GL030429.

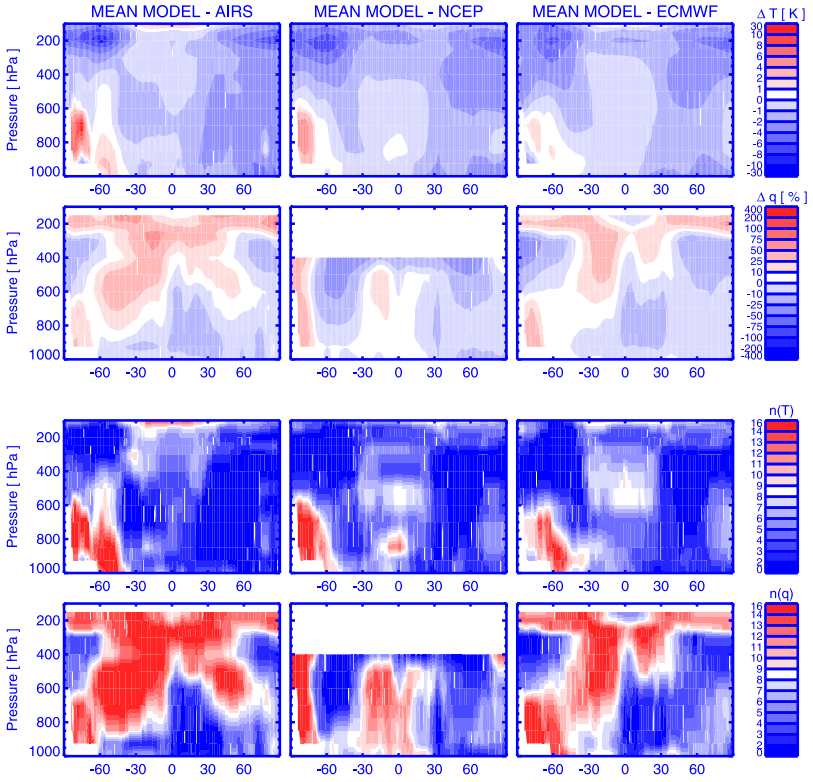

Figure 1. Vertical structure of difference in zonal averaged fields: (row 1) $\Delta T = T_{\text{model}} - T_{\text{obs}}$ and (row 2) $\Delta q = 100 \times \frac{q_{\text{model}} - q_{\text{obs}}}{q_{\text{obs}}}$. Note that in NCEP data set humidity values are not given above 300 hPa. These quantities for each coupled model are shown in Figures S1 and S2. Rows 3 and 4 show the number of models in which the simulated variable (T and q, respectively) is larger than the observed value. Values at either extreme of this range (all models or few models), indicate a robust positive or negative bias respectively.

tional to logarithm of the amount of water, the fractional error in water vapor provides a better measure of its impact on the radiative transfer than does the absolute error [*Soden et al.*, 2005]. Compared to both AIRS and ECMWF, the multi-model ensemble mean tends to show a dry bias in the lower troposphere (below 800 mb) and wet bias in the upper troposphere. This basic pattern of too much moisture in the free troposphere and too little moisture in the boundary layer is a common feature of many models (see Figure S2). This is also consistent with the analysis presented by *Pierce et al.* [2006]. In the lower troposphere (below 800 mb), the models' dry bias rarely exceeds 25%, which is about the expected retrieval precision of AIRS water vapor data. But the wet bias in the upper troposphere approaches 75% for the multi-model mean and can exceed 200% for some individual models (as seen in Figure S2).

[14] Table 1 presents vertically integrated global mean bias ($\int \frac{\Delta q}{q}$) from 1000 to 100 hPa for all models compared to AIRS and ECMWF. The NCEP reanalyses report humidity values only up to 300 hPa therefore vertically integrated biases are not reported in Table 1. $\int \frac{\Delta q}{q}$ shows large variability, ranging from −7.73 to 22.40%, indicating that the mean state of water vapor in the models differs significantly from one model to the next. To help discriminate between boundary layer and free tropospheric biases, Table 1 also presents $\int \frac{\Delta q}{q}$ computed with respect to AIRS for the 1000 to 850 hPa and 850 to 200 hPa layers. In the 1000–850 hPa layer $\int \frac{\Delta q}{q}$ is negative for all but one model, highlighting the prevalence of a dry bias in the boundary layer in current GCMs. The magnitude of this bias is relatively modest (typically less than 10% on a global mean) and does not vary substantially from model to model. In contrast the free tropospheric layer between 850–200 hPa is characterized by a much larger moist bias whose magnitude varies significantly among models. Of the models considered here, all but three contain a moist bias in the free troposphere relative to AIRS whose magnitude varies from 0 to >20%.

[15] It should be noted that the magnitude of the moist bias is reduced when using ECMWF analyses rather than AIRS retrievals, which may be indicative of a dry bias in the AIRS data. However, the basic structure of the model-inferred biases is very similar between the two observational data sets. Although, the structure of the biases with respect to NCEP differs noticeably from those in both AIRS and ECMWF (e.g., the dry bias in the extra-tropical upper troposphere). The similarity between the ECMWF and AIRS structure of errors suggests that this reflects a deficiency in the NCEP humidity distribution.

[16] The presence of both a cold and a moist bias in the free troposphere implies that the relative humidity is too high in most models. This is confirmed in Table 1 which shows the vertically integrated bias from 1000 to 100 hPa. The moist bias is larger with respect to AIRS due to dry bias

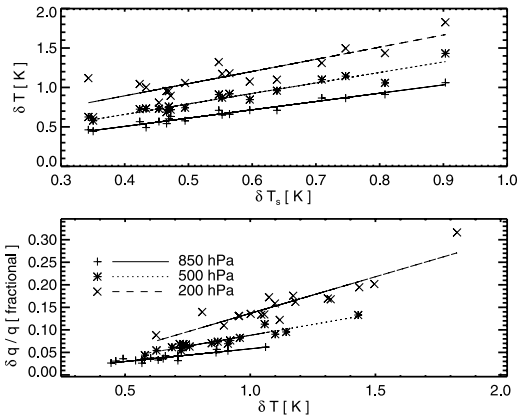

Figure 2. (top) Response of T at three different atmospheric levels (850, 500, and 200 hPa) to change in surface temperature (T_s). (bottom) Fractional response of q at three different atmospheric levels (850, 500, and 200 hPa) to change in T at those levels. Different symbols represent different coupled GCMs used in this study. Tropical means are used. δT, δT_s, and $\frac{\delta q}{q}$ are the difference between the first 10 year and the last 10 year means of 20th century of each variable.

in RH data calculated from AIRS (see auxiliary material for details).

3.3. Temperature and Humidity Response

[17] The above comparisons have identified significant biases in the ability of current climate models to simulate the zonal, annual mean distribution of water vapor and temperature. However, the influence of water vapor and temperature on the climate response depends primarily on the projected perturbations to these fields, rather than on their mean state. Thus the impact of these biases on the utility of the model for climate change studies remains unclear.

[18] In this section we analyze the impact of T and q biases in the mean state of the models on the response of these fields to an external forcing. Figure 2 (top) shows the response of T at three different levels of the atmosphere (850, 500, and 200 hPa) to change in surface temperature (T_s) for different coupled models (each symbol denotes a separate model). The scenario used here is the 20th century simulations (20C3M) and the data are averaged over the tropics (±30° latitudes). The changes are calculated by subtracting the first 10 years mean from the last 10 years mean of 20th century. Despite the large range of biases between models (see Table 1 and Figures S1 and S2), they all show a very similar rate of air temperature response at each level to a change in T_s. One can also see that the models predict an amplified rate of warming in the higher levels of the troposphere over that at the surface, i.e., a negative lapse-rate feedback [*Bony et al.*, 2006]. The slopes of least-square fit lines ($\frac{\partial \delta T}{\partial \delta T_s}$) are 1.05, 1.33, and 1.54 K/K at 850, 500, and 200 hPa, respectively. These results are consistent with the projected changes in lapse rate associated with a moist-adiabatic response to surface warming [*Santer et al.*, 2005].

[19] Figure 2 (bottom) shows fractional change in q with T (q and T for same level). Fractional increase in water vapor increases with altitude. The slopes, $\frac{\partial(\delta q/q)}{\partial \delta T}$, are 0.06, 0.10, and 0.16 K^{-1} at 850, 500, and 200 hPa, respectively. The amplified response at upper levels is consistent with a constant relative humidity moistening of the atmosphere. It stems from both the larger rate of warming simulated at these levels and the more rapid fractional increase of saturation vapor pressure at colder temperatures [*Held and Soden*, 2000]. Thus, despite large inter-model difference in q bias, models show highly robust q response to change in T which closely follows that expected from an assumption of constant relative humidity.

[20] The robust responses of tropospheric temperature and water vapor in the models irrespective of the mean state biases imply that lapse-rate and water vapor feedbacks simulated by the models are not sensitive to these biases. To illustrate this we show in Figure 3 the vertically integrated, global mean temperature (Figure 3, top) and fractional q (Figure 3, bottom) biases versus lapse-rate feedback (λ_T) and water vapor (λ_{H_2O}) feedback, respectively, simulated by the models. As expected, the strength of feedbacks simulated by the models does not depend on the bias in their mean states. Thus, the robustness of temperature and water vapor responses in the models suggests that their climate sensitivity is not affected by mean state biases in temperature and humidity fields.

4. Summary and Conclusions

[21] We assessed zonal averaged temperature and humidity fields simulated by coupled global climate models using those derived from AIRS measurements and also ECMWF and NCEP reanalysis data sets. Though models show systematic and consistent patterns of bias in tropospheric temperature and humidity compared to observations, we showed that these mean state biases do not impact lapse-rate or water vapor feedbacks.

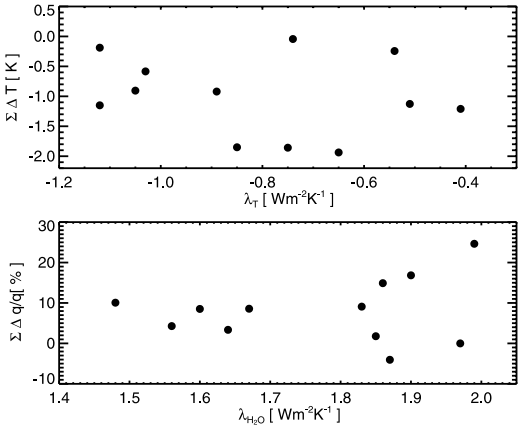

Figure 3. (top) Vertically integrated global mean T biases from 1000 to 100 hPa in the models versus temperature lapse-rate feedback (λ_T). (bottom) Vertically integrated fractional q biases from 850 to 200 hPa versus water vapor feedback (λ_{H_2O}). The feedback values simulated by the models are taken from *Soden and Held* [2006].

[22] Models in general have a cold bias in the troposphere. The vertically integrated global mean T bias with respect to AIRS ranges from -0.04 to -2.10 K and is consistent with reanalysis data sets. This bias generally increases with altitude in the free troposphere, with maxima located near 200 hPa in the extra-tropics where the bias exceeds 6 K compared to all three observational data sets.

[23] On average, models simulate less water vapor in the boundary layer (below 850 hPa) where the fractional bias in q is up to 25%. Although this discrepancy is within the expected accuracy of AIRS data, the bias in the boundary layer is robust across models with all but one model exhibiting a dry bias here. On the other hand, models simulate a much larger moist bias in the free troposphere. For the multi-model mean, the moist bias approaches 75% in the upper troposphere, but it can exceed 200% for individual models, reflecting the large intermodel variability. Among the models considered here, the vertically integrated free-tropospheric bias (from 850 to 200 hPa) varies from -4.89 to 24.64%. Given that the free troposphere is both too cold and too moist compared to observations, the mean biases are largely attributable to model deficiencies in simulating the climatological distribution of relative humidity. The reasons for these biases are unclear, we suggest detailed model sensitivity studies in future.

[24] To investigate the consequences of biases in models mean state on the utility of these models for climate change studies, we examined the dependence of the models temperature and moisture response to the magnitude of the bias in the mean state. We find that the model-simulated response of tropospheric temperature and water vapor to change in surface temperature is insensitive to biases in the mean state and show that water vapor and lapse-rate feedbacks in these models are uncorrelated to the base-state biases in water vapor and temperature.

[25] **Acknowledgments.** We thank AIRS team for their data. We acknowledge the modeling groups for making their model output available as part of the WCRP's CMIP3 multi-model dataset, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving this data, and the WCRP's Working Group on Coupled Modeling (WGCM) for organizing the model data analysis activity. The WCRP CMIP3 multi-model dataset is supported by the Office of Science, U.S. Department of Energy. This work was supported by grants from the NASA Energy and Water Cycle Study and the NASA Modeling Analysis and Prediction Program.

References

Allan, R. P., M. A. Ringer, and A. Slingo (2003), Evaluation of moisture in the Hadley Centre climate model using simulations of HIRS water vapour channel radiances, *Q. J. R. Meteorol. Soc.*, *129*, 3371–3389.

Bates, J. J., and D. L. Jackson (1997), A comparison of water vapor observations with AMIP-I simulations, *J. Geophys. Res.*, *102*(21), 21,837–21,852.

Bony, S., et al. (2006), How well do we understand and evaluate climate change feedback processes?, *J. Clim.*, *19*(15), 3445–3482.

Brogniez, H., R. Roca, and L. Picon (2005), Evaluation of the distribution of subtropical free tropospheric humidity in AMIP-2 simulations using METEOSAT water vapor channel data, *Geophys. Res. Lett.*, 32, L19708, doi:10.1029/2005GL024341.

Fetzer, E. J. (2006), Preface to special section: Validation of Atmospheric Infrared Sounder Observations, *J. Geophys. Res.*, *111*, D09S01, doi:10.1029/2005JD007020.

Forster, P., and M. Collins (2004), Quantifying the water vapour feedback associated with post-pinatubo cooling, *Clim. Dyn.*, *23*, 207–214.

Gates, W. L., et al. (1999), An overview of the results of the Atmospheric Model Intercomparison Project (AMIP-I), *Bull. Am. Meteorol. Soc.*, *80*, 29–55.

Held, I. M., and B. J. Soden (2000), Water vapor feedback and global warming, *Annu. Rev. Energy Environ.*, *25*, 441–475.

Held, I. M., and B. J. Soden (2006), Robust responses of the hydrological cycle to global warming, *J. Clim.*, *19*(21), 5686–5699.

Kalnay, E., et al. (1996), The NMC/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Pierce, D. W., T. P. Barnett, E. J. Fetzer, and P. J. Gleckler (2006), Three-dimensional tropospheric water vapor in coupled climate models compared with observations from the AIRS satellite system, *Geophys. Res. Lett.*, *33*, L21701, doi:10.1029/2006GL027060.

Santer, B. D., et al. (2005), Amplification of surface temperature trends and variability in the tropical atmosphere, *Science*, *309*, 1551–1556.

Soden, B. J., and F. P. Bretherton (1994), Evaluation of water vapor distribution in general circulation models using satellite observations, *J. Geophys. Res.*, *99*(D1), 1187–1210.

Soden, B. J., and I. M. Held (2006), An assessment of climate feedbacks in coupled atmosphere-ocean models, *J. Clim.*, *19*(14), 3354–3360.

Soden, B. J., D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, and X. Huang (2005), The radiative signature of upper tropospheric moistening, *Science*, *310*, 841–844.

Uppala, S. M., et al. (2005), The ERA-40 reanalysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.

V. O. John and B. J. Soden, Meteorology and Physical Oceanography, Rosenstiel School for Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, FL 33149, USA. (vojohn@rsmas.miami.edu)