Click Here for Full Article

A strict test in climate modeling with spectrally resolved radiances: GCM simulation versus AIRS observations

Yi Huang,[1] V. Ramaswamy,[2] Xianglei Huang,[3] Qiang Fu,[4] and Charles Bardeen[5]

[1] The spectrally resolved infrared radiances observed by AIRS provide a strict and insightful test for general circulation models (GCMs). We compare the clear- and total-sky spectra simulated from the Geophysical Fluid Dynamics Laboratory GCM using a high resolution radiation code with the AIRS observations. After ensuring consistency in the sampling of the observed and simulated spectra and a proper representation of clouds in the radiance simulation, the observed and simulated global-mean radiances are shown to agree to within 2 K in the window region. Radiance discrepancies in the water vapor $v2$ ($1300-1650$ cm^{-1}) and carbon dioxide $v2$ ($650-720$ cm^{-1}) bands are consistent with the model biases in atmospheric temperature and water vapor. The existence of radiance biases of opposite signs in different spectral regions suggests that a seemingly good agreement of the model's broadband longwave flux with observations may be due to a fortuitous cancellation of spectral errors. Moreover, an examination of the diurnal difference spectrum indicates pronounced biases in the model-simulated diurnal hydrologic cycle over the tropical oceans, a feature seen to occur in other GCMs as well. **Citation:** Huang, Y., V. Ramaswamy, X. Huang, Q. Fu, and C. Bardeen (2007), A strict test in climate modeling with spectrally resolved radiances: GCM simulation versus AIRS observations, *Geophys. Res. Lett.*, *34*, L24707, doi:10.1029/2007GL031409.

1. Introduction

[2] Since the launch of the NASA Aqua satellite in 2002, the Atmospheric Infrared Sounder (AIRS) instrument onboard has recorded an essential part of the outgoing longwave radiation (OLR) spectrum for over four consecutive years [*Chahine et al.*, 2006]. This data set has provided an invaluable reference for the outgoing infrared radiances. However, despite a few pioneering studies [e.g., *Haskins et al.*, 1997; *Huang and Yung*, 2005], it remains a largely unanswered question as to how such unprecedented spectrally resolved radiances, in addition to retrieval products, can be used to verify and ultimately improve climate models.

[3] A comparison of model and satellite spectra is as challenging as examining a variety of geophysical variables (temperature, water vapor, clouds, etc.) simultaneously, because the top of the atmosphere (TOA) radiances at different frequencies are sensitive to the different conditions prevailing at various levels. Thus, the ability of General Circulation Models (GCMs) to replicate the observed spectra provides a strict metric for model validation. Moreover, such a comparison offers more information than the broadband OLR flux, in that the radiance discrepancy, if it exists, should shed more insights into the physical cause(s) of the model's biases.

[4] Aqua is a sun-synchronous satellite which has equatorial crossing times at 1:30 a.m. and 1:30 p.m. (local times). This twice-per-day sampling enables two aspects of the OLR spectrum to be investigated: the mean and the difference at these two times. A direct comparison between observed and simulated mean spectra demonstrated significant discrepancies, especially in the total-sky case [*Haskins et al.*, 1997; *Huang et al.*, 2002]. Such a discrepancy could be as large as 10 K in the window region even though observed sea surface temperatures (SSTs) were used in the simulation; this has been inferred to be due to an unrealistic treatment of clouds. The diurnal radiance difference was not investigated in the early studies, but it is known that climate models generally do a poor job in simulating the diurnal hydrologic cycle [*Lin et al.*, 2000; *Yang and Slingo*, 2001; *Tian et al.*, 2004]. There exist large errors in both the phase and amplitude of the model-simulated diurnal cycle; interestingly, these errors are worse over the tropical oceans than over land [*Tian et al.*, 2004]. In this paper, outgoing radiances are simulated from the outputs of the state-of-the-art Geophysical Fluid Dynamics Laboratory (GFDL) GCM using a high resolution radiation code. The simulated radiances, sampled in the same manner as the satellite, are compared to the observation with respect to the above two issues.

2. Data and Model
2.1. Satellite Data

[5] AIRS is a grating spectrometer with 2378 channels ranging from wavenumber 650 to 2665 cm^{-1} and a nominal spectral resolution of $v/\Delta v = 1200$. The instrument has a nadir footprint of 13.5 km; and the observed radiance has a precision of 0.5 K at 250 K [*Chahine et al.*, 2006]. The AIRS scan angles range from $-49°$ to $49°$; however, we use the approach by *Huang and Yung* [2005], and only select the views within $\pm5°$ and treat them as nadir views. Among its channels, only those recommended for retrieval products and not subject to various static noises are used; dynamic noise flags are also used to eliminate spectra affected by

[1]Program of Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA.
[2]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA.
[3]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA.
[4]Department of Atmospheric Sciences, University of Washington, Seattle, Washington, USA.
[5]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado, USA.

random in-flight noise (E. Olsen, AIRS/AMSU/HSB version 4.0 L1B QA quick start, 2005, available at http://disc.gsfc. nasa.gov/AIRS/documentation/v4_docs/V4.0_L1B_QA_ Quick_Start.pdf). As a result, 1026 out of 1862 channels within our interested spectral range (650–1650 cm^{-1}) have been used. The AIRS L1B calibrated and L2 "cloud-cleared" radiances [*Chahine et al.*, 2006] are used as "total-sky" and "clear-sky", respectively. Over 1 billion AIRS spectra observed in the first operational year (September 2002 to October 2003) are used for this study.

[6] We also use the broadband longwave flux data from the Clouds and the Earth's Radiant Energy System (CERES) onboard Aqua (the Single Scanner Footprint product [*Loeb et al.*, 2005]).

2.2. Model

[7] Spectra are simulated from the profiles generated by the Geophysical Fluid Dynamics Laboratory (GFDL) AM2 (version AM2p13). AM2 is an atmospheric general circulation model with $2.5° \times 2°$ horizontal resolution. It uses a 24 vertical layer hybrid coordinate, with the top layer at about 3 hPa (about 35 km). The model is integrated over the same period as the selected AIRS data (September 2002 to August 2003), with prescribed SSTs from observations (J. W. Hurrell et al., A new sea surface temperature and sea ice boundary data set for the Community Atmosphere Model, submitted to *Journal of Climate*, 2006), and with well mixed greenhouse gas concentrations and ozone [*Randel and Wu*, 2007; *Horowitz*, 2006] distributions as appropriate to the time. Clouds at different vertical layers are assumed to be randomly overlapped. The SSTs do not vary diurnally, but the predicted land surface temperature does vary [*GFDL Global Atmospheric Model Development Team* (*GFDL GAMDT*), 2004].

[8] The 3-hourly instantaneous profiles of temperature, water vapor, cloud liquid, cloud ice and cloud fraction are archived. These profiles, patched with 5 layers from the standard atmosphere [*McClatchey et al.*, 1972] from 50 to 100 km at the top, are then fed into a radiative transfer model for the radiance simulation. The radiative transfer model used here is MODTRAN version 4 [*Berstein et al.*, 1996]. This model adopts a very narrow band parameterization with a computational resolution of 1 cm^{-1}. The cloud optical properties and the random overlapping assumption are treated consistently as in the AM2 radiation scheme. Because MODTRAN itself is not capable of implementing cloud overlapping schemes, the total-sky transmission function (T_{tot}) of each model layer at each frequency is combined manually from the MODTRAN computed cloudy (T_{cld}) and clear (T_{clr}) transmission functions ($T_{tot} = T_{cld} * A + T_{clr} * (1 − A)$, where A is the AM2 computed cloud fraction), and then is used to calculate the TOA radiances. To verify the consistency between the above radiative transfer computation and AM2's radiation scheme, broadband OLR fluxes are compared. For a randomly select

month (January 2003) with 137696 profiles of global coverage, the bias of OLR flux between the two simulations is -0.23 W m^{-2}, with a root-mean-square of 8.8 W m^{-2}, indicating an excellent agreement.

[9] In addition, auxiliary OLR fluxes from NCEP reanalysis [*Kalnay et al.*, 1996] and integrations of the NCAR Community Atmospheric Model (CAM3) [*Collins et al.*, 2004] are studied.

3. Model-Satellite Comparison

[10] We focus our investigation on the spectral range from 650 to 1650 cm^{-1}. This subset of the spectrum includes most of the CO_2 $v2$ band (650–720 cm^{-1}), the center of which is sensitive to the stratosphere; the atmospheric window region (833–985 cm^{-1} and 1085–1250 cm^{-1}) which is sensitive to surface, clouds and water vapor continuum absorption; the ozone $v3$ band (985–1085 cm^{-1}) which is affected by surface, troposphere and lower stratosphere; and majority of the water vapor $v2$ band (1300–1650 cm^{-1}) which is sensitive to the middle and upper tropospheric temperature and moisture.

[11] In the 650–1650 cm^{-1} spectral range, the spectral resolutions of AIRS channels gradually change from about 0.5 cm^{-1} to 1.3 cm^{-1} due to the fixed resolving power of $v/\Delta v = 1200$. Within this spectral range, AIRS does not cover some subintervals and occasionally some subintervals are covered twice by different channels. In order to make meaningful comparisons, both the observed and simulated spectra are degraded to 2 cm^{-1} resolution spectral grids by convolving with an appropriate triangular scanning function with a full-width-half-maximum of 2 cm^{-1}.

[12] To reduce the temporal sampling inconsistency, satellite tracking is applied, i.e., only those grid points overpassed by the Aqua satellite are selected in the radiance simulations. Averages are computed using this set of subsampled simulations, and compared to the averages from their observational counterparts. This ensures that the simulated spectra have similar spatial and temporal sampling pattern as the AIRS spectra, an essential criterion for the evaluation of the model. Also, in avoiding the uncertainties due to the land surface emissivity, we focus on ocean grid points only, where the MODTRAN default oceanic surface spectral reflectance is used.

[13] In the following section, the model simulation and the satellite observation are compared in terms of the mean and difference of the spectra at the two observational times (1:30 a.m. and 1:30 p.m. local crossing time at the Equator).

3.1. Mean Spectrum

[14] The model simulated global-mean spectra over oceans for clear- and total-skies are compared to the AIRS observation in Figure. In the clear-sky case, the window region discrepancy is about 1 K in terms of brightness temperature (hereafter, T_b). Considering the fact that SSTs

Figure 1. (a) Clear-sky global mean spectra (The major frequency bands discussed here are labeled on the plot). (b) Clear-sky model-satellite difference spectrum. (c) Error budget in the water vapor $v2$ band in b ("Jacobian" represents the model bias reproduced by using radiative Jacobians, which is further decomposed into contributions from the temperature (T) bias and the specific humidity (q) bias. See text and *Huang et al.* [2007]. (d) Total-sky global mean spectra. (e) Total-sky model-satellite difference spectrum.
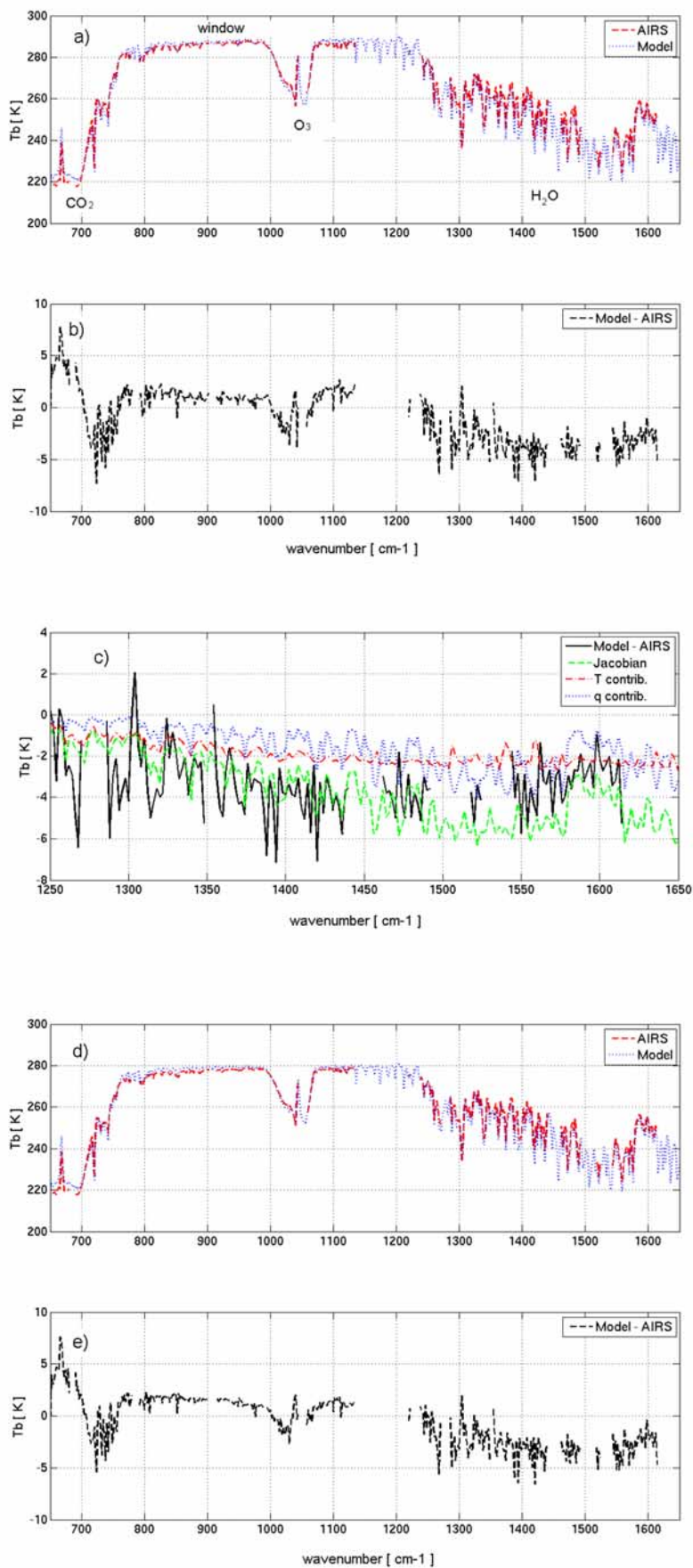
Figure 1

Table 1. Four-Month Mean Broadband Fluxes From CERES Observation and AM2 Simulation[a]

AM2-CERES	Total Sky	Clear Sky
OLR (50–2000 cm^{-1}), W m^{-2}	−1.10	−12.44
Window band (833–1250 cm^{-1}), W m^{-2}	7.05	4.28

[a]Months are January/April/July/October 2003.

are prescribed with the observed values, this is indicative of the quality of the simulated lower tropospheric temperature and water vapor. In the total-sky case where the window region T_b is affected by clouds, the model bias is about 2 K. An examination of several regions such as the tropical western Pacific and the tropical central Pacific exhibits discrepancies of the same sign and with similar amplitudes (not shown).

[15] The model has a prominent negative (cold) bias in the water vapor $v2$ band. When it is decomposed into different latitude belts, this cold bias is persistent in all latitudes, and is largest in the tropics (not shown). A direct comparison of temperature with the NCEP reanalysis data indicates a similar bias pattern and amplitude as shown by *GFDL GAMDT* [2004]. In terms of the global mean, the AM2 temperature profile has a −2 K bias, averaged from 100 hPa to 700 hPa, with a maximum of −3 K occurring at around 300 hPa. The AM2 simulated moisture field is compared to the NCEP reanalysis (up to 300 hPa) and the AIRS retrieval product (above 300 hPa), notwithstanding the uncertainties in each of these data sets. The global mean specific humidity profile has positive biases in most layers, about 10% between 400 and 700 hPa, and up to 150% between 100 and 400 hPa. The sign of the moisture bias is consistent with *Huang et al.* [2006]. By using the spectral Jacobians of temperature (dR/dT) and water vapor (dR/dq) as computed by *Huang et al.* [2007] and the above described temperature and moisture biases, the radiance biases can be approximately reproduced (Figure 1). It is evident that both the temperature and moisture biases contribute substantially to the radiance biases.

[16] The cold temperature bias in the model's troposphere is also indicated in the wing of the CO_2 band (700–760 cm^{-1}), while the warm bias in the model's stratosphere [*GFDL GAMDT*, 2004] emerges as a prominent discrepancy in the center of the CO_2 band (around 667 cm^{-1}). The biases described above are persistent in all the months of data analyzed (not shown).

[17] The results above, particularly the biases of opposite signs in the window region and the water vapor band, bring up a significant question about the consequences of model tuning. If a model has an intrinsic upper tropospheric bias as in AM2, then, in tuning the model to match the observed broadband flux, the underestimation of outgoing radiation in the water vapor band would have to be compensated by an overestimation in some other spectral region(s). Apparently, this is so in AM2: the cold bias in the water vapor band is offset by the warm bias in the window region. It must be noted that the radiation from the window and the water vapor rotational band (0–560 cm^{-1}) comprises the bulk of the OLR; further, the outgoing radiation in the water vapor rotational band, which is not observed by AIRS, is sensitive to the middle and upper troposphere, and is

positively correlated with the radiation in the water vapor $v2$ band that is observed by AIRS.

[18] To test the hypothesis regarding the cancellation of errors, the 4-month mean (January, April, July and October, 2003) broadband longwave flux from the AM2 simulation is compared to the CERES observation (Table 1). In spite of a good agreement (−1.1 W m^{-2} difference) in the total-sky broadband flux between the model and CERES, a much larger positive bias of 7.05 W m^{-2} exists in the window region. This implies that a negative bias of about −8 W m^{-2} must exist in other spectral regions, which, as indicated by the above radiance comparison, is likely to be in the water vapor bands. This is further supported by the much larger discrepancy (−12.44 W m^{-2}) in the clear-sky OLR, in which case the positive bias in the window region is smaller due to the constraint of prescribed SST while the negative bias in the water vapor band is more prominent. Furthermore, similar cancellation of errors at different major frequency bands is also evident from the radiance derived band-by-band fluxes (X. L. Huang et al., Spectrally resolved fluxes derived from collocated AIRS and CERES measurements and their application in model evaluation: 1. Clear sky over the tropic oceans, submitted to *Journal of Geophysical Research*, 2007) simulated by this particular GFDL model.

3.2. Difference Spectrum

[19] By subtracting the nighttime radiances from the daytime radiances, an index of the diurnal difference spectrum is obtained. The difference spectra from observation and simulation are compared in various oceanic regions. It is found that in the equatorial oceans such as the warm pool and the west equatorial Indian where deep convection frequently occurs, the simulated total-sky spectra show a strong diurnal contrast which does not exist in the observation. This is illustrated by the example of the warm pool region (90E–150E, 10S–10N) in Figure 2. Here, the one-year mean AIRS daytime (1:30 p.m.) radiances are slightly smaller than the nighttime (1:30 a.m.) radiances, while the AM2 daytime radiances are much larger than the nighttime values (up to 3 K warmer in the window region). A similar day-night contrast also emerges in the broadband fluxes. In the same region, CERES has a day-night difference of −2.82 W m^{-2} while AM2 yields 8.17 W m^{-2}. It is evident from the AIRS observations that this region is cloudier and thus results in less upwelling radiation at 1:30 p.m. than at 1:30 a.m., which is consistent with the previous diagnosis [*Tian et al.*, 2004]. However, the model yields the opposite result.

[20] A survey of the NCAR CAM3 and the NCEP reanalysis data shows that the same diurnal contrast also exists in these models. As listed in Table 2, for the equatorial ocean region (10S–10N, 150E–150W), in contrast to a small negative day-night difference of OLR in the CERES observation, the numerical models all yield large positive differences. Note that the time steps closest to the two local times, 1:30 p.m. and 1:30 a.m., from the AM2 3-hourly outputs, the hourly CAM3 outputs and the 6-hourly NCEP reanalysis, respectively, are used to compute the day-night differences. Moreover, both AM2 and CAM3 have a very regular diurnal cycle from day to day, with the maximum of cloudiness occurring shortly after noon time and the minimum after midnight (not shown). Thus, if one looks at the
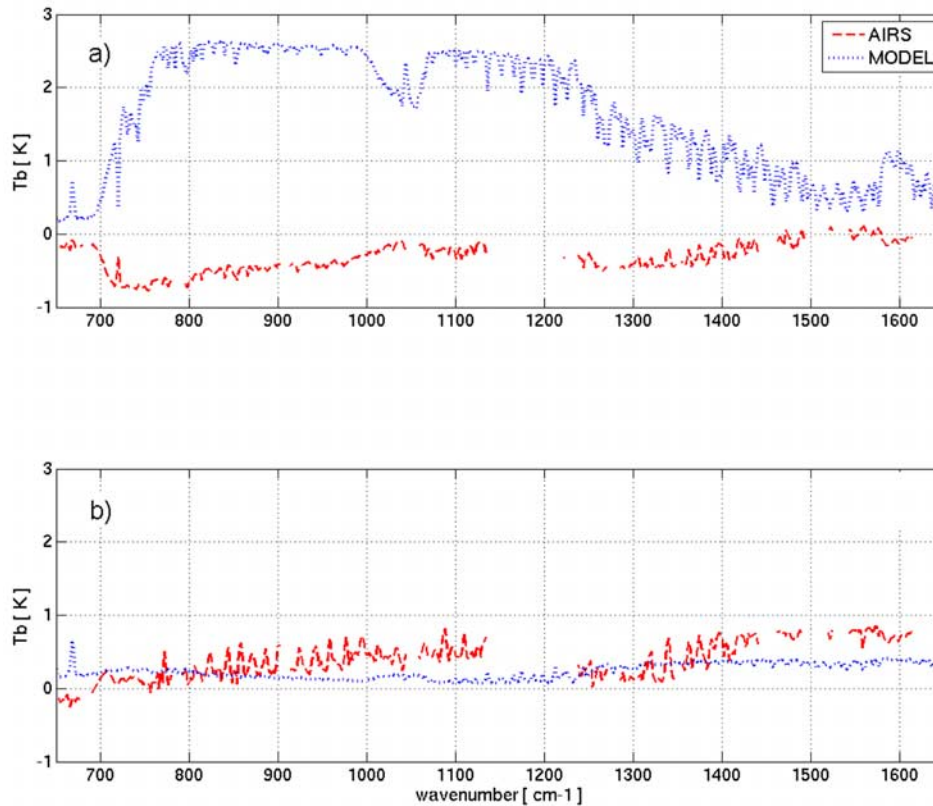
Figure 2. (a) Total-sky and (b) clear-sky diurnal difference spectra (day – night) over the tropical ocean (90E–150E, 10S–10N).

1:30 p.m.–1:30 a.m. difference, it would always be positive, with a mean of about 10 W m^{-2}. But in reality, as in the CERES observation (not shown), the difference at these two times may vary from a large negative value to a large positive value, with a mean close to zero. The irregular variation of the day-night difference of outgoing radiation is not properly represented in these numerical models.

[21] It is evident from the prominent T$_b$ contrast in the window band (Figure 2) that the large bias in the total-sky OLR, which does not occur in the clear-sky, is mostly due to the bias in clouds. In fact, by comparing with the high cloud and deep convection indices derived from geostationary satellite observations [*Tian et al.*, 2004], it is apparent that the model used in that study (an earlier version of AM2) generates an unrealistic midnight high cloud maximum in the tropics, which in the observed data occurs in the early morning. This model bias in the phasing of the diurnal cycle leaves a large signature in the diurnal difference spectrum. Despite measurements available only at two local times from AIRS and CERES, the gap is remarkable between model and observation.

4. Conclusion

[22] Satellite observed infrared radiances provide a more rigorous and insightful test for climate models than the broadband radiation flux alone. An agreement of the spectrum between simulation and observation would provide a very high level of confidence in the climate model. Such a

straightforward method avoids uncertainties associated with retrieval methods, and supplements the verification and improvement of climate models.

[23] It is shown that, with appropriate sampling and treatment of clouds, the GCM along with a narrow band radiation code can reproduce fairly well the observed clear- and total-sky spectra. However, systematic radiance biases do exist and reveal the biases in model simulated physical variables. For instance, the cold bias in the water vapor band, as emerging from the AM2-AIRS comparison, is a strong evidence of model bias in the upper troposphere, while the warm bias in the window band indicates an underestimation of cloud forcing. The existence of spectral errors of opposite signs suggests a cancellation in the broadband flux, which is confirmed by the independent CERES broadband longwave flux observations. Hence, although model simulated broadband flux can be tuned to match the observation over a certain period, and such tuning can yield satisfactory simulation (at least in the global-mean sense) for other periods, the resulting apparent agreement could be due to a cancellation of errors of opposite signs occurring in different spectral regions. This offers a plausible explanation to the cause of the AM2 clear-sky OLR bias that

Table 2. Four-Month Mean Day-Night OLR Difference (δOLR)

δOLR	CERES	AM2	CAM3	NCEP
Clear sky, W m^{-2}	2.22	0.55	−0.01	0.48
Total sky, W m^{-2}	−0.91	8.06	7.56	12.86

was explored by *Huang et al.* [2006]. Thus, the AIRS data affords a higher-grade evaluation of the model TOA radiation field and its dependence on temperature, water vapor and clouds than is enabled by the broadband observations.

[24] Although AIRS observations are limited to two particular local times, the index of diurnal difference using just these two times serves as a meaningful index for assessing the model-simulated diurnal cycle. The artificial contrast in the simulated OLR spectrum, as well as broadband OLR flux at these two times, indicates considerable errors in the diurnal phase of the modeled convection and cloudiness over equatorial oceans. During the iterative process to rectify model errors, the mean and the diurnal difference in the radiance observations provide a straightforward yet strict constraint for the model development process and contribute an important metric for model validation.

References

Berstein, L. S., et al. (1996), Very narrow band model calculations of atmospheric fluxes and cooling rate, *J. Atmos. Sci.*, *53*, 2887–2904.

Chahine, M., et al. (2006), AIRS: Improving weather forecasting and providing new data on greenhouse gases, *Bull. Am. Meteorol. Soc.*, *87*, 896–909.

Collins, W. D., et al. (2004), Description of the NCAR Community Atmosphere Model (CAM3), 226 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.

GFDL Global Atmospheric Model Development Team (GFDL GAMDT) (2004), The new GFDL global atmospheric and land model (AM2-LM2), evaluation with prescribed SST simulations, *J. Clim.*, *17*, 4641–4673.

Haskins, R. D., R. M. Goody, and L. Chen (1997), A statistical method for testing a general circulation model with spectrally resolved satellite data, *J. Geophys. Res.*, *102*(D14), 16,563–16,582.

Horowitz, L. W. (2006), Past, present, and future concentrations of tropospheric ozone and aerosols: Methodology, ozone evaluation, and sensitivity to aerosol wet removal, *J. Geophys. Res.*, *111*, D22211, doi:10.1029/2005JD006937.

Huang, X., J. Farrara, S. S. Leroy, Y. L. Yung, and R. M. Goody (2002), Cloud variability as revealed in outgoing infrared spectra: Comparing model to observation with spectral EOF analysis, *Geophys. Res. Lett.*, *29*(8), 1270, doi:10.1029/2001GL014176.

Huang, X., and Y. L. Yung (2005), Spatial and spectral variability of the outgoing thermal IR spectra from AIRS: A case study of July 2003, *J. Geophys. Res.*, *110*, D12102, doi:10.1029/2004JD005530.

Huang, X., V. Ramaswamy, and M. D. Schwarzkopf (2006), Quantification of the source of errors in AM2 simulated tropical clear-sky outgoing longwave radiation, *J. Geophys. Res.*, *111*, D14107, doi:10.1029/2005JD006576.

Huang, Y., V. Ramaswamy, and B. Soden (2007), An investigation of the sensitivity of the clear-sky outgoing longwave radiation to atmospheric temperature and water vapor, *J. Geophys. Res.*, *112*, D05104, doi:10.1029/2005JD006906.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–470.

Lin, X., D. Randall, and L. Fowler (2000), Diurnal variability of the hydrologic cycle and radiative fluxes: Comparisons between observations and a GCM, *J. Clim.*, *13*, 4159–4179.

Loeb, N., et al. (2005), Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Terra satellite. part I: Methodology, *J. Atmos. Oceanic Technol.*, *22*, 338–351.

McClatchey, R. A., et al. (1972), Optical properties of the atmosphere, *Environ. Res. Pap. 411*, Air Force Cambridge Res. Labs., Cambridge, Mass.

Randel, W. J., and F. Wu (2007), A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data, *J. Geophys. Res.*, *112*, D06313, doi:10.1029/2006JD007339.

Tian, B., B. J. Soden, and X. Wu (2004), Diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere: Satellites versus a general circulation model, *J. Geophys. Res.*, *109*, D10101, doi:10.1029/2003JD004117.

Yang, G., and J. Slingo (2001), The diurnal cycle in the tropics, *Mon. Weather Rev.*, *129*, 784–801.

————————

C. Bardeen, Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO 80309, USA.

Q. Fu, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640, USA.

X. Huang, Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, 2455 Hayward St., Ann Arbor, MI 48109-2143, USA.

Y. Huang, Program of Atmospheric and Oceanic Sciences, Princeton University, 300 Forrestal Rd., Princeton, NJ 08544, USA. (yi.huang@noaa.gov)

V. Ramaswamy, NOAA Geophysical Fluid Dynamics Laboratory, 201 Forrestal Rd., Princeton, NJ 08540, USA.