JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 111, D14107, doi:10.1029/2005JD006576, 2006

Quantification of the source of errors in AM2 simulated tropical clear-sky outgoing longwave radiation

Xianglei Huang,[1] V. Ramaswamy,[2] and M. Daniel Schwarzkopf[2]

Received 8 August 2005; revised 24 February 2006; accepted 24 April 2006; published 26 July 2006.

[1] The global and tropical means of clear-sky outgoing longwave radiation (hereinafter OLRc) simulated by the new GFDL atmospheric general circulation model, AM2, tend to be systematically lower than ERBE observations by about 4 W m^{-2}, even though the AM2 total-sky radiation budget is tuned to be consistent with these observations. Here we quantify the source of errors in AM2-simulated OLRc over the tropical oceans by comparing the synthetic outgoing IR spectra at the top of the atmosphere on the basis of AM2 simulations to observed IRIS spectra. After the sampling disparity between IRIS and AM2 is reduced, AM2 still shows considerable negative bias in the simulated monthly mean OLRc over the tropical oceans. Together with other evidence, this suggests that the influence of spatial sampling disparity, although present, does not account for the majority of the bias. Decomposition of OLRc shows that the negative bias comes mainly from the H_2O bands and can be explained by a too humid layer around 6–9 km in the model. Meanwhile, a positive bias exists in channels sensitive to near-surface humidity and temperature, which implies that the boundary layer in the model might be too dry. These facts suggest that the negative bias in the simulated OLRc can be attributed to model deficiencies, especially the large-scale water vapor transport. We also find that AM2-simulated OLRc has $\sim$1 W m^{-2} positive bias originating from the stratosphere; this positive bias should exist in simulated total-sky OLR as well.

Citation: 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.

1. Introduction

[2] The total-sky radiative fluxes observed by Earth Radiation Budget Experiment (ERBE) [*Barkstrom*, 1984] are widely used in the climate modeling community to validate general circulation models (GCMs). Some cloud-related parameters in GCMs that are not well constrained by observations and theories are adjusted within acceptable ranges to achieve a net balance of total-sky radiative fluxes at the top of the atmosphere and maximize the agreements of total-sky radiation budgets between GCM simulations and ERBE observations. This is known as ''tuning'' in the climate modeling community. Typical parameters used in such tuning include the threshold of conversion from cloud droplet to raindroplet, cloud erosion parameter, and the precipitation efficiency. Such tuning, however, does not guarantee a satisfactory agreement between simulated and observed clear-sky radiative fluxes at the same time because it might improve model performance for the wrong reason. For example, the long-term global annual mean of clear-sky outgoing longwave radiation (hereinafter OLRc) simulated by AM2, the new GFDL AGCM, is 4.8 W m^{-2} lower than ERBE observations [*GFDL Global Atmospheric Model Development Team*, 2004] even though the total-sky global radiation budget is in good balance and in good agreement with ERBE observations.

[3] Figure 1a shows the time series of the global monthly mean OLRc from ERBE observations and AM2 simulation from 1985–1989. It can be seen that the difference between the simulated and observed OLRc is systematically negative with a mean of -4.03 W m^{-2} and a standard deviation of 0.99 W m^{-2}. The tropical monthly mean OLRc simulated by AM2 (Figure 1b) has a similar negative bias (mean: -4.86 W m^{-2}, standard deviation: 0.88 W m^{-2}). Such a bias is not small and could influence the longwave cloud radiative forcing estimated from the model simulation. Therefore it is necessary to understand the causes of such negative bias in greater detail.

[4] Part of the negative bias can be attributed to the different ways used by the model and ERBE to obtain the monthly mean OLRc. In model simulations, OLRc is computed from temperature and humidity profiles at each grid box no matter whether the grid box is cloud-free or not. The monthly mean OLRc over a geographical region is obtained by averaging over all grid boxes inside this region and then averaging over the whole month. For ERBE, monthly mean OLRc is estimated from measurements over cloud-free pixels only. Neither spatial nor temporal distributions of these cloud-free pixels are guaranteed to be

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA.
[2]Geophysical Fluid Dynamics Laboratory, NOAA, Princeton University Forrestal Campus, Princeton, New Jersey, USA.

0148-0227/06/2005JD006576$09.00

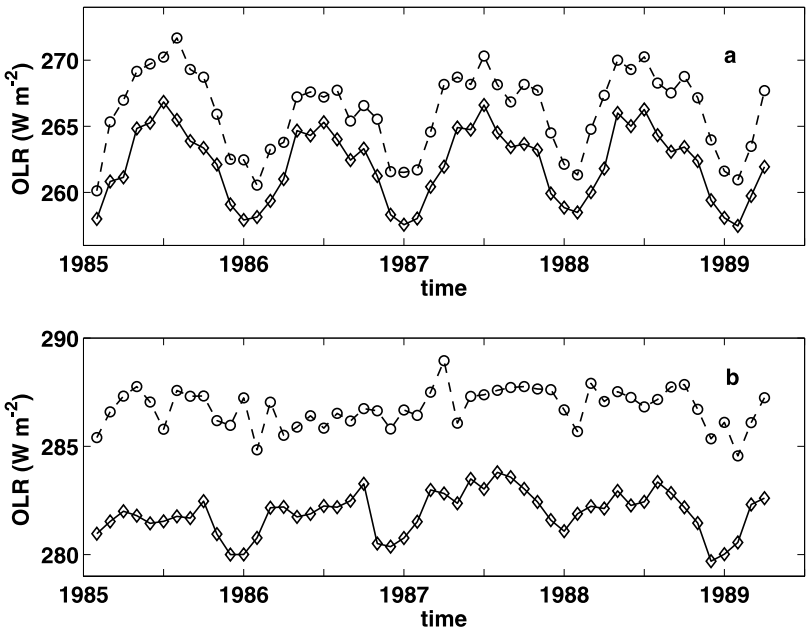

Figure 1. (a) Time series of the globally averaged monthly mean of clear-sky OLR from ERBE observations (the dashed line with circles) and AM2 simulations (the solid line with diamonds). (b) Same as Figure 1a except for the tropically averaged monthly mean. The spatial average for both AM2 and ERBE is done by averaging over the area with valid ERBE clear-sky measurements. No attempt is made to fill the missing values in the ERBE data set. Given that most missing values were at high latitudes, the globally averaged monthly means shown in Figure 1a are higher than the corresponding real global means.

uniform. Moreover, a typical tropical grid box in GCM is about 200 km by 200 km while the field of view (hereinafter FOV) of ERBE measurements is only about 40 km in diameter. Therefore, when an area comparable to a GCM grid box is partly covered by cloud, a cloud-free ERBE pixel inside this area could be drier than the average of the whole area. As a result, the OLRc derived from this cloud-free pixel is expected to be higher than the OLRc averaged over the whole area. In other words, monthly mean OLRc estimated from satellite could have a potential positive bias (dry bias) compared to the monthly mean OLRc computed by models, especially over the regions of active convection.

[5] This potential dry bias in satellite estimated OLRc has long been recognized [*Cess and Potter*, 1987; *Kiehl and Briegleb*, 1992; *Collins and Inamdar*, 1995; *Slingo et al.*, 1998; *Allan and Ringer*, 2003; *Allan et al.*, 2004]. For example, *Kiehl and Briegleb* [1992] calculated OLRc over the tropical oceans on the basis of temperature and humidity fields from ECMWF analyses and found the monthly mean ERBE OLRc is higher than calculated by 10–15 W m^{-2} in the tropical convective regions. *Allan and Ringer* [2003] compared CERES OLRc with OLRc on the basis of the ECMWF ERA-40 reanalysis and found 6–8 W m^{-2} differences between CERES and ERA-40 OLRc over regions of warm sea surface temperature (SST) and strong ascent. The potential dry bias in ERBE OLRc means that the difference between simulated OLRc and ERBE OLRc can be systematically negative even if the model is perfect, and can definitely contribute to the negative bias shown in Figure 1. However, the magnitude of this contribution to the total negative bias remains to be understood. Recently *Allan et al.* [2005] have conducted a near real-time comparison between Geostationary Earth Radiation Budget (GERB) data and output from a numerical weather prediction (NWP) model initialized with analyses from a 3-D variational assimilation. Their results showed agreements between modeled and observed OLRc over the oceans within the expected uncertainty.

[6] Systematic biases in the AM2 model and calibration uncertainties of ERBE instruments can also contribute to the bias in the AM2 versus ERBE comparison. Limited by the information contained in the broadband outgoing longwave radiation (OLR) measurement, it is difficult to further assess the relative importance of these factors from a model versus ERBE comparison alone. On the other hand, the outgoing infrared spectrum has much more information content than OLR and therefore can help us further understand this problem. There are two straightforward ways to make use of the rich information contained in the outgoing infrared spectra: (1) We can decompose the OLRc differences between model and observation into different absorption bands and study the differences band by band and (2) given that different infrared channels are sensitive to emissions from different vertical layers [*Goody and Yung*, 1989], we can also group channels according to the vertical layers that they are most sensitive to and study the differences group by group. By doing this, we can quantitatively understand the error budgets in each individual absorption bands and the contributions from groups of channels sensitive to different altitudes. Moreover, if the footprint of the instrument which measured the outgoing IR spectra is significantly different from that of ERBE, the dry bias due to the spatial sampling

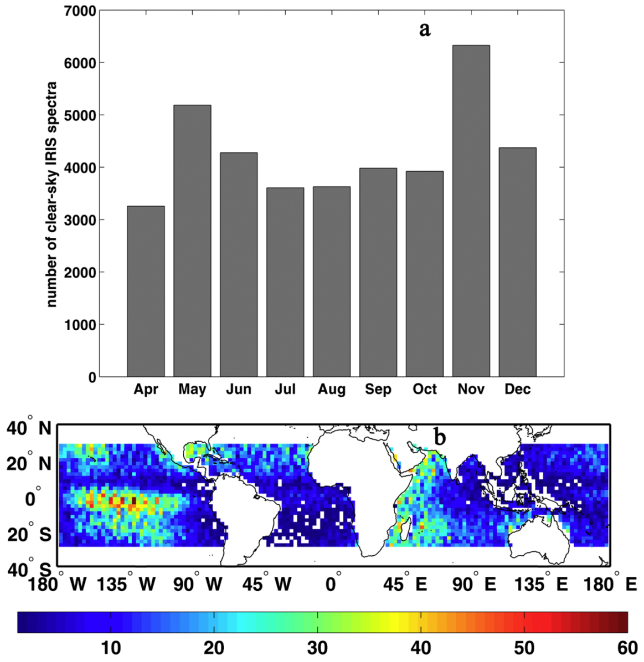

Figure 2. (a) Number of clear-sky IRIS spectra over the tropical oceans in each month. Clear-sky IRIS spectra are identified using the algorithm described in section 2.1. (b) Number of clear-sky IRIS spectra in each 2.5° longitude by 2° latitude grid box of the tropical oceans. A white grid box means that there is no IRIS spectrum inside that grid box identified as clear-sky spectrum.

disparity would be different from the case of ERBE measurements. This could help us evaluate the importance of the potential dry bias due to the spatial sampling disparity. Therefore, by using measurements of outgoing IR spectra, we can gain more insights about the causes of the discrepancies between simulated and observed OLRc as shown in Figure 1.

[7] In this study, we use the outgoing IR spectra measured by IRIS (Infrared Interferometer Spectrometer) in 1970 [*Hanel et al.*, 1972] to study the bias in AM2 simulated OLRc. For comparison, we also simulate OLRc on the basis of six-hourly meteorological fields from the ECMWF ERA-40 reanalysis [*Uppala et al.*, 2005]. In section 2, we describe the model and data used in this study as well as the data manipulation. Comparisons among AM2 simulated OLRc, OLRc inferred from IRIS observations, and OLRc based on the ECWMF reanalysis are presented in section 3. Section 3 also presents the AM2-IRIS and ECMWF-IRIS flux differences band by band. In section 4, IRIS channels are categorized into several groups according to the peaks of their weighting functions and the flux difference in each group is examined. Conclusions are given in section 5.

2. Model and Data Manipulation

2.1. IRIS Observations

[8] IRIS-D (hereinafter IRIS) was aboard Nimbus-7 and collected data from April 1970 to January 1971 [*Hanel et al.*, 1972], a moderate La Niña period. It was a Michelson FTIR spectrometer covering 400–1600 cm^{-1} with an apodized spectral resolution of 2.8 cm^{-1}. The signal-to-noise ratio was higher than 100 at the midpoint of the spectrum and gradually degraded to about 20 at the frequency endpoints. The estimated in-flight calibration uncertainty was about 0.5 K in brightness temperature [*Hanel et al.*, 1972; *Harries et al.*, 2001]. The FOV of IRIS was a circle with a diameter of 95 km, 6.25 times larger than the FOV of ERBE and about 15% of a typical AM2 grid box in the tropics. In total IRIS collected about 700,000 good-quality spectra during its 10-month operation. Its excellent instrument performance and rich information content made it a valuable data set that has been studied for more than three decades [for example, *Kunde et al.*, 1974; *Prabhakara et al.*, 1976, 1988; *Iacono and Clough*, 1996; *Haskins et al.*, 1997; *Harries et al.*, 2001; *Huang et al.*, 2002; *Anderson et al.*, 2004]. This study, to our knowledge, is the first study to use IRIS to help understand the biases in model-simulated broadband OLR. Given the complicated spectral dependence of land surface emissivity and the consistency between observed SSTs and SSTs prescribed in the model, in this study we focus on the tropical oceans (30°S–30°N) only. Because of a signal-to-noise consideration [*Haskins et al.*, 1997], we only use IRIS spectra over 400–1400 cm^{-1} (720 channels in total) from April to December 1970.

[9] Since we are interested in using IRIS observations to study OLRc, two issues have to be addressed: (1) identifying clear-sky spectra and (2) estimating OLRc from an individual IRIS spectrum which does not cover the whole range of longwave emission. Our approaches are described in the following two paragraphs.

[10] For issue 1, clear-sky spectra are identified on the basis of the following two criteria: (1) The 11 μm brightness temperature must be close to the climatological SST (within ±6 K) at the sampling location [*Haskins et al.*, 1997; *Harries et al.*, 2001] and (2) no fingerprints of ice or water clouds are apparent upon examination of the brightness temperature difference between 8 μm and 11 μm bands (ΔBT_{8-11}) as well as the brightness temperature difference between 11 μm and 12 μm bands (ΔBT_{11-12}) [*Ackerman et al.*, 1990]. As pointed out by *Ackerman et al.* [1990], there is a weak water vapor absorption line in the 8 μm band (8.3–8.4 μm). 11 μm (11.06–11.25 μm) and 12 μm (11.93–12.06 μm) bands are between H_2O absorption lines. The absorption coefficient of ice increases from 11 μm to 12 μm more rapidly than that of liquid water. Therefore, by this trispectral approach, clear-sky and ice and water clouds can be distinguished. After applying these two criteria, an average of 4415 spectra (30% of the total tropical spectra) sampled over the tropical oceans per month are identified as clear-sky spectra. Figure 2 shows the number of clear-sky spectra in each month identified by this algorithm and the spatial distribution of these clear-sky spectra. It can be seen that the spatial coverage of these clear-sky spectra is fairly good except for areas with frequent convection such as the maritime continent in the western Pacific.

[11] As for issue 2, estimating OLRc from an IRIS spectrum, we use the radiance (R_ν) at each IRIS channel to estimate the spectral flux (F_ν) at the same channel by assuming a linear relation between R_ν and F_ν

$$F_\nu = s_{1\nu}R_\nu + s_{2\nu}. \quad (1)$$

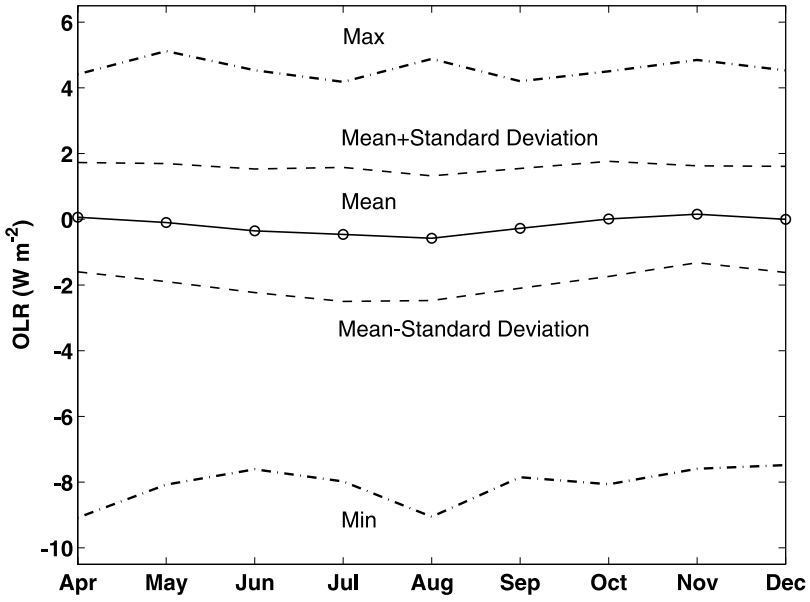

Figure 3. Difference between predicted OLR based on synthetic IRIS spectra and directly computed OLR from MODTRAN. About 40,000 random tropical profiles are used. Refer to section 2.1 for further details. The dash-dotted lines are the maximum and minimum differences for all random profiles in a given month. The solid line with open circles is the monthly mean difference averaged over these random profiles. The dashed lines show the monthly mean ± standard deviation of each month.

Two IR spectral regions, <400 cm^{-1} and >1400 cm^{-1}, are not covered by IRIS spectra used in this study. Given that both regions are primarily sensitive to mid and upper tropospheric water vapor, we estimate the fluxes over these regions ($F_{uncovered}$) by a linear combination of the fluxes in the H_2O $v2$ band (F_{v2}) covered by IRIS (1320–400 cm^{-1}) and in a narrow window region (F_{win}) transparent to atmospheric absorption (889–904 cm^{-1}):

$$F_{uncovered} = aF_{v2} + bF_{win} + c. \quad (2)$$

The coefficients, s_{1v}, s_{2v}, a, b, and c, are obtained in the following way: first ~40,000 tropical vertical profiles of temperature and humidity from the AM2 simulation described in the following subsection (section 2.2) are randomly chosen. These profiles are then fed into a radiative transfer model, MODTRAN, to compute synthetic IRIS spectra and OLRc. The synthetic IRIS spectra and OLRc are used to first estimate s_{1v} and s_{2v} by linear regression of equation (1) and then to estimate a, b, and c by linear regression of equation (2). To test how good this estimate of OLRc is, we randomly choose another set of ~40,000 tropical profiles and use MODTRAN to compute another set of synthetic IRIS spectra and OLRc (hereinafter the "computed" OLRc). Then we use each synthetic IRIS spectrum to estimate $F_{uncovered}$ as described above and, in this manner, we obtain a "predicted" OLRc based on each individual synthetic IRIS spectrum. This predicted OLRc is then compared to the "computed" OLRc directly obtained from MODTRAN. The comparison between the "predicted" OLRc and "computed" OLRc is shown in Figure 3. It can be seen that for all ~40,000 randomly chosen profiles, the maximum difference between predicted and computed OLRc is less than ±9 W m^{-2}. For each month, the standard deviation of the differences is about 1.8 W m^{-2}. For monthly averaged OLRc, the difference is about −0.17 W m^{-2}. These facts give us confidence to use this prediction scheme to estimate OLRc from IRIS spectra, and then compare the estimated monthly mean OLRc with the counterpart simulated by AM2.

2.2. Model

[12] In this study, we use AM2, an atmospheric GCM (AGCM) recently developed at Geophysical Fluid Dynamics Lab (GFDL). In brief, AM2 employs a hydrostatic, finite difference dynamical core using the staggered Arakawa B-grid with 2.5° longitude by 2° latitude resolution. The standard configuration of AM2 has 24 vertical levels with the lowest model level about 30 meters above the surface and five levels in the stratosphere, the top level being at ~3 hPa. Cloud liquid water, cloud ice amount, and cloud fraction are treated as prognostic variables. The relaxed Arakawa-Schubert scheme is used for cumulus parameterization with several modifications. The longwave and shortwave radiation parameterizations follow *Schwarzkopf and Ramaswamy* [1999] and *Freidenreich and Ramaswamy* [1999], respectively. The detailed description of AM2 can be found in *GFDL GAMDT* [2004].

[13] Four-member AM2 ensemble runs are forced by observed monthly SSTs from 1966 to 1971 [*Rayner et al.*, 2003]. Results reported here are averages over the four members. Observed CO_2 and other greenhouse gas concentrations appropriate for the IRIS period are used in the runs. Three-hourly instantaneous outputs over the IRIS period (April–December 1970) are archived. To minimize the temporal sampling disparity between IRIS and the model, these 3-hourly instantaneous outputs are further sampled to the same time and location as those IRIS clear-sky spectra

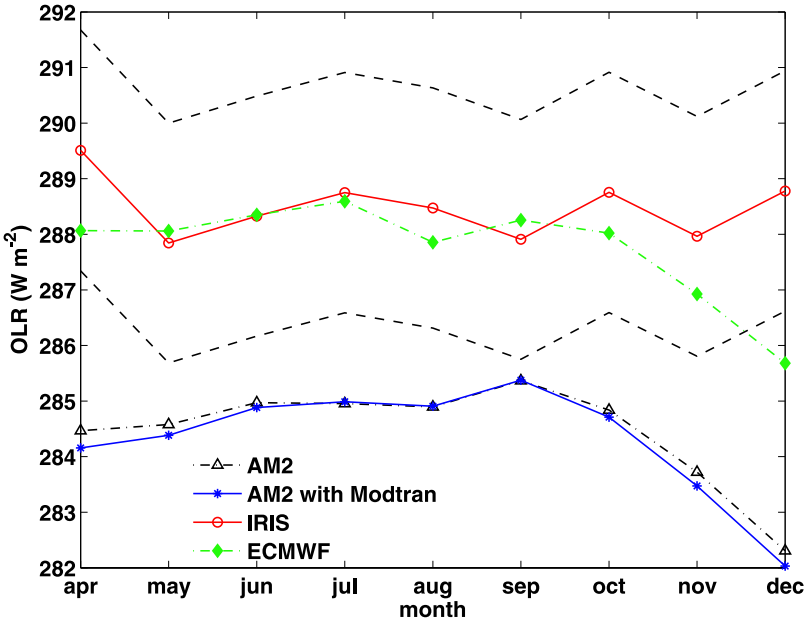

Figure 4. Monthly mean of clear-sky OLR inferred from IRIS clear-sky spectra over the tropical oceans (the red solid line with open circles), computed from ECMWF ERA-40 reanalysis (the green dash-dotted line with solid diamonds), and obtained from MODTRAN calculation based on the corresponding AM2 simulation (the blue solid line with stars). The clear-sky OLR computed by the AM2 radiation scheme is shown as the dash-dotted line with open triangles. The two dashed lines show the range of uncertainty in the IRIS clear-sky OLR due to IRIS calibration uncertainties.

identified in section 2.1. Then the subsampled outputs are fed into a very narrow band radiative transfer model, MODTRAN v4.1 [*Bernstein et al.*, 1996], to compute OLRc and IRIS-like clear-sky spectra. MODTRAN (Moderate transmission code) was developed by the Air Force Research Laboratory with a high computational efficiency and satisfactory performance compared to line-by-line radiative transfer code [*Bernstein et al.*, 1996].

[14] For comparison, 6-hourly outputs over the same period from the ECMWF ERA-40 reanalysis [*Uppala et al.*, 2005] are sampled and fed into MODTRAN in the same way to generate OLRc and synthetic clear-sky spectra. No IRIS observations were assimilated by ECMWF ERA-40. Therefore IRIS and ECMWF ERA-40 are two independent data sets. The spatial resolution of the ECMWF ERA-40 reanalysis data is 2.5° by 2.5°, comparable to the spatial resolution of AM2.

3. Differences in the Clear-Sky OLR and Different Absorption Bands

[15] In this section, we first present the comparisons of OLRc among the AM2, IRIS and ECMWF. Then the flux differences in each individual absorption bands are discussed.

3.1. Differences in the Clear-Sky OLR (OLRc)

[16] Figure 4 shows the monthly averaged OLRc over the tropical oceans inferred from IRIS clear-sky spectra and computed from the corresponding AM2 simulation and ECMWF reanalysis data, respectively. As mentioned in section 2, AM2 output (ECWMF reanalysis product) are sampled to the same location and the same time as IRIS clear-sky spectra used in this study and then the AM2 (ECMWF) monthly mean OLRc is derived only from those subsampled data sets. It can be seen that, for all 9 months, AM2 has a consistent negative bias in OLRc compared to IRIS observation (-4.02 ± 1.14 W m^{-2}). This bias is larger than the instrumental calibration uncertainty of IRIS. Moreover, this bias is comparable to the negative bias shown in the AM2 versus ERBE comparison (Figure 1b, -4.86 ± 0.88 W m^{-2}). On the other hand, ECMWF has a generally good agreement with the IRIS observations: the mean difference over 9 months is -0.72 W m^{-2} with a standard deviation of 1.07 W m^{-2}. The largest difference between ECMWF and IRIS occurs in December. Figure 4 also shows that the OLRc computed by MODTRAN has a good

Table 1. List of Eight Spectral Bands Used to Decompose Clear-Sky OLR[a]

	Spectral Range, cm^{-1}	Major Absorber	Flux, W m^{-2}
1	<400 or >1400	H_2O[b]	60.2
2	400–560	H_2O	52.2
3	560–800	CO_2, N_2O[c]	58.0
4	800–985	H_2O continuum	59.7
5	985–1080	O_3	18.0
6	1080–1200	H_2O continuum	23.5
7	1200–1320	H_2O, N_2O, CH_4[c]	12.4
8	1320–1400	H_2O	4.5

[a]The flux in each band observed by or inferred from IRIS measurements is also listed (averages of 9-month IRIS clear-sky spectra over the tropical oceans).

[b]Not covered by IRIS observations, flux is inferred as described in section 2.

[c]N_2O has two bands centered at 589 cm^{-1} and 1285 cm^{-1} and CH_4 has a strong band centered at 1306 cm^{-1}.

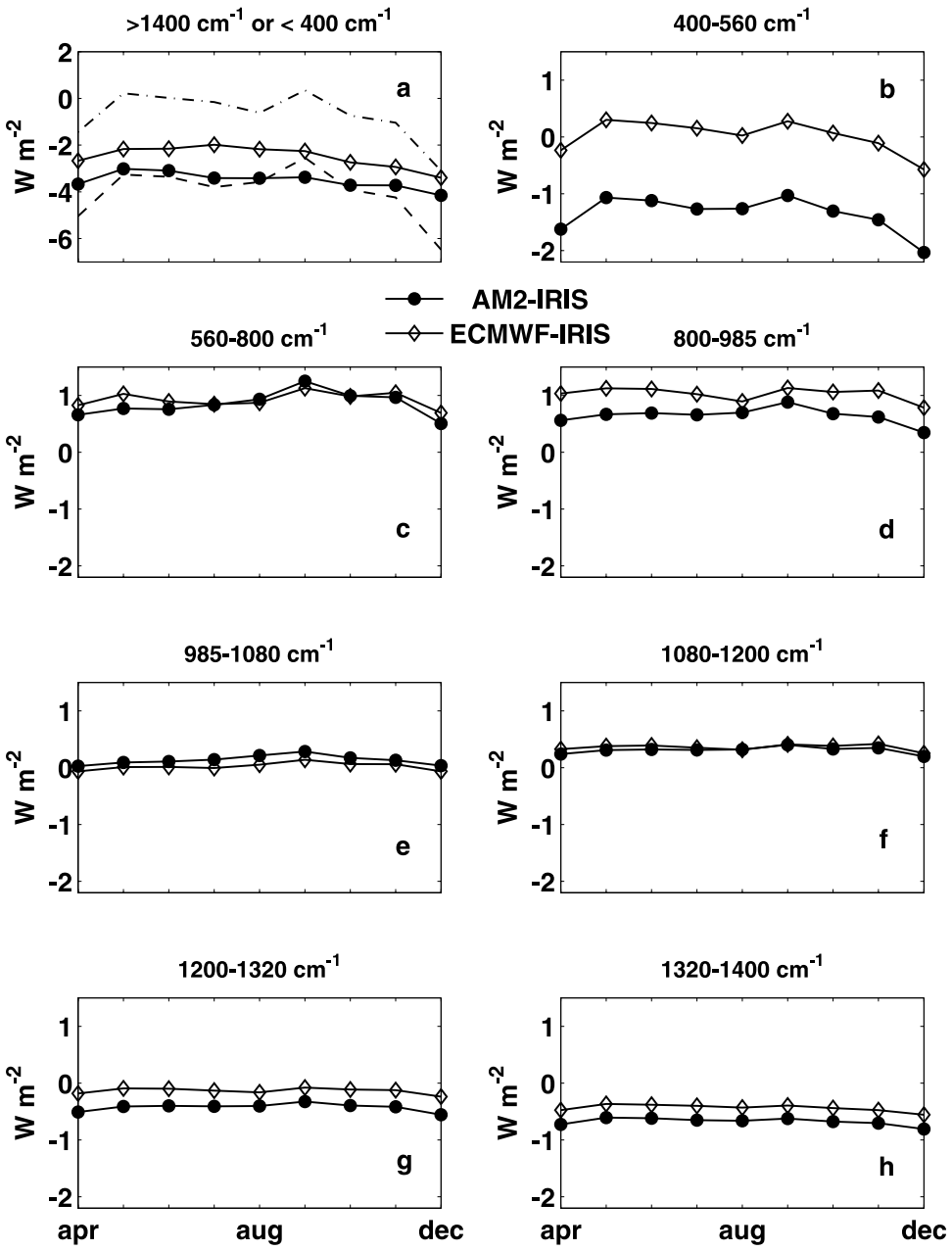

Figure 5. (a) Solid line with solid circles is the AM2-IRIS difference of monthly mean clear-sky flux in band 1 as listed in Table 1. Solid line with open diamonds is the corresponding ECMWF-IRIS difference in band 1. The spectral range of band 1 is labeled. The clear-sky OLR difference between AM2 and IRIS (dashed line) and between ECMWF and IRIS (dash-dotted line) are also shown. (b–h) Same as Figure 5a except for bands 2–8 as listed in Table 1. The spectral range of each band is labeled.

agreement with the OLRc calculated from the AM2 radiation scheme, which confirms the robustness of the clear-sky longwave parameterizations used in AM2.

[17] As described in section 2, the IRIS FOV is about six times larger than that of ERBE. As a result, the bias caused by the spatial sampling disparity between GCM and satellite observations should be different for AM2-IRIS and AM2-ERBE. If the spatial sampling disparity were a major contributor to the negative bias in the simulated OLRc, the amplitude of the bias in the AM2-IRIS comparison would be notably different for that in the AM2-ERBE comparison. Moreover, such bias would be prominent in the ECMWF-IRIS comparison as well. Yet the comparable biases from the AM2-IRIS and AM2-ERBE comparisons (Figures 4 and 1b) and the good agreement between ECMWF and IRIS (Figure 4) imply that this is not the case. These results indicate that, although the spatial sampling disparity between the model and satellite can definitely contribute to the negative bias in the AM2-simulated OLRc, it cannot account for the majority of the negative bias seen in the simulated OLRc.

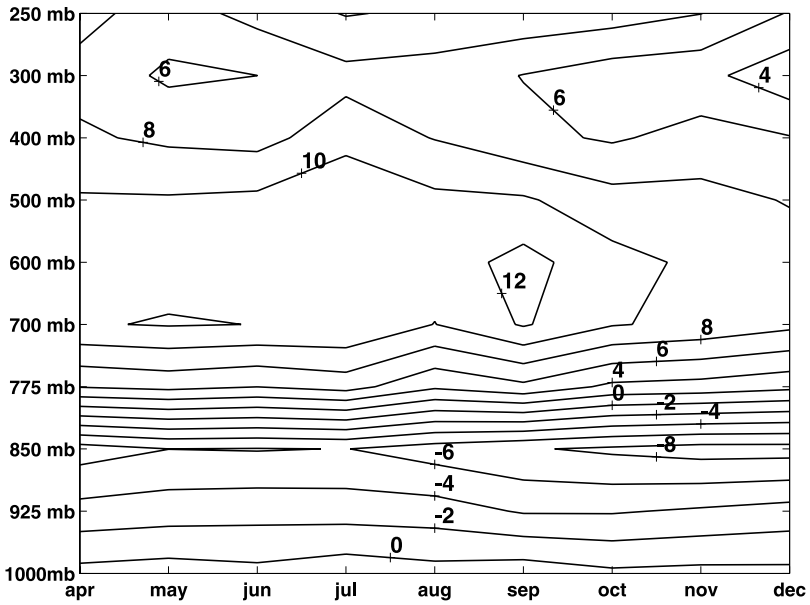

Figure 6. AM2-ECMWF difference of the tropical monthly mean of relative humidity from April to December 1970. The unit of relative humidity is percentage. Contour interval is 2%.

3.2. Flux Differences at Different Absorption Bands

[18] In this subsection, we decompose the OLRc differences into eight bands as listed in Table 1. The monthly mean of the AM2-IRIS difference in each band is shown in Figure 5. The corresponding difference between ECMWF and IRIS is also shown in Figure 5. Some features to be noted in Figure 5 are the following:

[19] 1. For the CO_2 band and two window regions (Figures 5c, 5d, and 5f), the AM2-IRIS flux differences are positive, as are the ECMWF-IRIS flux differences in these bands. For these three bands, the difference between AM2 and ECMWF flux is small ($\sim$0.2 W m^{-2}).

[20] 2. For all H_2O bands (Figures 5a, 5b, 5g, and 5h), the AM2-IRIS flux differences are negative and they are about twice the magnitude of the ECMWF-IRIS flux differences. For the water vapor $\nu 2$ band (1200–2000 cm^{-1}) which does not contribute much to OLRc (Figures 5g and 5h and a small portion in Figure 5a), the AM2 flux is about 0.52 W m^{-2} smaller than the ECWMF flux and 1.2 W m^{-2} smaller than the IRIS flux. However, for the water vapor rotation band (<560 cm^{-1}) (Figure 5b and the majority in Figure 5a), a band contributing $\sim$37% to OLRc flux, the AM2 flux is about 2.51 W m^{-2} smaller than the ECMWF flux and 4.99 W m^{-2} smaller than the IRIS flux.

[21] For ECMWF-IRIS, the positive differences in feature 1 are, to a very large extent, offset by the negative differences in feature 2. As a result, the OLRc difference between the ECMWF reanalysis and IRIS is small. For AM2-IRIS, the positive differences in feature 1 can only offset part of the negative differences in feature 2 and, therefore, the OLRc differences between AM2 and IRIS are still notably negative. The water vapor rotation and $\nu 2$ band are both primarily sensitive to the mid and upper tropospheric humidity. Therefore the discrepancies between AM2-IRIS and ECMWF-IRIS suggest that, for all nine months examined here, the tropical middle and upper troposphere in AM2 is more humid than both the real atmosphere and the ECMWF reanalysis. This point is further supported by looking at the difference in tropical mean relative humidity between AM2 and ECMWF for the IRIS period as shown in Figure 6. It can be seen that, above 780 mb, AM2 is more humid than ECMWF all through the troposphere with a maximum about 12% around 650 mb.

[22] Given that the observed SST is used in the simulation and that we only compare the spectral sampling over the tropical oceans, the surface temperature difference between model and observations is too small to explain the positive biases in the two window regions (Figures 5d and 5f). Possible candidates for such positive biases are (1) the algorithm that we use to identify clear-sky spectra and (2) a drier lower part of the troposphere (<3 km) in the model than in the observations. For candidate 1, when a very small cloud fraction was present inside a FOV, an IRIS spectrum sampled at this FOV might pass both the brightness tem-

Table 2a. Properties of Groups of IRIS T-Channels (CO_2 Band and Two Window Regions) Used in This Study[a]

	T1	*T2*	*T3*	*T4*	*T5*	*T6*	*T7*	*T8*
Range of z_{max} in km	<1	[1, 3]	[3, 5]	[5, 7]	[7, 9]	[9, 11]	[11, 13]	>13
Number of channels	216	73	13	19	7	12	6	46
Nadir flux, W m^{-2}	86.7	31.1	5.8	7.2	2.2	3.1	1.2	8.8

[a]The nadir flux, which is the nadir-view radiance multiplied by π, is derived from 9-month IRIS clear-sky observations over the tropical oceans. z_{max} is the altitude at which the nadir-view weighting function of a IRIS channel has its maximum.

Table 2b. Properties of Groups of IRIS RH-Channels (Water Vapor Bands Covered by IRIS Measurement) Used in This Study[a]

	RH1	*RH2*	*RH3*	*RH4*	*RH5*	*RH6*
Range of z_{max} in km	<1	[1, 3]	[3, 5]	[5, 7]	[7, 9]	>9
Number of channels	26	56	55	85	29	7
Nadir flux, W m^{-2}	5.6	14.2	18.8	25.9	5.9	1.8

[a]The nadir flux, which is the nadir-view radiance multiplied by π, is derived from 9-month IRIS clear-sky observations over the tropical oceans. z_{max} is the altitude at which the nadir-view weighting function of a IRIS channel has its maximum.

perature threshold test and the water/ice spectral fingerprint test. Because of the attenuation caused by clouds, such an observed spectrum would have less radiance in the CO_2 band and the two window regions than a corresponding clear-sky spectrum. So even if the simulated spectrum is perfect, this would still lead to positive biases in differences between simulated and observed fluxes in these three bands.

[23] For candidate 2, the major clear-sky absorption in the window regions is due to H_2O continuum, which is only important in the boundary layer and the bottom part of the free troposphere (<3 km). H_2O continuum absorption is also not negligible in the CO_2 band. Therefore a dry bias in the model's boundary layer and the bottom part of the free troposphere could also produce positive biases in these bands. Indeed, the relative humidity difference between AM2 and ECMWF ERA-40 (Figure 6) does show that, from 1000 mb to 780 mb, AM2 is drier than ECMWF ERA-40. The two factors discussed here are not exclusive to each other; it is impossible to unambiguously distinguish them from the flux differences presented here. In the next section, we will discuss further the positive biases in the window regions.

4. Flux Differences in Different Groups of IR Channels

[24] We classify IRIS infrared channels into two categories: channels sensitive to temperature (hereinafter T-channels) and channels sensitive to relative humidity (hereinafter RH-channels). Spectral regions not covered by IRIS spectra (band 1 in Table 1) are not included. To facilitate the comparison and interpretation, the 9.6 μm ozone band is

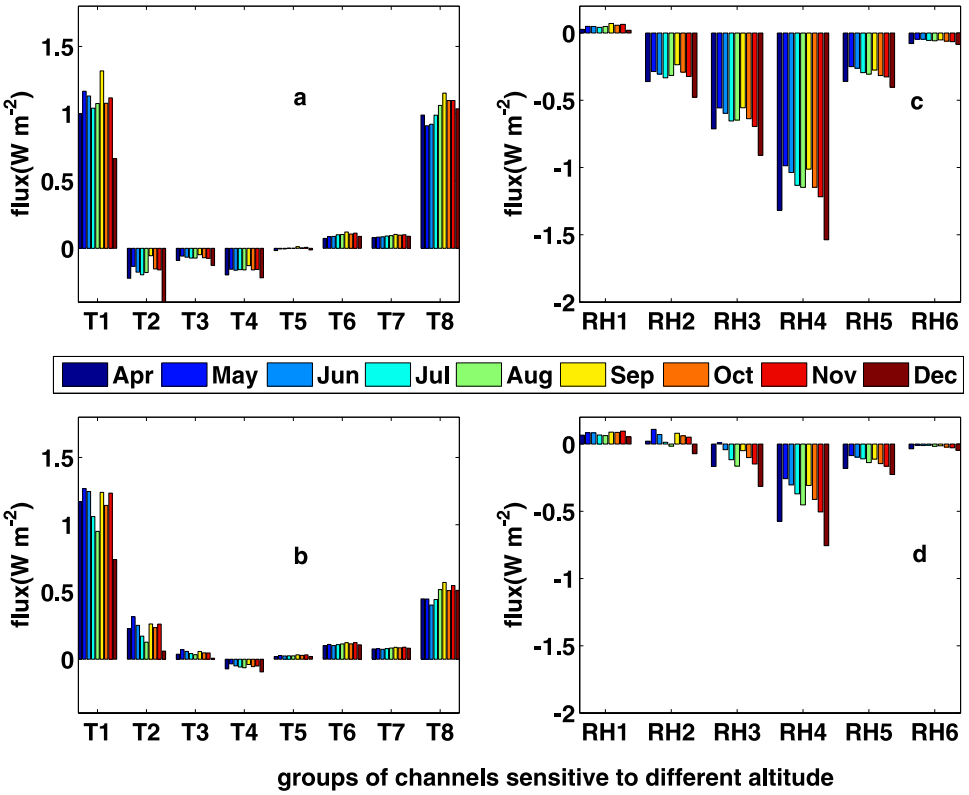

Figure 7. (a) AM2-IRIS nadir flux differences in eight groups of T-channels (channels sensitive to temperature) as defined in Table 2a. Refer to the context for the definition of nadir flux. Each bar represents the flux difference of one month. The T1–T8 groups are defined in Table 2a. (b) Same as Figure 7a except for ECWMF-IRIS. (c) Same as Figure 7a except for six groups of RH-channels (channels sensitive to relative humidity) as defined in Table 2b. (d) Same as Figure 7c except for ECMWF-IRIS.

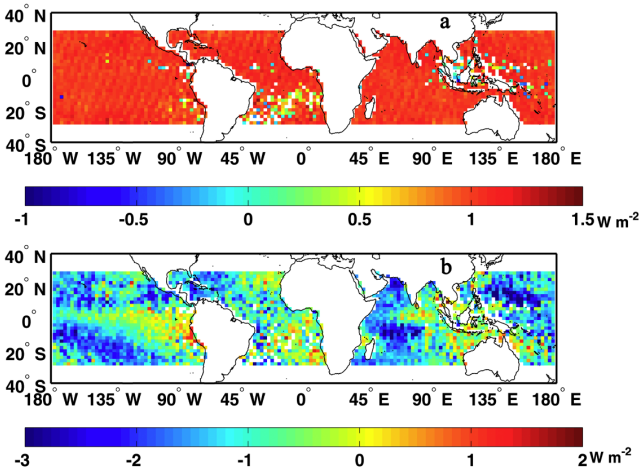

Figure 8. (a) Nine-month composite of AM2-IRIS nadir flux difference in the T8 group (sensitive to the temperature above 13 km). Each grid box is 2.5° longitude by 2° latitude. The unit is W m^{-2}. (b) Same as Figure 8a except for the RH4 group (sensitive to the relative humidity at 5–7 km). White pixels over the oceans are grid boxes with no single clear-sky spectrum identified during the IRIS period.

not included here because of the complexity and uncertainty of ozone profiles. Several channels dominated by CH_4 and N_2O absorption are also excluded. T-channels include channels in the CO_2 band and two window regions (bands 3, 4 and 6 in Table 1) while RH-channels consist of those in the H_2O bands (bands 2, 7 and 8 in Table 1). This classification is based on the most important contributor to each channel. Inevitably, it is a crude classification in several aspects. For example, T-channels in the window regions are actually also sensitive to humidity in the boundary layer and the bottom part of the free troposphere because of H_2O continuum absorption.

[25] For each channel in these two categories, the nadir-view weighting function, $w(z)$ (also known as the kernel function), is calculated using a typical tropical sounding profile [*Anderson et al.*, 1986] and z_{max}—the altitude where $w(z)$ reaches its maximum—is identified. Then channels with similar z_{max} are grouped together as listed in Table 2: T-channels are divided into 8 groups and RH-channels are divided into 6 groups. The weighting function changes with respect to the zenith angle and so does z_{max}. In order to avoid errors introduced in the mapping from nadir-view radiance to angle-integrated flux, here we only examine the "nadir flux," which is simply the nadir-view radiance multiplied by π.

4.1. Temperature Channels

[26] Figure 7a shows the differences between AM2 and IRIS in the T-channel groups (T1–T8, as defined in Table 2a). In terms of the absolute value, the two largest contributors to the AM2-IRIS OLRc difference are the T1 (z_{max} < 1 km) and T8 (z_{max} > 13 km) groups, each contributing about +1 W m^{-2}. However, as shown in Table 2a, the total nadir flux of the T1 group is about 86.7 W m^{-2} and that of the T8 group is only 8.8 W m^{-2}. Therefore the relative difference in the T8 group (12.5%) is much larger than that in the T1 group (1.1%). Channels in the T8 group are primarily sensitive to stratospheric temperature and less affected by tropospheric variations. A 9-month composite map of the differences in the T8 group is shown in Figure 8a. It can be seen that the differences are fairly uniform over the whole tropical oceans, which further suggests that tropospheric contributions to the differences in this group are unimportant. Because the channels in the T8 group are predominantly sensitive to the stratosphere and the variation in the stratosphere is not sensitive to cloud variations in the troposphere, a similar positive bias in the T8 group should also exist in the simulated total-sky OLR. This means that, when cloud-related parameters are tuned to match simulated total-sky OLR with the ERBE observations, these parameters might have been "overtuned" to compensate for the bias originating from the stratosphere.

[27] As discussed in section 2, sea surface temperature is prescribed with observed values, so the positive AM2-IRIS difference in the T1 group reflects either a drier boundary layer in the model or a failure of our algorithm to exclude some cloudy spectra. The AM2-IRIS differences in the T2–T4 groups are negative. These two facts suggest that, if cloud contamination is responsible for the AM2-IRIS differences seen in the window regions (Figures 5d and 5f), the major contamination should be boundary layer clouds rather than high clouds. Otherwise, the differences in the T2–T4 group should be positive as well.

[28] The ECMWF-IRIS difference in each group is shown in Figure 7b. The difference in the T1 group is comparable to that of AM2-IRIS. The difference in the T8 group is about half of the AM2-IRIS counterpart. For both AM2-IRIS and ECMWF-IRIS differences, monthly variations in all groups are small.

4.2. Relative Humidity Channels

[29] Figure 7c shows the differences between AM2 and IRIS in RH-channel groups (RH1-RH6, as defined in Table 2b). The AM2-IRIS difference is only positive in the RH1 group, a group sensitive to boundary layer humidity. This is consistent with the explanation given in section 4.1 about the flux differences in the T1 groups. The positive differences of AM2-IRIS in the RH1 and T1 groups and the window regions (Figures 5d and 5f) are also consistent with the discussion of the spatial sampling disparity between satellite and AM2 in section 3.1. In a mostly cloudy region with an area comparable to that of a GCM grid box, the boundary layer humidity of cloud-free pixels is expected to be less than that averaged over the whole area. If the contribution of sampling disparity overshadowed other contributions, AM2-IRIS difference would be negative in the RH1 and T1 groups and the window regions. Therefore the positive differences at these channels are consistent with the argument that the spatial sampling disparity is not the dominant contributor to the bias in OLRc comparison.

[30] The largest absolute difference shown in Figure 7c is −1.2 W m^{-2} from the RH4 group, a group which also has the largest flux among all RH-channel groups. The negative differences in the humidity channels sensitive to the free troposphere (the RH2–RH6 groups) suggest that the simulated troposphere is wetter than observed, especially the middle and upper troposphere. The difference between

Table 3. Normalized Dimensionless Monthly AM2-IRIS Nadir Flux Differences in the Different RH Groups[a]

	RH1	RH2	RH3	RH4	RH5	RH6
Mean	−0.043	0.28	0.57	1.00	0.27	0.05
Std	0.017	0.024	0.016	0.0	0.008	0.004
Pert	0.031	0.26	0.54	1.00	0.26	0.05
Real diff, W m^{-2}	0.0476	−0.3253	−0.6627	−1.1703	−0.3105	−0.0602

[a]The normalization is done with respect to the monthly nadir flux difference in the RH4 groups. The mean and standard deviation for each RH group are listed as Mean and Std, respectively. Pert shows the normalized flux differences caused by a uniform perturbation to relative humidity from 6.5 to 8.5 km in a typical tropical profile [*Anderson et al.*, 1986]. The real 9-month mean flux differences before normalization are shown as "Real diff" in W m^{-2}.

ECMWF and IRIS in the RH1 group (Figure 7d) is comparable to that between AM2 and IRIS. For all the other groups, the amplitudes of ECMWF-IRIS differences are systemically smaller than those of AM2-IRIS differences by about a factor of two. The differences between ECMWF and IRIS in the RH3–RH6 groups are negative, implying that ECMWF is more humid than the observations in the middle and upper troposphere. Given that AM2 is more humid than ECMWF in these regions by about 6–12% (as shown in Figure 6), AM2 should be even more humid than the real atmosphere in these regions.

[31] If we normalize the monthly nadir flux differences in the RH2–RH6 groups with respect to the difference in the RH4 group as shown in Table 3, the ratio of the differences between any pair of RH2–RH6 groups changes little from month to month. In addition, if we perturb a typical tropical sounding profile by changing its relative humidity within a vertical layer from 6.5 km to 8.5 km, the corresponding normalized flux changes (Pert in Table 3) due to this perturbation are similar to the normalized monthly flux differences of AM2-IRIS. This suggests that a more humid layer around 6.5–8.5 km (about 500–350 mb) in the model could explain most of the AM2-IRIS flux differences in the RH2–RH6 groups.

[32] The 9-month spatial composite of the AM2-IRIS nadir flux differences in the RH4 group (sensitive to relative humidity at 5–7 km) is shown in Figure 8b. Although the monthly mean difference averaged over the tropical oceans is always negative, the difference is actually positive over many regions with subsidence such as the coasts off Peru and Namibia. Regions with large negative differences, on the other hand, are overlapped with the ITCZ (Intertropical convergence zone) and the SPCZ (Southern Pacific convergence zone) in the model. The correlation between the difference in the RH4 group and large-scale vertical movement can also be found by examining the difference with respect to the corresponding 500 mb vertical velocity (500 mb ω) simulated by AM2. Figure 9 shows the mean

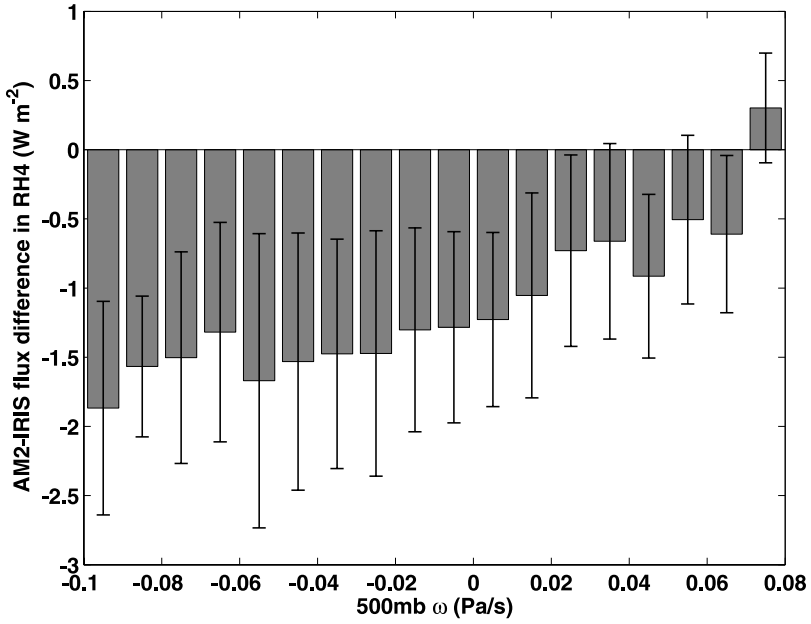

Figure 9. AM2 –IRIS nadir flux difference in the RH4 group (sensitive to the relative humidity at 5–7 km) with respect to the corresponding 500 mb vertical velocity simulated by AM2. The vertical velocity is grouped into 18 bins from −0.095 Pa/s to 0.075 Pa/s with a width of 0.01 Pa/s. For each bin, the shaded bar represents the average of flux differences for all samplings falling into that bin, and the mean ± standard deviation is shown as the ticked solid line.

flux differences in the RH4 group and the corresponding standard deviations for 18 bins of 500 mb ω (from -0.095 Pa/s to 0.075 Pa/s). It can be seen that, for strong ascending motion (large negative ω), the corresponding AM2-IRIS flux differences in the RH4 group tend to be negative and have large magnitudes. Only for strong descending motion (large positive ω) do the differences tend to be positive, but the magnitudes are smaller than those associated with strong ascending motion. Regions with strong descending motion tend to have high occurrence of marine stratus and, as discussed in section 3 and 4.1, some IRIS spectra contaminated by low clouds might pass our clear-sky screening procedure. However, our criteria to identify clear-sky spectra guarantee that, for any significant low-cloud contamination, the cloud tops should be quite low (around or less than 1 km). Therefore the impact of this kind of low-cloud contamination to channels in the RH4 group should be limited since these channels are sensitive to relative humidity around 5–7 km. Meanwhile, in some regions with high occurrence of low clouds, such as the Arabian Sea and the Indian Ocean west of Australia, ω_{500} tends to be largely positive but the AM2-IRIS flux difference is actually negative, as shown in Figure 8b. On the basis of these facts, we believe the impact of low-cloud contamination to Figures 8b and 9 should be limited.

[33] These results suggest that, compared to the observation, the middle and upper troposphere in the model tends to be more humid in the tropical convective regions but drier in the subsidence regions. This could be further related to deficiencies in simulating moisture transport in the tropical meridional and zonal circulations.

5. Conclusion

[34] In this study we use IRIS observations to quantify the bias in AM2-simulated OLRc over the tropical oceans. The spectrally resolved radiances recorded by IRIS and the wide FOV of IRIS make IRIS a valuable data set to study this problem. After the temporal sampling disparity between model and observation is minimized, there is still considerable negative bias in the simulated monthly mean OLRc over the tropical oceans. The amplitude of such negative bias (Figure 4) is out of the range of IRIS calibration uncertainties and comparable to the bias seen from the AM2 versus ERBE comparison (Figure 1b). This suggests that the spatial sampling disparity due to different fields of view between model and satellite, which does contribute to the clear-sky OLR difference between model and satellite, does not account for the majority of the negative bias in AM2 simulated tropical OLRc. The good agreement between IRIS and ECMWF and the positive AM2-IRIS differences in channels sensitive to boundary layer moisture are consistent with this argument. The quantitative understanding of this spatial sampling disparity and how this bias varies with instrument footprint, might be important for evaluating simulated clear-sky radiation budget and cloud forcing against counterparts from satellite observations. Such an investigation would require long enough model output with spatial resolution comparable to or smaller than the satellite footprint (e.g., mesoscale model output for at least 1 month), which is beyond the scope of this study.

[35] Spectral decomposition of OLRc into different absorption bands shows that the negative bias comes mainly from water vapor rotation bands while the biases from window regions and from the CO_2 band are actually positive. When IRIS channels are grouped according to the altitudes that their weighting functions peak at, and OLRc is decomposed into these groups, the major findings are the following: (1) The AM2 mean profiles over the tropical oceans are more humid than observations in the middle and upper troposphere (at least > 8% in relative humidity over 300–500 mb) and this humid layer could explain most differences in the RH-channels sensitive to the free troposphere; (2) to a large extent, the spatial distribution of the AM2-IRIS difference in midtropospheric humidity channels is correlated with the large-scale circulation—large negative differences in ascending regions and positive differences in strong descending regions; (3) the differences in the T1 and RH1 groups suggest a possible drier planetary boundary layer in the model; and (4) there is a $\sim$1 W m^{-2} bias in the simulated OLRc which originates from the stratosphere. Findings 1, 2, and 3 together imply that the majority of the negative bias in OLRc can be attributed to model deficiencies, especially the vertical and large-scale transport of moisture which further relates to the strength of the circulation. For finding 4, it is conceivable that because of the insensitivity of the stratospheric channels to the presence of tropospheric clouds, a similar positive bias in the stratospheric channels (the T8 group) should exist in the total-sky OLR comparison as well. This implies a potential ''overtuning'' when only parameters related to the tropospheric physics are adjusted in the tuning procedure. Given that the channels in the T8 group are all in the CO_2 fundamental band centered at 667 cm^{-1} and that most modern GCMs parameterize flux in this band, perhaps future tuning procedures could omit this band and only use the total fluxes from other bands. This would eliminate the potential overtuning problem.

[36] IRIS recorded outgoing infrared spectra with a high spectral resolution in a period much earlier than the current era of satellite observations which began in the late 1970s. This fact makes IRIS a unique data set to evaluate GCMs. The high spectral resolution ensures rich information of thermodynamic variables, such as temperature, humidity, and clouds, contained in the spectra. Besides the approach adopted in this study, IRIS and other data sets of spectrally resolved infrared radiances can be used to validate models in other ways [*Haskins et al.*, 1997; *Huang et al.*, 2002; *Huang and Yung*, 2005]. As demonstrated in this study, the observation of spectrally resolved radiances like IRIS can help us understand the discrepancies between broadband measurements and model simulation in more detail so that we can further narrow down the cause of such discrepancies. Limited by the data availability, this study is confined to the seasonal scale. With AIRS in operation and IASI to be launched in the future, spectrally resolved radiances can be used to test climate models over a variety of timescales. Moreover, the collocation of AIRS, MODIS and CERES measurements on NASA EOS platforms makes it possible for the first time to simultaneously use such collocated broadband, narrowband and spectrally resolved radiances to validate climate models. With a careful treatment of the sampling issues between satellite and climate

models, such hierarchy of diagnostic analyses will bring more insight to the deficiencies in the models and eventually help the improvement of climate simulation.

[37] **Acknowledgments.** We thank R. Goody, C. Crevoisier and B. Soden for valuable comments. We also wish to thank two anonymous reviewers for improving the quality of this paper. One of the authors, X.L. Huang, is supported by the AOS postdoctoral program at Princeton University.

References

Ackerman, S. A., W. L. Smith, J. D. Spinhirne, and H. E. Revercomb (1990), The 27–28 October 1986 FIRE IFO cirrus case study: Spectral properties of cirrus clouds in the 8–12 μm window, *Mon. Weather Rev.*, *118*(11), 2377–2388.

Allan, R. P., and M. A. Ringer (2003), Inconsistencies between satellite estimates of longwave cloud forcing and dynamical fields from reanalyses, *Geophys. Res. Lett.*, *30*(9), 1491, doi:10.1029/2003GL017019.

Allan, R. P., M. A. Ringer, J. A. Pamment, and A. Slingo (2004), Simulation of the Earth's radiation budget by the European Centre for Medium-Range Weather Forecasts 40-year reanalysis (ERA40), *J. Geophys. Res.*, *109*, D18107, doi:10.1029/2004JD004816.

Allan, R. P., A. Slingo, S. F. Milton, and I. Culverwell (2005), Exploitation of Geostationary Earth Radiation Budget data using simulations from a numerical weather prediction model: Methodology and data validation, *J. Geophys. Res.*, *110*, D14111, doi:10.1029/2004JD005698.

Anderson, G. P., S. A. Clough, F. X. Kneizys, J. H. Chetwynd, and E. P. Shuttle (1986), AFGL atmospheric constituent profiles (0–120 km), *Tech. Rep. AFGL-TR-86-0110*, 43 pp., Air Force Geophys. Lab., Hanscom Air Force Base, Mass.

Anderson, J. G., J. A. Dykema, R. M. Goody, H. Hu, and D. B. Kirk-Davidoff (2004), Absolute, spectrally-resolved, thermal radiance: A benchmark for climate monitoring from space, *J. Quant. Spectrosc. Radiat. Transfer*, *85*(3–4), 367–383.

Barkstrom, B. R. (1984), The Earth Radiation Budget Experiment (ERBE), *Bull. Am. Meteorol. Soc.*, *65*(11), 1170–1185.

Bernstein, L. S., A. Berk, P. K. Acharya, D. C. Robertson, G. P. Anderson, J. H. Chetwynd, and L. M. Kimball (1996), Very narrow band model calculations of atmospheric fluxes and cooling rates, *J. Atmos. Sci.*, *53*(20), 2887–2904.

Cess, R. D., and G. L. Potter (1987), Explorator studies of cloud radiative forcing with a general circulation model, *Tellus, Ser. A*, *39*, 460–473.

Collins, W. D., and A. K. Inamdar (1995), Validation of clear-sky fluxes for tropical oceans from the Earth radiation budget experiment, *J. Clim.*, *8*(3), 569–578.

Freidenreich, S. M., and V. Ramaswamy (1999), A new multiple-band solar radiative parameterization for general circulation models, *J. Geophys. Res.*, *104*(D24), 31,389–31,409.

GFDL Global Atmospheric Model Development Team (2004), The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations, *J. Clim.*, *17*(24), 4641–4673.

Goody, R. M., and Y. L. Yung (1989), *Atmospheric Radiation: Theoretical Basis*, 519 pp., Oxford Univ. Press, New York.

Hanel, R. A., V. Salomons, G. Wolford, I. Revah, C. Prabhaka, V. G. Kunde, and B. J. Conrath (1972), Nimbus 4 infrared spectroscopy experiment: 1. Calibrated thermal emission-spectra, *J. Geophys. Res.*, *77*(15), 2629–2641.

Harries, J. E., H. E. Brindley, P. J. Sagoo, and R. J. Bantges (2001), Increases in greenhouse forcing inferred from the outgoing longwave radiation spectra of the Earth in 1970 and 1997, *Nature*, *410*(6826), 355–357.

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), 16,563–16,581.

Huang, X. L., 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. L., 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.

Iacono, M. J., and S. A. Clough (1996), Application of infrared interferometer spectrometer clear sky spectral radiance to investigations of climate variability, *J. Geophys. Res.*, *101*(D23), 29,439–29,460.

Kiehl, J. T., and B. P. Briegleb (1992), Comparison of the observed and calculated clear sky greenhouse effect: Implications for climate studies, *J. Geophys. Res.*, *97*(D9), 10,037–10,049.

Kunde, V. G., B. J. Conrath, R. A. Hanel, W. C. Maguire, C. Prabhaka, and V. Salomons (1974), Nimbus-4 infrared spectroscopy experiment: 2. Comparison of observed and theoretical radiances from 425–1450 cm^{-1}, *J. Geophys. Res.*, *79*(6), 777–784.

Prabhakara, C., E. B. Rodgers, B. J. Conrath, R. A. Hanel, and V. G. Knude (1976), NIMBUS 4 infrared spectroscopy experiment: 3. Observations of lower stratospheric thermal structure and total ozone, *J. Geophys. Res.*, *81*(36), 6391–6399.

Prabhakara, C., R. S. Fraser, G. Dalu, M. L. C. Wu, R. J. Curran, and T. Styles (1988), Thin cirrus clouds—Seasonal distribution over oceans deduced from NIMBUS-4 IRIS, *J. Appl. Meteorol.*, *27*(4), 379–399.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi:10.1029/2002JD002670.

Schwarzkopf, M. D., and V. Ramaswamy (1999), Radiative effects of CH4, N2O, halocarbons and the foreign-broadened H_2O continuum: A GCM experiment, *J. Geophys. Res.*, *104*(D8), 9467–9488.

Slingo, A., J. A. Pamment, and M. J. Webb (1998), A 15-year simulation of the clear-sky greenhouse effect using the ECMWF reanalyses: Fluxes and comparisons with ERBE, *J. Clim.*, *11*(4), 690–708.

Uppala, S. M., et al. (2005), The ERA-40 reanalysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012, doi:10.1256/qj.04.176.

X. Huang, Program in Atmospheric and Oceanic Sciences, Princeton University, 300 Forrestal Road, Sayre Hall, P.O. Box CN710, Princeton, NJ 08544-0710, USA. (xianglei.huang@noaa.gov)

V. Ramaswamy and M. D. Schwarzkopf, Geophysical Fluid Dynamics Laboratory, NOAA, Princeton University Forrestal Campus, 201 Forrestal Road, Princeton, NJ 08540, USA.