Popular summary for
"The Continual Intercomparison of Radiation Codes: Results from Phase I"

by
Lazaros Oreopoulos (Lazaros.Oreopoulos@nasa.gov)
Eli Mlawer (AER Inc.) and co-authors

Prepared for the Journal of Geophysical Research-Atmospheres, August 2011

The computer codes that calculate the energy budget of solar and thermal radiation in Global Climate Models (GCMs), our most advanced tools for predicting climate change, have to be computationally efficient in order to not impose undue computational burden to climate simulations. By using approximations to gain execution speed, these codes sacrifice accuracy compared to more accurate, but also much slower, alternatives. International efforts to evaluate the approximate schemes have taken place in the past, but they have suffered from the drawback that the accurate standards were not validated themselves for performance. The manuscript summarizes the main results of the first phase of an effort called "Continual Intercomparison of Radiation Codes" (CIRC) where the cases chosen to evaluate the approximate models are based on observations and where we have ensured that the accurate models perform well when compared to solar and thermal radiation measurements. The effort is endorsed by international organizations such as the GEWEX Radiation Panel and the International Radiation Commission and has a dedicated website http://circ.gsfc.nasa.gov where interested scientists can freely download data and obtain more information about the effort's modus operandi and objectives.

In a paper published in the March 2010 issue of the Bulletin of the American Meteorological Society only a brief overview of CIRC was provided with some sample results. In this paper the analysis of submissions of 11 solar and 13 thermal infrared codes relative to accurate reference calculations obtained by so-called "line-by-line" radiation codes is much more detailed. We demonstrate that, while performance of the approximate codes continues to improve, significant issues still remain to be addressed for satisfactory performance within GCMs. We hope that by identifying and quantifying shortcomings, the paper will help establish performance standards to objectively assess radiation code quality, and will guide the development of future phases of CIRC.

The Continual Intercomparison of Radiation Codes: Results from Phase I

Lazaros Oreopoulos[1], Eli Mlawer[2], Jennifer Delamere[2], Timothy Shippert[3], Jason Cole[4], Boris Fomin[5], Michael Iacono[2], Zhonghai Jin[6,7], Jiangnan Li[4], James Manners[8], Petri Räisänen[9], Fred Rose[6,7], Yuanchong Zhang[10], Michael J. Wilson[11, 1], William Rossow[12]

1. Laboratory for Atmospheres, NASA-Goddard Space Flight Center, Greenbelt, Maryland, USA
2. Atmospheric and Environmental Research Inc., Lexington, Massachusetts, USA
3. Pacific Northwest National Laboratory, Richland, Washington, USA
4. Canadian Centre for Modeling and Analysis, Environment Canada, Victoria, British Columbia, Canada
5. Central Aerological Observatory, Moscow, Russia
6. Science Systems and Applications Inc., Hampton, Virginia, USA
7. Science Directorate, NASA Langley Research Center, Hampton, Virginia USA
8. British Met Office, Exeter, UK
9. Finnish Meteorological Institute, Helsinki, Finland
10. Dept. of Applied Physics and Applied Mathematics, Columbia University, New York, New York, USA
11. University of Maryland Baltimore County, Baltimore, Maryland, USA
12. Cooperative Remote Sensing Science and Technology Center, City College of New York, New York, New York, USA

Prepared for the Journal of Geophysical Research (Atmospheres)

August 2011

Corresponding author:
Lazaros Oreopoulos
NASA-GSFC, Code 613.2
Greenbelt MD 20771, USA
Lazaros.Oreopoulos@nasa.gov

1. Introduction

While we have high confidence in radiative forcing calculations due to enhanced concentrations of greenhouse gases when using high-resolution spectral radiation algorithms, their steep computational cost makes them presently unaffordable in Global Climate Models (GCMs). Calculations from such algorithms have instead been used as the basis for designing approximate but much faster codes to perform efficient gaseous radiative transfer (RT) in GCMs. When clouds, aerosols and reflective/emitting surfaces also partake in radiative interactions, uncertainties in radiative fluxes increase, not only because the spatiotemporal distribution of their physical properties is dubious, but also because their radiative properties themselves are approximated and parameterized. In the end, the simulation of solar and thermal radiative processes becomes a rather complex endeavour that burdens climate simulations with a substantial degree of uncertainty. Still, before tackling radiative transfer involving clouds and aerosols, the more straightforward and well-defined problem of gaseous absorption needs to be advanced. Unfortunately, despite the relatively well-settled status of spectrally detailed clear-sky radiative transfer and the significant share of CPU resources allocated to radiation in GCMs, the radiation codes in these models may still be inadequate in reproducing the radiative effects of increased greenhouse gases obtained by more spectrally detailed codes. For example, a recent intercomparison [*Collins et al.,* 2006] of well-mixed greenhouse gas forcing calculations between line-by-line (LBL) RT models and their speedier, but coarser, counterparts of GCMs used in the Intergovernmental Panel for Climate Change (IPCC) 4th Assessment Report, reported that for many of the cases analyzed, GCM codes exhibited "substantial discrepancies" relative to the detailed spectral LBL standards. The against which RT code performance is documented in scientific publications and coordinated joint modeling activities such as GCM intercomparisons. While it is understood that CIRC reference calculations at any time reflect current spectroscopic knowledge and may not be perfect, by keeping CIRC up-to-date with algorithmic and spectral database improvements as they become available, and by gradually expanding the effort with new cases, a valuable service to the radiation modeling community will exist for years to come.

This paper presents results from Phase I of CIRC, designed to test RT codes under relatively non-complex atmospheric conditions, i.e., either cloudless skies or skies completely covered by homogeneous liquid clouds. Submissions from 13 solar and 11 thermal infrared codes are analyzed against reference LBL calculations. Besides overall performance, we also delve into particular aspects of RT model behavior exposed by the specifications and requirements of individual cases. Simplified versions of the cases help isolate contributions of individual components to the overall errors. Before presenting the results, a detailed description of how the CIRC input and reference RT calculations were generated is provided in the next section.

2. The CIRC Phase I dataset

a. The cases

Table 1 provides a summary of the seven primary (baseline) cases used in Phase I to test RT algorithm performance. Two of the seven cases (Cases 6 and 7) include overcast liquid phase clouds with very different condensate amounts. The cloudy cases were selected for their apparent homogeneity as indicated by low temporal variability of the SW downwelling fluxes at the surface as measured by pyranometers. The high liquid radiance residuals of each RRTM band in order to convert them to flux residuals. Band residuals wer added to obtain broadband flux residuals. These results and accompanying plots can be found at the project website, http://circ.gsfc.nasa.gov/CIRC_cases.html. Further details on the construction of the cases are provided below.

With the intercomparison underway, CIRC participants suggested that the interpretation of model performance will be aided by including additional "subcases" that are simplified variants of the above seven baseline cases. These extra cases with simpler atmospheric and surface specifications would, of course, be no longer radiatively constrained by observations. The subcases were constructed by imposing one or more of the following simplifications: (a) spectrally invariant SW albedo; (b) no aerosol; (c) no cloud. A complete list of the subcases is provided in Table 2. A total of 16 SW subcases and 2 LW subcases resulted after applying various combinations of these simplifications (changes in surface albedo and aerosol only affected SW cases).

b. Input

A full list of input variables needed to perform RT calculations are provided at the project website, http://circ.gsfc.nasa.gov/CIRC_input.html. The input for six of the seven cases (the exception being Case 7) is based on v_1.4.1 of the ARM BBHRP evaluation dataset [*Mlawer et al., 2002*]. The features and data content of BBHRP most relevant to CIRC as well as the modifications employed to adapt the cases for the purposes of CIRC are provided below.

For the clear cases, the atmospheric column is discretized in layers of varying physical thickness, ranging from 54 m near the surface to 4 km for the uppermost layers.

September ozone values for this case. For all other species, mixing ratios are taken from the US Standard Atmosphere.

The cloud of Case 6 is based on ARM's Active Remotely-Sensed Clouds Locations (ARSCL) product [*Clothiaux et al., 2000*], which provides height distributions of hydrometeor reflectivity (and cloud boundaries) every 10 seconds based on observations from a Millimeter Cloud Radar (MMCR) and Micropulse Lidar (MPL). These ARSCL products are combined with thermodynamic profiles from radiosondes and column integrated water vapor estimates from the MWR and inserted into the Microbase cloud property retrieval algorithm [*Miller et al., 2003; U.S. Department of Energy, 2006*], which computes a time-height grid of the liquid water concentration, liquid effective radius, ice water concentration, and ice effective radius. Within the Microbase retrieval, the initial liquid water concentration data are integrated to produce an estimate of the Liquid Water Path (LWP) and then scaled by the ratio of the LWP retrieved from coincident MWR measurements and the initial Microbase LWP estimate. The retrieved cloud properties for each time and height are averaged over a 20-minute interval empirically deemed to encompass the cloud fields affecting the irradiance measurements used for the comparisons. For Case 7 (Pt. Reyes), the cloud property retrievals are based on the MIXCRA inversion algorithm [*Turner, 2005; Turner, 2007*]. The cloud was assumed to be vertically homogeneous and its top and base were determined from WACR (a W-band Doppler radar operating at 95 GHz) measurements.

For the SGP clear-sky cases (Cases 1-3) the aerosol optical depths are derived from spectral solar irradiance measurements of the Multi-Filter Rotating Shadowband Radiomenter (MFRSR) at 6 wavelengths below 1 μm. These MFRSR measurements, single tower is available for surface classification and albedo estimation. Based on a satellite (Multispectral Thermal Imager) image analysis of the region surrounding the NSA site in Barrow, Alaska on the day corresponding to Case 4, it was determined that an appropriate surface albedo would require taking the weighted average of the surface albedo below the tower (with a 85% weight) and the albedo of open water (with a 15% weight). For PYE Case 7 MODIS-derived surface (i.e. land) albedos are used to generate pseudo-MFR albedos that can subsequently yield a spectral albedo function in the exact same manner as for the SGP cases. For all cases, the surface reflectance is assumed to be Lambertian. Since different radiation codes have their own band structure, the surface albedo was provided at relatively high spectral resolution (1 cm^{-1}). Two additional spectral functions were provided, the product of surface albedo and the extraterrestrial spectral solar irradiance [*Kurucz et al.* 1992] in the 0.2-12.2 μm range, and the product of surface albedo with the 1 cm^{-1} downwelling surface irradiances from our reference LBL SW calculations. These spectral functions allow participants to calculate both unweighted or weighted (by the either the TOA or SFC spectral flux) albedos within the (wide) spectral bands appropriate for their RT code. The impact of the SFC albedo weighting scheme on TOA SW fluxes is discussed in subsection 4e.

Finally, the participants were provided with the downwelling broadband flux at the TOA calculated from the [*Kurucz et al.* 1992] spectral solar irradiance for the Sun-Earth distance and solar zenith angle at the day and time of each case, which is consistenst with the downwelling TOA broadband of the SW LBL calculations. When using this flux, any errors in the SW flux calculations can not be due to excessive or missing energy, but only factor and single-scattering albedo remained spectrally invariant. The aerosol information used in the CHARTS calculation was therefore exactly the same as that provided as input to the participants. The aerosol phase function was assumed to follow the Henyey-Greenstein function. Aerosol effects were ignored in the LW. The surface albedo was resolved at 1 cm^{-1} (as that provided to the participants) and was linearly interpolated to the wavenumber of the calculation. Finally, the surface emissivity was set to unity across the LW spectrum. The output fluxes of the calculations were integrated using a boxcar function into 1 cm^{-1} wide bins.

d. Radiative observations

The LW and SW surface observed irradiances for the CIRC SGP and NSA cases are 5-minute averages of values provided by ARM Best Estimate Radiative Flux VAP as included in the BBHRP v_1.4.1 and BBHRP v_1.4.1tK (Case 3) datasets. The 5-minute averaging window was centered at the time provided for each case. For the PYE case (Case 7) the LW irradiance measurement value is the mean of the 5-minute average from two pyrgeometers at the site. The PYE SW value is taken from the 5-minute average of measurements from the single collocated shaded pyranometer.

The observed TOA irradiance values are from the spatially and temporally closest Clouds and the Earth's Radiant Energy System (CERES) measurement for Case 4. For Cases 3 and 7, the broadband TOA fluxes are inferred from the spatially and temporally closest Geostationary Operational Environmental Satellite (GOES) using a radiance-to-flux and narrowband-to-broadband conversion algorithm. For the remaining SGP cases surface, i.e., without contributions of scattered radiation along the direction of the solar beam. No spectral results were solicited. The spectral range of LW and SW calculations was left to the participants, but under the condition that the radiation source is exclusively the Earth's surface and atmosphere for the LW and the Sun for the SW.

Some of the input information provided to conduct the CIRC runs is typically not available in an operational GCM environment, for example, spectral surface albedo. On the other hand, the available input may be incomplete for some RT algorithms, e.g., those requiring separate albedos for the ground and an overlying vegetation canopy. While the submissions solicited should ideally come from runs where the model uses as much of the information provided as possible, this may have required modifications of the RT algorithms from their standard operational configuration, so submissions where the algorithms operate with assumptions and input that more closely resemble default operational configuration were not discouraged. For example, the participants were allowed to perform runs where they would employ modifications such using the incident solar flux corresponding to the total solar irradiance (TSI) of the host model, or leaving scattering active in their LW algorithm even though there is no scattering in the reference LBL calculations.

Submissions have been received from 11 LW and 13 SW RT algorithms. The algorithms are identified in Tables 3 and 4, along with some pertinent information, such whether the algorithm is currently implemented in a Large Scale Model (LSM) and who the submitting party is. For some of the models multiple submissions were received, corresponding to different configurations of the runs (e.g., high/low accuracy modes, with and without LW scattering, different weightings of the spectral surface albedo

$$e_{n,i,j}(\%) = 100 \times \frac{F_{n,i}^{j} - F_{n,0}^{j}}{F_{n,0}^{j}} \qquad (1)$$

$F_{n,0}^{j}$ is the flux (irradiance) calculated with the reference LBL radiation code (LBLRTM for LW and LBLRTM/CHARTS for SW) for case *n*, and $F_{n,i}^{j}$ is the corresponding flux by RT model *i*. The flux types for LW are the upwelling radiation at the TOA, the downwelling at the SFC and the net flux divergence of the atmospheric column defined as the net (down minus up) at TOA minus net at SFC. For the SW, the flux types are as above, with the addition of the diffuse component of the downward flux at the SFC defined as the total flux minus the direct-horizontal flux. It should be pointed out that for the SW, the net flux divergence corresponds to the flux absorbed by the atmospheric column and underlying surface.

In the LW, the best overall simulated flux type appears to be the downward flux at the SFC, but there are two instances of flux errors outside the ±3% range (Cases 4 and 5 for Model 3). The TOA flux error is never outside the ±3% bounds. Model 11 exhibits a persistent overestimate of ~1.5% in the downward flux and an underestimate of ~2% in the TOA upward flux (including cloudy cases). Since the upward flux at the surface depends solely on surface temperature according to the Stephan-Boltzmann law and the downward LW flux at TOA is zero, the errors in net flux divergence (which relates to the total heating/cooling rate of the entire atmospherc column) can be smaller than the errors in the TOA up or SFC down fluxes if there is error cancellation. For about 70% of model-case combinations the net flux divergence indeed exhibits smaller errors than either of the TOA and SFC flux errors. A few models, namely models 1, 2, 4, 9, 10 (and 8 if cloudy Case 6 is excluded) maintain SFC and TOA flux errors within ±1% of LBL. But only are significantly better (with the exception of Model 11, Case 2b, the errors are always within ±5%). However, in the absence of clouds and aerosols, the underestimation of gaseous absorption is revealed more clearly, so the absorbed flux errors become larger especially in Cases 6 and 7. Only models 2 and 5 achieve absorption errors within ±2.5% for all the pristine cases.

b. Overall errors per model and flux quantity

A more compact view of model performance can be obtained by averaging errors over all cases, and these are shown in Figures 3 and 4. Specifically for each model *i* and flux type *j* we show the mean error, the mean absolute error, and the error standard deviation (shown as error bars) over all cases, defined respectively as:

$$\overline{e_i^j} = \frac{\sum_{n=1}^{N_c}\left(F_{n,i}^j - F_{n,0}^j\right)}{N_c} \tag{2}$$

$$\overline{\left|e_i^j\right|} = \frac{\sum_{n=1}^{N_c}\left|F_{n,i}^j - F_{n,0}^j\right|}{N_c} \tag{3}$$

$$\sigma_i^j(e) = \sqrt{\frac{\sum_{n=1}^{N_c}\left(F_{n,i}^j - F_{n,0}^j\right)^2}{N_c} - \overline{e_i^j}^2} \tag{4}$$

where N_c is the number of cases (including subcases), i.e., 9 for LW and 23 for SW (7 and 21 for model 6). All quantities are expressed in $\mathrm{Wm^{-2}}$. Cancellation of errors is allowed in the calculation of the means to recognize the fact that in an operating environment in which the RT codes are applied on a wide range of atmospheric input, the average performance should also be evaluated. On the other hand, the mean absolute where N_f is the number of flux types, and the subscript 0 refers to the LBL flux. We calculate the above error for $N_f=3$ flux types for both the SW and LW, i.e., we exclude the poorly simulated diffuse SW surface flux which suffers also from not being available for models 5 and 6. In addition to calculating this error for all 9 LW and 23 SW cases, we also perform a second calculation for only the subcases of pure molecular atmospheres, i.e., excluding cases 6 and 7 for the LW and excluding all but subcases 1(b), 2(b), 3(b), 4(b), 5(b), 6(d) and 7(b) for the SW.

The top two panels of Fig. 5 show the errors calculated by eq. (5). This type of error metric could potentially serve as a simple way to rank model performance and to set thresholds of acceptable RT code performance for particular applications. The bottom panels tell us whether the errors calculated from eq. (5) are smaller or larger when only simplified atmospheres and surfaces are considered. Reminding ourselves that in the LW the simplified atmospheres differ only in that clouds have been removed, it is not surprising that performance does not in general improve in the absence of purely absorbing clouds thick enough to have near unity emissivity. One the other hand, all models, with one exception, predictably improve their performance in the SW for the simplified cloudless atmospheres with no aerosols and spectral variations of surface albedo. The exception, model 12, while not consistently performing better for the pristine cases for all flux types, seems to suffer primarily from the fact that its absorbed flux performs notably worse, enough to push the error metric of eq. (5) above the values corresponding to the full basket of cases.

Another metric that can be illuminating is the percentage local mean deviation (LMD(%)), introduced originally by *Pinty et al.* [2004] in the RAMI intercomparison of when spectrally variably surface albedos are used in place of a spectrally flat albedo; (b) the difference in flux between two different ways of averaging the spectrally detailed albedo into coarse spectral intervals.

To gain insight into the first issue we calculate the TOA flux difference between the baseline and "a" subcases (i.e., the difference between cases with spectrally resolved albedo and spectrally flat albedo) averaged separately for the clear (Cases 1-5) and cloudy cases (Cases 6-7), i.e.,

$$\overline{\Delta F_i^{TOA}} = \frac{\sum_{n=1}^{N_c}\left(F_{n,i}^{TOA} - F_{n(\mathrm{a}),i}^{TOA}\right)}{N_c} \tag{7}$$

where N_c=5 for clear and N_c=2 for cloudy. Figure 7 shows these average differences, with the LBL results shown as dashed lines. To look into the second issue, we estimate for Cases 1-4 the TOA flux difference between calculations where the coarse-band albedo comes from weighting the spectral surface albedo values with the LBL spectral downwelling flux reaching the surface (these calculations were used for $F_{n,i}^{TOA}$ in eq. 7) and calculations where the incoming spectral solar radiation at the TOA was instead used to perform the weighting. Figure 8 shows these differences (positive indicates that the flux from surface flux weighting is larger) for those models that made results for both those surface albedo weighting options available.

The LBL results in Fig. 7 indicate that the average effect of surface albedo spectral variations is ~1.5 Wm^{-2} for cloudless atmospheres and ~-0.4 Wm^{-2} for cloudy atmospheres. The sign of the TOA flux difference is negative for Case 4 with its ice-driven surface albedo, and the two cloudy cases, and positive for all other cases resolved spectral albedo functions. The subtle effects of surface albedo averaging will emerge again in the CO_2 forcing analysis of subsection 4f.

e. Cloud and aerosol radiative effect errors

Cloud and aerosol radiative effect SW errors by the participating models can be best isolated by using the spectrally constant surface albedo subcases. We therefore define the error in LW and SW cloud radiative effect for model *i* and flux type *j* as:

$$e^{j}_{n-n(\mathrm{y}),i} = \left(F^{j}_{n,i} - F^{j}_{n(\mathrm{y}),i}\right) - \left(F^{j}_{n,0} - F^{j}_{n(\mathrm{y}),0}\right) \tag{8a}$$

$$e^{j}_{n(\mathrm{x})-n(\mathrm{y}),i} = \left(F^{j}_{n(\mathrm{x}),i} - F^{j}_{n(\mathrm{y}),i}\right) - \left(F^{j}_{n(\mathrm{x}),0} - F^{j}_{n(\mathrm{y}),0}\right) \tag{8b}$$

For the LW y = “a” in eq. (8a). To evaluate the SW cloud radiative effect error for Case 6 we select the aerosol-free subcases, so x = “b” and y = “d” in eq. (8b); for Case 7 which does not include aerosols x = “a” and y = “b”. These cloud radiative effect errors (both SW and LW) are shown in Fig. 9. LW cloud radiative effect errors remain within approximately ±5 Wm^{-2} for all flux types for both cloudy cases, with the exception of model 7 which is slightly outside this range, and are generally larger for the less optically thick cloud of Case 7. SW cloud radiative effect errors are much larger than their LW counterparts and are generally of greater magnitude for the optically thicker cloud of Case 6. Due to smaller absolute values and occassional error cancellations, absorbed flux radiative effect errors can be small even if TOA and SFC radiative effect errors are large. Overall, it appears that the radiative effect of the downwelling SFC flux is simulated slightly worse than the TOA radiative effect. Interestingly, for both the LW and SW provided for participants to average appropriately within their model's band structure, perhaps smaller cloud radiative effect errors would have resulted. However, this would not necessarily have been a better approach to evaluate cloud radiative effect estimation capabilities since in an operational environment such a level of detailed information on cloud radiative properties would not have been available.

f. Carbon dioxide forcing errors

The CO_2 forcing can be defined as either:

$$f_{5-4,i}^{j} = F_{5,i}^{j} - F_{4,i}^{j} \tag{10a}$$

or:

$$f_{5(x)-4(x),i}^{j} = F_{5(x),i}^{j} - F_{4(x),i}^{j} \tag{10b}$$

Eq. (10a) applies for both the LW and SW; eq. (10b) applies only for the SW with either x="a" or x="b". In the SW therefore three CO_2 forcing error calculations are possible, one that corresponds to the baseline case and two that correspond to the two subcases (spectrally flat albedo with aerosol and spectrally flat albedo with no aerosol). We chose to show in this subsection not the errors, but the forcings themselves in order to highlight an issue related to the sign of the SW TOA forcing. These are shown in Fig. 11 where the dashed lines depict the LBL reference results. In the LW, models divide in almost equal numbers to those that underestimate and those that overestimate the SFC and COL (flux divergence) forcing. When it comes to TOA forcing, however, only one model yields a notable overestimate. The magnitudes of the LW CO_2 forcing are quite good for most models, typically less than 0.5 Wm^{-2} different from LBL. When comparing the forcing (i.e. Case 5 reflecting more than Case 4 despite the larger CO_2 concentration) for this exact reason, a higher near infrared surface albedo for Case 5 due to downwelling surface flux weightings that have changed from Case 4 (this has been previously pointed out by *Oreopoulos and Mlawer* 2010). Since in an operational GCM environment, the effect of CO_2 on band-average surface albedo would most probably be neglected a zero CO_2 TOA forcing would result.Overall, the quality of the results in Figure 11 varies greatly. Even when concentrating on the CO_2 forcing of the spectrally flat albedo subcases, deviations from LBL calculations are quite large, as was previously found by *Collins et al.* [2006]. The poorest performer appears to be model 9. For this model CO_2 doubling apparently implicitly increases the concentration of some of the other uniformly mixed absorbing gases, O_2, N_2O, CH_4 and CO. When looking at the spectral response of this model to doubling CO_2, one sees flux changes in spectral intervals where CO_2 should not be active. Obviously, it should be made a priority that deficiencies like this are eliminated from radiation codes used in climate models.

g. Longwave scattering

The LBLRTM code used to generate the LW reference results does not account for cloud scattering. Still, it is instructive to have a general idea on the impact of scattering for our particular cloudy cases. Here we examine the scattering effects of those participating models (4, 5 and 10) that provided submissions with and without LW scattering effects. The comparison is shown in Fig. 12; the no-scattering results from the LBL code are also included for reference. To facilitate visualization and comparison of the differences between scattering and no-scattering results for both cloudy cases and the two fluxes at

h. LW heating rate errors

Heating (cooling) rate errors are only calculated for the LW since no reference LBL SW heating rate profiles are available. We use only the five original clear-sky cases and the two cloudy cases in order to put models 5 and 6, which did not submit results for the two LW subcases, on equal footing with the other models. The added vertical dimension makes the evaluation of heating rate profile errors somewhat more challenging than the column boundary fluxes we have been dealing with so far. We settled on using a mass-weighted [e.g., *Räisänen and Barker*, 2004] heating rate (HR) root mean square error (rmse) for each model *i* calculated as follows:

$$rmse_i(HR) = \sqrt{\frac{\sum_{n=1}^{N_c}\sum_{l}\left(HR_{n,i}^{l} - HR_{n,0}^{l}\right)^2 \Delta p_n^l}{\sum_{n=1}^{N_c}\sum_{l}\Delta p_n^l}} \qquad (11)$$

where the heating rate $HR_{n,i}^{l}$ (cooling rate when a negative value is obtained) of model *i* in layer *l* for case *n* is given in (K/day) by:

$$HR_{n,i}^{l} = 86400(s/day) \times \frac{g}{c_p}\frac{\Delta F_{n,i}^{l}}{\Delta p_n^l} \qquad (12)$$

Δp_n^l is the pressure thickness of layer *l* for case *n*, c_p is the specific heat of air at constant pressure, *g* is the acceleration of gravity constant, and $\Delta F_{n,i}^{l}$ is the flux divergence of model *n* in layer *l* of case *n*. As before, the index *0* is reserved for the reference LBL model. For the clear-sky cases we calculate the HR rmse separately for the parts of the atmosphere below and above 200 hPa (a proxy separator between troposphere and evaluation of submissions are publicly available for download on the CIRC website http://circ.gsfc.nasa.gov. Provided that CIRC participants will consent in having their submissions be posted on the CIRC website, any interested individual will be able to perform their own code evaluation and examine aspects of code performance that we did not cover in this paper.

Our analysis has revealed a number of intriguing findings. We found that errors in SW simulations, which have more degrees of freedom and parameters to specify, were larger than LW errors. We also found that diffuse and absorbed SW fluxes are particular areas of concern. Obtaining the correct breakdown of total to direct and diffuse may be important for the simulation of chemical or surface processes in climate models. Previously found underestimates of SW absorption by less spectrally detailed models [*Ackerman et al., 2003*] seem to be confirmed here. Another finding was that the number of bands available to resolve spectral surface albedo and other details of wide-band averaging can be important, so model developers need to pay attention to their representations of spectral albedo variations. SW CO_2 forcing, which should be included in all models, needs to be better simulated for those models that include it; LW CO_2 forcing has been more scrutinized and is quite well simulated by thermal radiation codes. Finally, while LW fluxes may be quite adequately simulated at the atmospheric column boundaries, net flux divergences within the atmosphere that determine heating (or cooling) rates may need further attention. In addition, scattering by clouds in the LW can potentially have measurable contributions and should ideally be included. We hope that scattering capability will soon be available to all LBL codes as well (some, like model 5 of this study already have it).

(project number 127210). The endorsement of the GEWEX Radiation Panel and the International Radiation Commission is critical for the success of CIRC and we extend our thanks to their respective leaders C. Kummerow (GRP) and R. Cahalan (IRC) for their guidance. We would like to acknowledge the large number of people involved in CIRC indirectly by developing the ARM data products used specifically, M. Miller for leading the cloud retrieval effort used in BBHRP, D. Turner and C. Chiu for CIRC Case 7 cloud and surface input, B. Zak for providing satellite images relevant to CIRC Cases 4 and 5, and M. Khaiyer and P. Minnis for the GOES and CERES satellite-based radiative fluxes.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown (2005), Atmospheric radiative transfer modeling: a summary of the AER codes, *J. Quant. Spectrosc. Radiat. Transfer*, *91*, 233-244.

Collins, W. D. (2001), Parameterization of generalized cloud overlap for radiative calculations in general circulation models, *J. Atmos. Sci.*, *58*, 3224–3242.

Collins, W. D., J. K. Hackney, and D. P. Edwards (2002), An updated parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model, *J. Geophys. Res.*, *107(D22)*, 4664, doi:10.1029/2001JD001365.

Collins, W. D. and coauthors (2004), Description of the NCAR Community Atmosphere Model (CAM 3.0). NCAR Techinical Note NCAR/TN-464+STR, 226 pp.

Collins W. D., and coauthors (2006), Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4), *J. Geophys. Res.*, *111*, D14317, doi:10.1029/2005JD006713.

Costa S. M. S. and K. P. Shine (2006), An estimate of the global impact of multiple scattering by clouds on outgoing long-wave radiation, *Quart. J. Roy. Meteor. Soc.*, *132*, 885-895.

Delamere, J. S., S. A. Clough, V. H. Payne, E. J. Mlawer, D. D. Turner and R. R. Gamache (2010), A far-infrared radiative closure study in the Arctic: Application to water vapor, *J. Geophys. Res.*, *115*, D17106, doi:10.1029/2009JD012968.

Fomin, B.A. (2004), A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 1. FKDM, fast k-distribution model for the longwave, *J. Geophys. Res., 109*, D02110, doi:10.1029/2003JD003802.

Fomin, B.A. and M.P. Correa (2005), A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 2. FKDM, fast k-distribution model for the shortwave, *J. Geophys. Res., 110*, D02106, doi:10.1029/2004JD005163.

Fomin, B.A. (2006), Monte-Carlo algorithm for line-by-line calculations of thermal radiation in multiple scattering layered atmospheres, *J. Quant. Spectr. Rad. Trans. 2471*, doi:10.1016/j.jqsrt.2005.05.078, 2006.

Hu, Y. X and K. Stamnes (1993), An accurate parameterization of the radiative properties of water clouds suitable for use in climate models, *J. Clim. 6*, 728-742.

Iacono M.J., J.S. Delamere, E. J. Mlawer, M.W. Shephard, S.A. Clough, and W. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res., 113*, D13103, doi:10.1029/2008JD009944.

Illingworth, A. J., and coauthors (2007), Cloudnet—Continuous evaluation of cloud profiles in seven operational models using ground-based observations, *Bull. Amer. Meteor. Soc., 88*, 883–898.

IPCC (2007), Climate Change (2007), The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)], *Cambridge University Press*, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

Mlawer, E.J., S.J. Taubman, P.D. Brown, M.J. Iacono and S.A. Clough (1997), RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res., 102*, 16,663-16,682.

Mlawer, E.J., and coauthors (2002), The Broadband Heating Rate Profile (BBHRP) VAP, paper presented at Twelfth Atmospheric Radiation Measurement (ARM) Science Meeting, U.S. Dep. of Energy, Washington, D. C.

Moncet, J.-L., and S. A. Clough (1997), Accelerated monochromatic radiative transfer for scattering atmospheres: Application of a new model to spectral radiance observations, J. Geophys. Res., 102, 21,853– 21,866, doi:10.1029/97JD01551.

Oreopoulos, L., and E. Mlawer, (2010), The Continual Intercomparison of Radiation Codes (CIRC): Assessing anew the quality of GCM radiation algorithms, *Bull. Amer. Meteor. Soc.*, *91*, 305-310.

Pinty, B., and coauthors (2004), Radiation Transfer Model Intercomparison (RAMI) exercise: Results from the second phase, *J. Geophys. Res.*, *109*, D06210, doi:10.1029/2003JD004252.

Räisänen, P., and H. W. Barker (2004), Evaluation and optimization of sampling errors for the Monte Carlo Independent Column Approximation. *Quart. J. Roy. Meteor. Soc., 130*, 2069–2085.

Stokes, G. M., and S. E. Schwartz (1994), The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed, *Bull. Amer. Meteor. Soc., 75*, 1201–1222.

Tarasova, T. A., and B.A. Fomin (2007), The Use of New Parameterizations for Gaseous Absorption in the CLIRAD-SW Solar Radiation Code for Models, *J. Atmos. Oceanic Technol., 24,* 1157–1162.

Table 1. Synopsis of the seven CIRC Phase I primary (baseline) cases. The four rightmost columns show observed and LBL-calculated (in bold) flux values (in Wm^{-2}) at the surface (SFC) and the top-of-the atmosphere (TOA) for both the thermal/longwave (LW) and solar/shortwave (SW) part of the spectrum. Observed TOA fluxes are from GOES using narrowband to broadband conversion algorithms or from CERES (case 4), while observed SFC fluxes come from ARM instruments. The first six columns provide some essential input information (SZA=Solar Zenith Angle, PWV=Precipitable Water Vapor, LWP=Liquid Water Path). The aerosol optical depth (τ_{aer}) is for 0.55 μm. Case 5 is as Case 4, but with doubled CO_2.

Date (Site)	**Case**	***SZA***	***PWV* *(cm)***	τ_{aer}	***LWP* *(gm^{-2})***	LW_{SFC}	LW_{TOA}	SW_{SFC}	SW_{TOA}
September 25, 2000	1	47.9°	1.23	0.04		289.7	301.7	705.9	169.8
(SGP)						**288.2**	**304.3**	**701.2**	**175.0**
July 19, 2000 (SGP)	2	64.6°	4.85	0.18		441.8	288.6	345.4	127.8
						439.3	**292.6**	**348.0**	**117.1**
May 4, 2000 (SGP)	3	40.6°	2.31	0.09		336.4	277.6	772.5	159.6
						333.0	**280.8**	**773.1**	**173.6**
May 3, 2004 (NSA)	4	55.1°	0.32	0.13		194.7	229.1	638.9	425.8
						192.4	**230.5**	**642.8**	**422.9**
May 3, 2004 (NSA, CO_2)	5	55.1°	0.32	0.13					
						195.7	**229.2**	**641.3**	**422.7**
March 17, 2000 (SGP)	6	45.5°	1.90	0.24	263.4	339.0	234.8	97.6	623.2
						335.2	**241.8**	**92.1**	**628.8**
July 6, 2005 (PYE)	7	41.2°	2.42		39.1	373.2	284.0	479.8	356.0
						372.6	**280.2**	**473.7**	**356.4**

Table 3. LW codes participating in CIRC Phase I.

Model Index	Brief Model Description	In LSM?	Experiment variants	Submitted By	Reference(s)
0	LBLRTM v.11.1/HITRAN 2004, MT_CKD_2.0, AER_V_2.0	No	None	Delamere, Mlawer	Clough et al. (2005)
1	RRTM-LW, 10-3000 cm^{-1}, CKD, 16 bands, 256 g-points	No	None	Iacono, Mlawer	Mlawer et al. (1997); Clough et al. (2005);
2	RRTMG-LW, 10-3000 cm^{-1}, CKD, 16 bands, 140 g-points	Yes	None	Iacono	Mlawer et al. (1997); Iacono et al. (2008)
3	CLIRAD-LW, 0-3000 cm^{-1}, k-distribution and one-parameter scaling, 10 bands, 85/113 k-points	Yes	"High/Low" Accuracy	Oreopoulos	Chou et al. (2003)
4	CCC 0-2500 cm^{-1}, CKD, 9 bands, 56 g-points	Yes	With/without scattering	Cole, Li	Li (2002); Li and Barker (2002); Li and Barker (2005);
5	FLBLM, 40-3000 cm^{-1}, line-by-line,	No	None	Fomin	Fomin (2006)
6	FKDM, 40-3000 cm^{-1}, CKD, 23 g-points	No	None	Fomin	Fomin (2004)
7	CAM 3.1, 0-2000 cm^{-1}, absorptiviy-emissivity approach	Yes	Treatment of lowest level air temperature	Oreopoulos	Collins et al. (2004)
8	FLCKKR (LW), 0-2200 cm^{-1}, CKD, 12 bands, 67 g-points	No	None	Rose, Kratz, Kato, Charlock	Fu and Liou (1992); Fu et al. (1997)
9	RRTMG-LW (as implemented in FMI/ECHAM5.4), 10-3000 cm^{-1}, 16 bands, 140 g-points	Yes	None	Räisänen	Mlawer et al. (1997); Iacono et al. (2007)
10	ES, 10-3000 cm^{-1}, 9 bands/33 g-points, ESF of band transmissions	Yes	With/without scattering	Manners	Edwards and Slingo (1996); Edwards (1996)
11	NASA-GISS, 50-2000 cm^{-1}, CKD, 33 g-points	Yes	None	Zhang, Rossow, Lacis	Zhang et al. (2004)

List of Figures

Figure 1. Percentage errors as defined by Eq. (1) of each participating model for each case for LW upward flux at TOA, downward flux at SFC and flux divergence. Gray indicates unavailability of submissions. Errors outside the colorbar range are assigned the extreme colors of the colorbar.

Figure 2. As Fig. 1, but for SW. In addition, the diffuse flux at the surface (difference between total flux and the direct solar beam flux) is shown. The flux divergence in this case correspons to the atmospheric absorption. Gray indicates unavailability of submissions. Errors outside the colorbar range are assigned the extreme colors of the colorbar. Model 6a is a variant of Model 6 where the optical properties of the scatterers are described in greater spectral detail within the k-distribution bands following Fomin and Correa (2005).

Figure 3. Mean errors in Wm^{-2} over all cases of the three flux types of Fig. 1 depicted as gray bars, their associated standard deviations depicted as error bars, and mean absolute errors depicted as black bars. There are calculated from eq. (2) , (4), and (3) respectively for each participating model and the three radiative fluxes of Fig. 1.

Figure 4. As in Fig. 3, but for the four flux types of Fig. 2.

Figure 5. The total error of each participating model according to eq. (5) for all LW and SW cases (upper panels, left for LW and right for SW). The bottom panels (again, left for LW and right for SW) show the difference between total errors calculated for the full basket of cases (i.e., top panels) and total errors calculated for a subset of the cases, namely the seven cloudless (clear) LW cases and the seven pristine SW cases (no aerosol, spectral flat surface albedo) of Table 2.

aerosol minus fluxes with no aerosol, specifically the difference between subcase "a" minus subcase "b" fluxes for Cases 1-5 and subcase "c" minus subcase "d" fluxes for Case 6 (there is no aerosol in Case 7), see. The radiative effect errors depicted are averages over all 6 cases, see eq. (9). TOA, SFC and COL flux designations are as in the previous figure.

Figure 11. CO_2 forcing (flux difference between cases with 375 ppm CO_2 and 750 ppm CO_2) for LW (upper left panel) and SW (remaining panels). The LBL forcings are indicated by the dashed lines: for the LW plot the black line indicating positive value is for the forcing in atmospheric flux divergence ("COL"), the gray line is for forcing of upwelling TOA flux, and the black line indicating negative value is for the downwelling flux at SFC; for the SW plots, the lines indicating positive values are for TOA (gray) and SFC (black), and the line indicating negative values is for absorptance ("COL"). Further explanations are given in the text.

Figure 12. Comparison between LW fluxes at TOA and SFC with and without scattering for the participating models that submitted both types of calculations. The LBL reference fluxes also shown were calculated only without scattering.

Figure 13. The mass-weighted LW heating rate (HR) root mean square error of each particpating model derived from eq. (9). Three HR rmse's are calculated: For the original five clear cases below 200 mb; for the original five clear cases above 200 mb; and for the two cloudy cases below 200 mb.

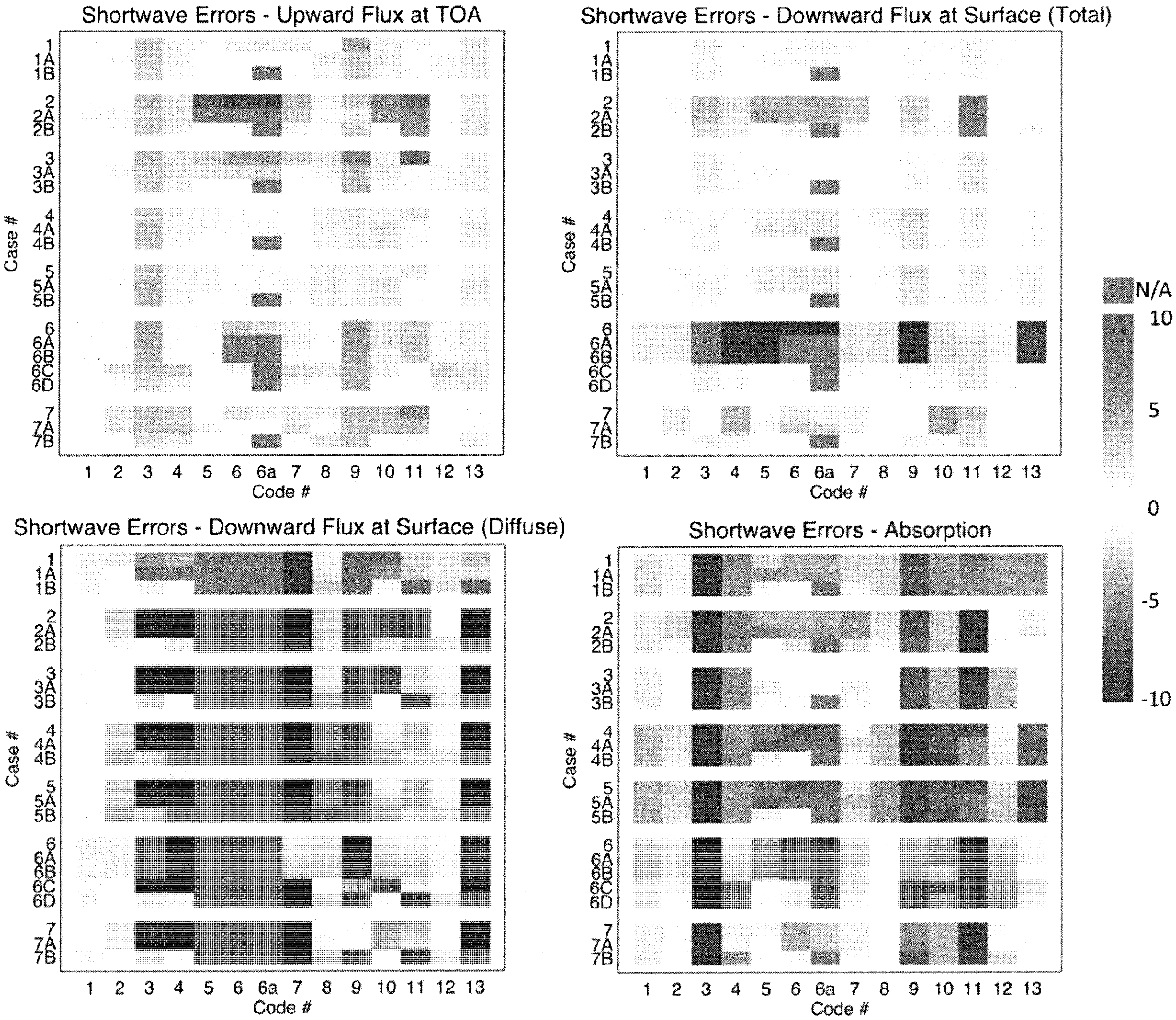
Shortwave Errors - Upward Flux at TOA
Shortwave Errors - Downward Flux at Surface (Total)
Shortwave Errors - Downward Flux at Surface (Diffuse)
Shortwave Errors - Absorption
Case #
1
1A
1B
2
2A
2B
3
3A
3B
4
4A
4B
5
5A
5B
6
6A
6B
6C
6D
7
7A
7B
Code #
1
2
3
4
5
6
6a
7
8
9
10
11
12
13
N/A
10
5
0
-5
-10

Figure 2

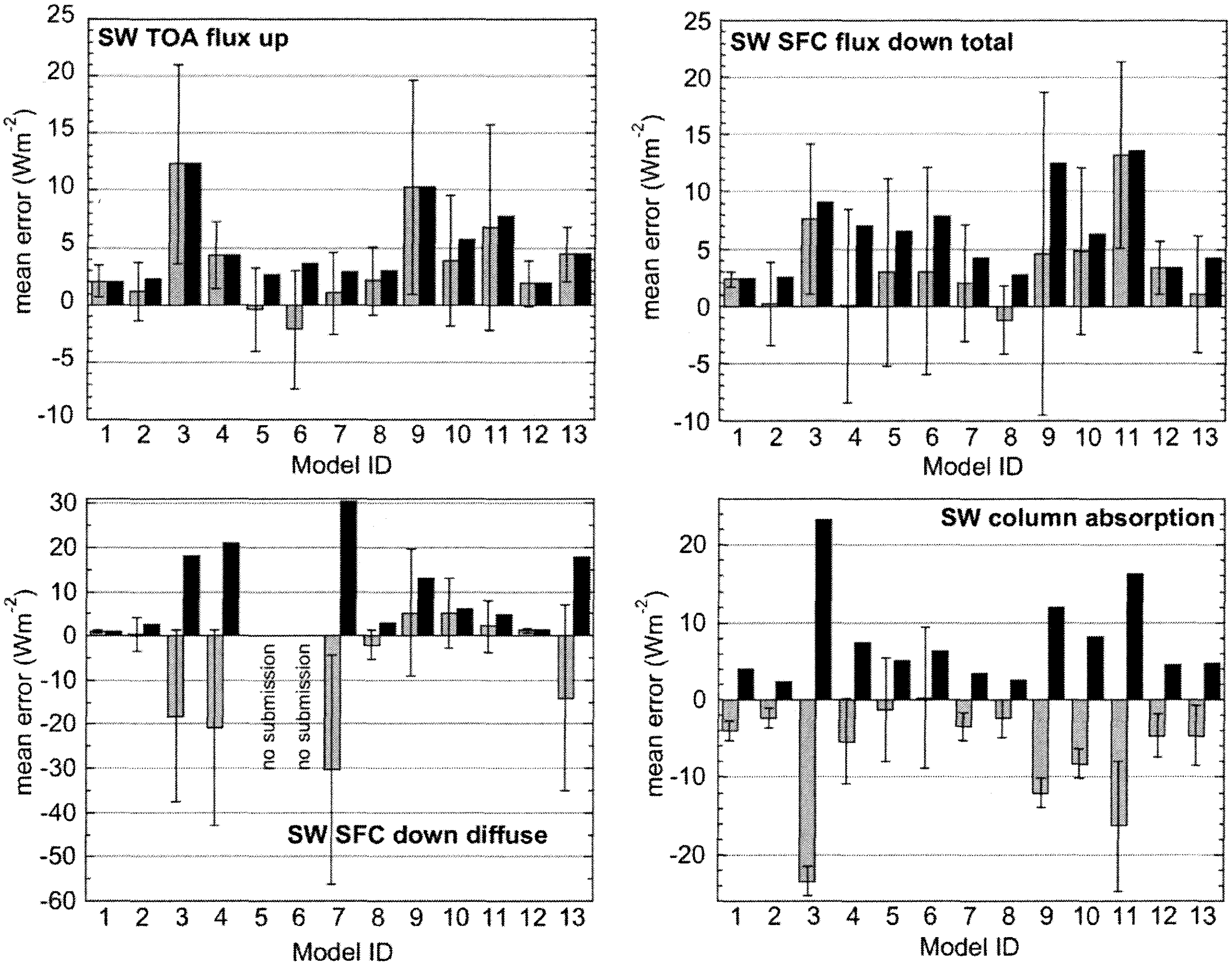
SW TOA flux up
SW SFC flux down total
SW SFC down diffuse
SW column absorption
mean error (Wm^{-2})
Model ID
1 2 3 4 5 6 7 8 9 10 11 12 13
25
20
15
10
5
0
-5
-10
30
-20
-30
-40
-50
-60
no submission
no submission

Figure 4

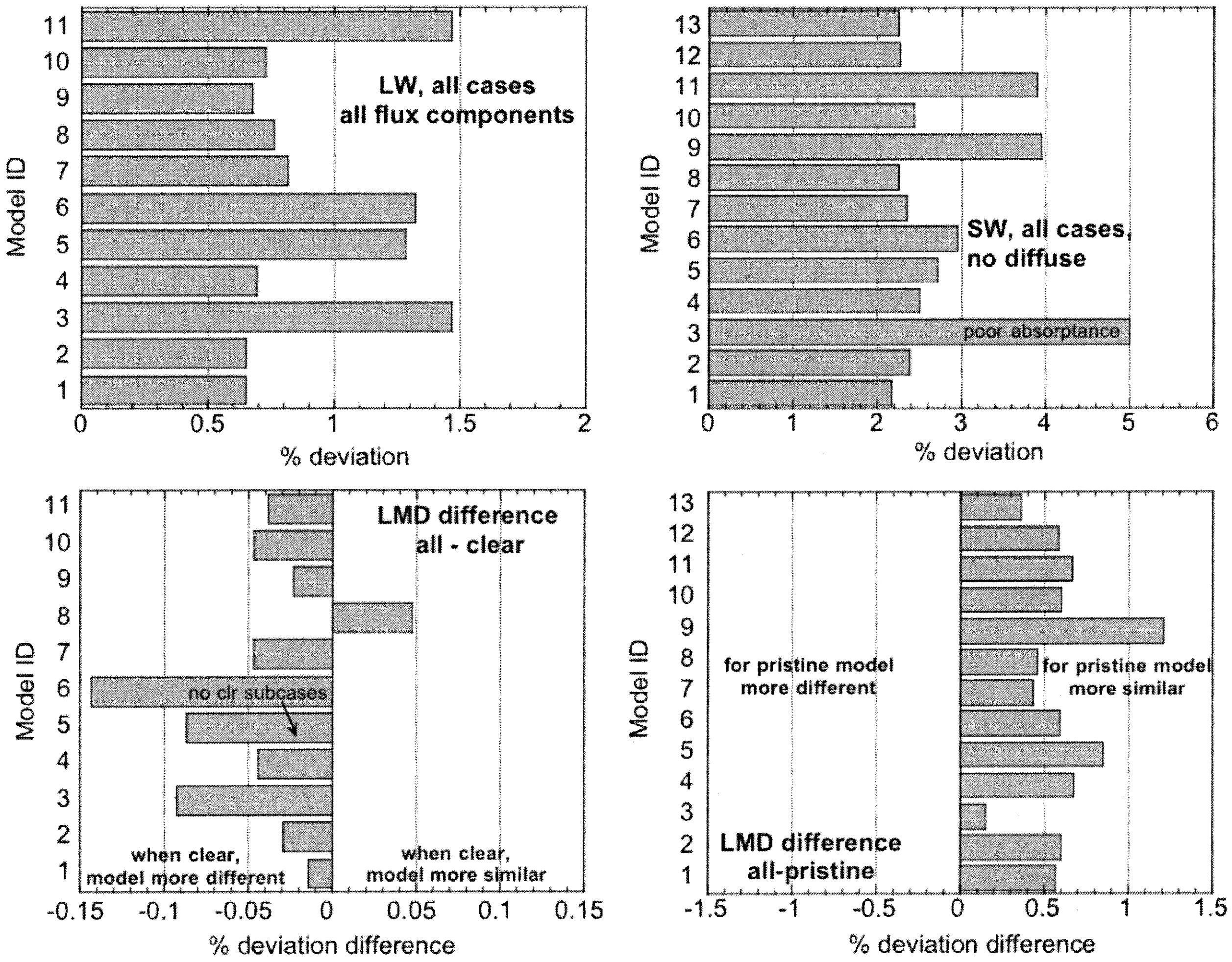
LW, all cases
all flux components
Model ID
% deviation
0
0.5
1
1.5
2
SW, all cases,
no diffuse
poor absorptance
0
1
2
3
4
5
6
LMD difference
all - clear
no clr subcases
when clear,
model more different
when clear,
model more similar
% deviation difference
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
for pristine model
more different
for pristine model
more similar
LMD difference
all-pristine
-1.5
-1
-0.5
0
0.5
1
1.5

Figure 6

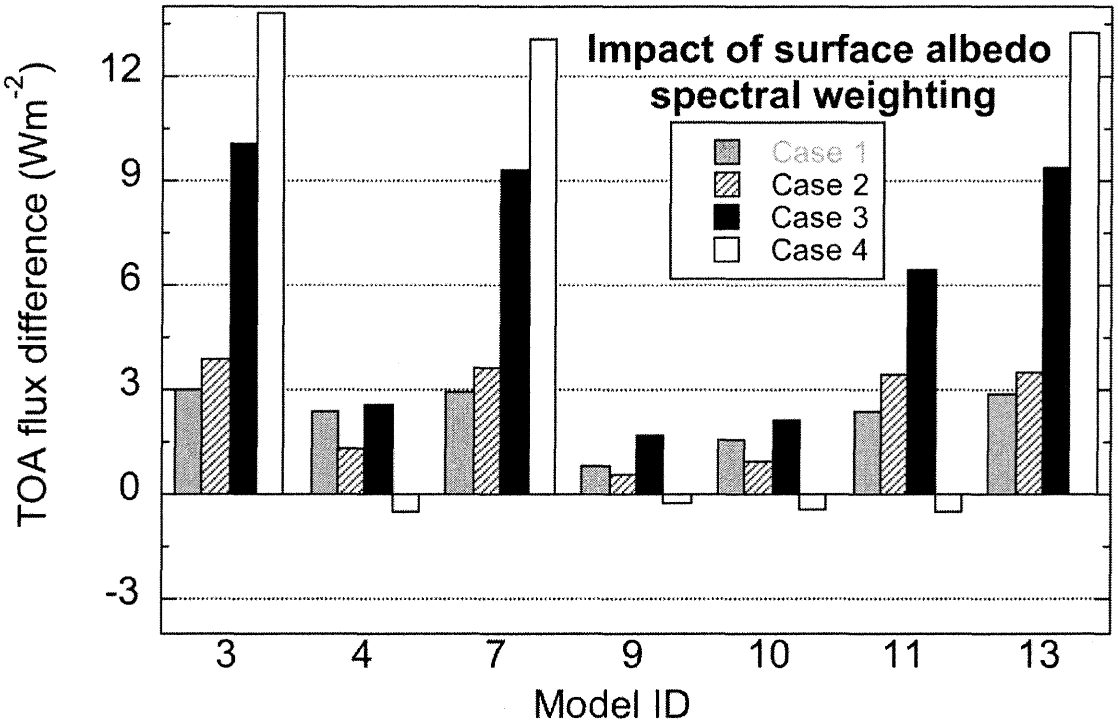
Impact of surface albedo spectral weighting
Case 1
Case 2
Case 3
Case 4
TOA flux difference (Wm^{-2})
12
9
6
3
0
-3
3
4
7
9
10
11
13
Model ID

Figure 8

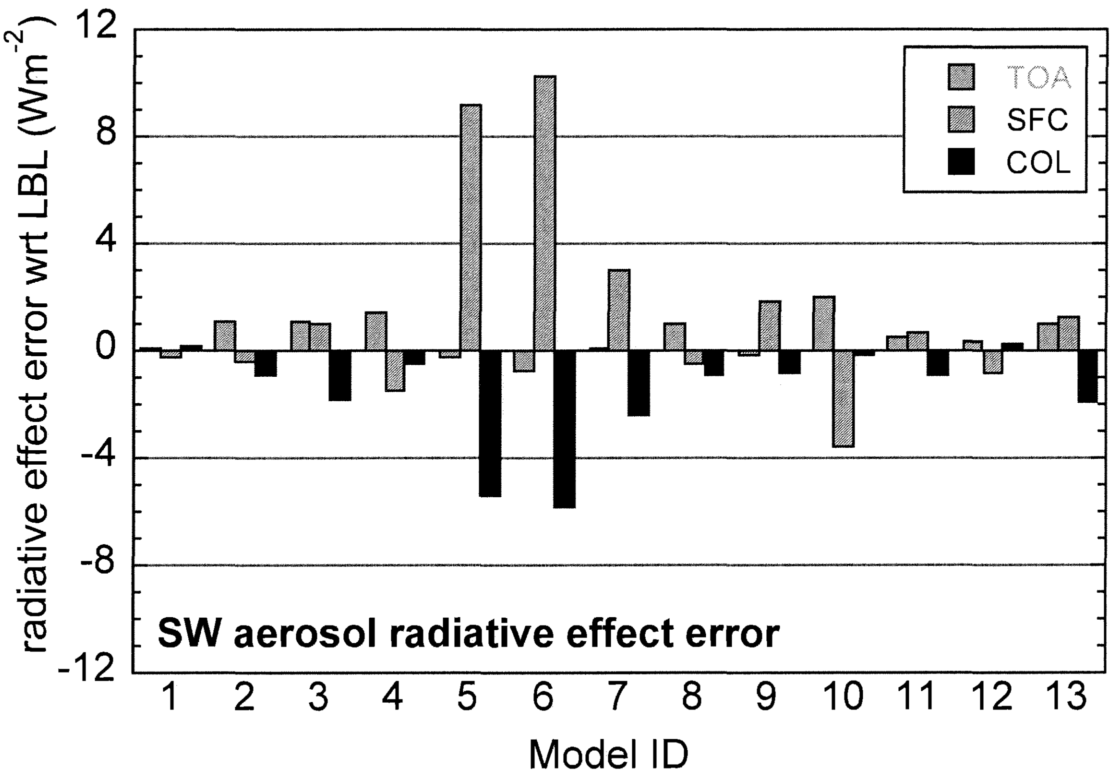
radiative effect error wrt LBL (Wm^{-2})
12
8
4
0
-4
-8
-12
TOA
SFC
COL
SW aerosol radiative effect error
1
2
3
4
5
6
7
8
9
10
11
12
13
Model ID

Figure 10

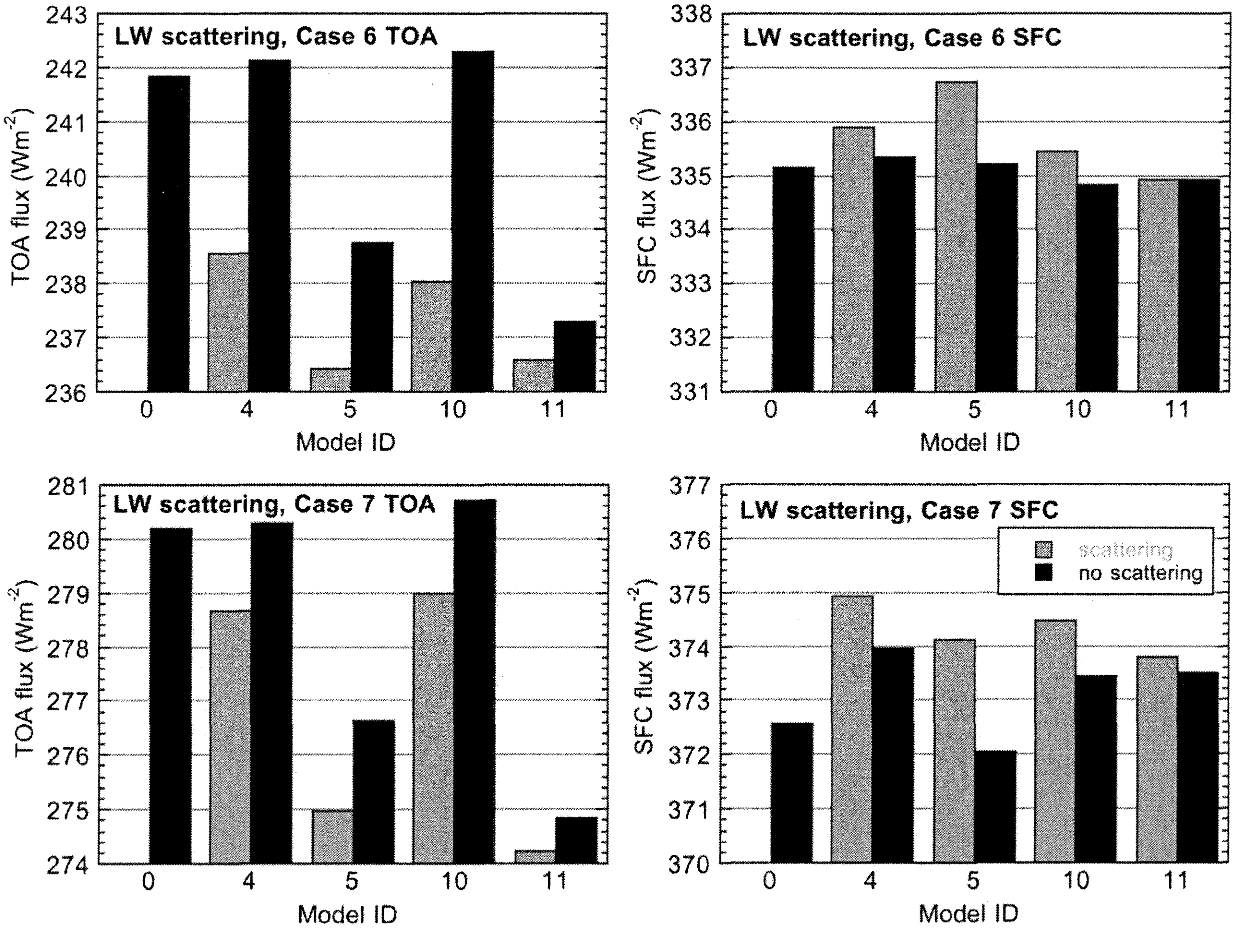
LW scattering, Case 6 TOA
TOA flux (Wm^{-2})
243
242
241
240
239
238
237
236
0
4
5
10
11
Model ID
LW scattering, Case 6 SFC
SFC flux (Wm^{-2})
338
337
336
335
334
333
332
331
0
4
5
10
11
Model ID
LW scattering, Case 7 TOA
TOA flux (Wm^{-2})
281
280
279
278
277
276
275
274
0
4
5
10
11
Model ID
LW scattering, Case 7 SFC
SFC flux (Wm^{-2})
377
376
375
374
373
372
371
370
0
4
5
10
11
Model ID
▨ scattering
■ no scattering

Figure 12