JGR Atmospheres

RESEARCH ARTICLE
10.1029/2019JD032323

Key Points:

- Using radiative forcing model intercomparison project (RFMIP's) protocol to assess biases in aerosol instantaneous radiative effect (IRE), contrasting sources of error found between the GFDL AM4 and RRTMG CESM 1.2.2 models
- Two sources of error in modeled IRE relative to benchmarks are: scattering methods and spectral resolution of shortwave parameterizations
- Results demonstrate how the RFMIP protocol can help determine the origins of model biases in aerosol IRE for other participating models

Correspondence to:
D. Paynter,
David.Paynter@Noaa.gov

Citation:
Freidenreich, S., Paynter, D., Lin, P., Ramaswamy, V., Jones, A. L., Feldman, D., & Collins, W. D. (2021). An investigation into biases in instantaneous aerosol radiative effects calculated by shortwave parameterizations in two Earth System Models. *Journal of Geophysical Research: Atmospheres*, *126*, e2019JD032323. https://doi.org/10.1029/2019JD032323

Received 26 DEC 2019
Accepted 28 APR 2021

An Investigation Into Biases in Instantaneous Aerosol Radiative Effects Calculated by Shortwave Parameterizations in Two Earth System Models

S. Freidenreich[1], D. Paynter[1], P. Lin[1], V. Ramaswamy[1], A. L. Jones[2], D. Feldman[3], and W. D. Collins[4]

[1]NOAA/GFDL, Princeton, NJ, USA, [2]University of Maryland, College Park, MD, USA, [3]Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [4]Department of Earth and Planetary Science, University of California, Berkeley, CA, USA

Abstract Because the forcings to which Coupled Model Intercomparison Project - Phase 5 (CMIP5) models were subjected were poorly quantified, recent efforts from the Radiative Forcing Model Intercomparison Project (RFMIP) have focused on developing and testing models with exacting benchmarks. Here, we focus on aerosol forcing to understand if for a given distribution of aerosols, participating models are producing a radiometrically-accurate forcing. We apply the RFMIP experimental protocol for assessing flux biases in aerosol instantaneous radiative effect (IRE) on two participating models, GFDL AM4 and CESM 1.2.2. The latter model contains the RRTMG radiation code which is widely used among CMIP6 GCM's. We conduct a series of calculations that test different potential sources of error in these models relative to line-by-line benchmarks. We find two primary sources of error: two-stream solution methods and the techniques to resolve spectral dependencies of absorption and scattering across the solar spectrum. The former is the dominant source of error for both models but the latter is more significant as a contributing factor for CESM 1.2.2. Either source of error can be addressed in future model development, and these results both demonstrate how the RFMIP protocol can help determine the origins of parameterized errors relative to their equivalent benchmark calculations for participating models, as well as highlight a viable path towards a more rigorous quantification and control of forcings for future CMIP exercises.

1. Introduction

Despite the tremendous efforts expended by modeling centers and information obtained regarding climate modeling from the Coupled Model Intercomparison Project - Phase 5 (CMIP5) (Taylor et al., 2012), there were some notable areas where future CMIP exercises could be improved, including the specification and reporting of radiative forcing (Soden et al., 2018; Stouffer et al., 2017). The coordinated activities and experimental design of the Radiative Forcing Model Intercomparison Project (RFMIP) are configured to improve understanding of modeled radiative forcing (Pincus et al. 2016). As part of CMIP6, RFMIP seeks to provide a basis for quantifying uncertainties in radiative forcing among models, including the specific contribution from aerosols, through controlled experiments and benchmark calculations. The motivation behind this is that large differences in radiative forcing exist among models even when the model ensemble members have identically-prescribed changes in the surface or atmospheric states. The differences can be partially the result of how models translate perturbations into forcing through individual model radiative parameterizations. To ensure that future model intercomparison experiments test model variations in response to identical forcing rather than testing a convolution of varied forcing and varied response, accurate clear-sky shortwave radiative transfer parameterizations relative to accurate benchmarks is a necessary first step. Unlike previous efforts such as the Continual Intercomparison of Radiation Codes (Oreopoulos & Mlawer, 2010), which evaluate code performance under controlled conditions, it is necessary to evaluate these parameterizations as they are being implemented in models, with native error diagnostics, to understand and address the dominant source(s) of error (Jones et al., 2017). This is particularly important where forcing is spatially and/or temporally heterogeneous.

Benchmark calculations are conceptually straightforward, because a solution that is accurate to within 2% is achievable using line-by-line (LBL) resolution to capture the absorption of gases (Pincus et al., 2015),

along with an accurate scattering solution, such as the doubling-adding (DA) method (Hunt & Grant, 1969). However, the achievement of such a calculation at the spatial resolution of an Atmospheric Global Climate Model is challenged by the extreme computational expense of such an endeavor.

With the computer resources available at GFDL and parallel processing architecture, it has become computationally feasible to develop comprehensive, accurate radiative transfer benchmarks (though not at the resolution needed to account for individual lines; explained in the next section), and to perform them at scales that diagnose the corresponding parameterization errors over a climatologically representative range of atmospheric conditions. While the resources at GFDL enable the computationally ambitious scope of these diagnostics, the open and protocol-based nature of CMIP6 allows for participating models to upload their aerosol and other meteorological pertinent data to the Earth System Grid Federation, and so any institution with access to the requisite resources can perform these calculations.

Such an analysis was demonstrated in Jones et al. (2017) (hereinafter J17), where benchmark computations of aerosol instantaneous radiative effect (IRE) (i.e., fluxes calculated with and without aerosol present) and the resulting parameterization errors are shown for two different climate models, GFDL's AM4 model (Zhao et al., 2018) and the NCAR CESM model version 1.2.2 (Hurrell et al., 2013). The GFDL AM4 radiation code is specific to that model and is documented by Zhao et al. (2018), while CESM contains the widely used RRTMG radiation code (Mlawer et al., 1997). These computations were done utilizing each model's individually prognosed aerosol fields, meteorological thermodynamic conditions, and surface properties for a single, equinoctial day. Comparisons were done for the reflected flux at the TOA, the absorbed flux in the atmosphere and the downward flux at the surface. J17 showed that both models underestimated the absorbed flux in the atmosphere, and overestimated the scattered upward flux at the TOA, with differing signs in the downward flux to the surface. Also the CESM model parameterized result exhibited much less overall error relative to benchmark calculations than that corresponding to the AM4 model for the absorbed and downward fluxes. However, while these parameterized errors and differences between the two models were documented, the source(s) of error were not further identified.

In this paper, we investigate the source(s) of model error in the two models to show how the RFMIP protocols using benchmark calculations with native model error diagnostics enable the specific identification and quantification of their errors in the aerosol radiative forcing. These protocols are adopted for examining the instantaneous errors that arise from model parameterizations. Differences in two principal parametric approximations between the two models' radiation parameterizations are shown in this study to contribute to corresponding differences in their resultant global aerosol IRE biases. These consist of a) the partitioning of the solar spectrum into wider band regions for approximating the spectral dependencies of absorption and scattering, and b) the two stream method used in approximating the scattering process. To separate these two factors, additional calculations not performed in J17 are performed here, where the benchmark-computed gaseous absorption is coupled to the two stream scattering method utilized by each model. For AM4, the delta Eddington (DE) method (Joseph et al., 1976) is used and for the RRTMG code of CESM, the Practical Improved Flux Method (PIFM) (Zdunkowski et al., 1980) is used. Differencing the parameterized results with these additional calculations highlights the error due to spectral partitioning, while differencing the two stream and DA results coupled to the benchmark computed gaseous absorption highlights the error due to two-stream method implementation in each of these models.

In the first part of this study, this approach is used to examine the total spectrally integrated clear-sky error, which is comprised of the clear-clean-sky error (i.e., that due to atmospheric gases and Rayleigh scattering in the absence of clouds and aerosols), referred to hereinafter as just "clean-sky" and the clear-sky aerosol IRE error shown in the J17 paper. This procedure would be expected to be followed by other participating models as part of the RFMIP protocol. In the second part of this study, a finding that arose as a result of questions raised by the output from executing the RFMIP protocol, the errors in the aerosol IRE due to partitioning alone are examined.

Note that since the publication of J17, the global results for the CESM benchmark case shown in Figure 4 panels a, c and e of that paper were discovered to have a slight error, due to an anomaly in how aerosol radiative properties are reported in CESM and subsequently delivered to the benchmark code calculation

routines. This also affected the corresponding error in the parameterized aerosol IRE error. Please see Appendix A for a more detailed discussion of this anomaly and the changes due to its correction.

1.1. Methodology for Reference Calculations

The important characteristics pertaining to the benchmark calculations performed is summarized here. Some of these characteristics were discussed in J17, but some additional detail is provided here. The spatial distribution of gas concentrations, temperatures, pressures, geometric thicknesses, and aerosol optical properties at every grid point and radiative time step for each model are inputted into a parallelized version of the Reference Forward Model (Dudhia, 2016) a LBL code discussed in J17. The code uses line parameters from HITRAN 2012 (Rothman et al., 2013) and the MTCKD 2.5 H_2O, O_2 and N_2 continue to compute the gaseous optical depths at 10^{-4} cm^{-1}.

There is still some uncertainty surrounding the spectroscopy of absorbing gases in the shortwave spectral regions. This is mostly related to the water vapor continuum, but there is also uncertainty in the role of N_2–N_2 and O_2–O_2 collisions. This is important because differences in clean-sky fluxes between our benchmark code and those of other global climate models (GCM's) might represent different choices made about the spectroscopy. For the LBL benchmark code, our choice of the MTCKD 2.5 continuum reflects what is also used within the CESM radiation code. However, the GFDL GCM parameterized radiation code uses a version of the BPS continuum (Paynter et al., 2014) which results in stronger continuum absorption of ~0.7 W/m^2 over the MT CKD 2.5. While not used in either of the codes here, an even larger version of the BPS water vapor continuum is estimated to increase absorption by ~1.3 W/m^2 compared to MT CKD 2.5. This highlights an important distinction that should be made between the clean-sky and aerosol IRE errors presented in this paper. In the case of the clean-sky, we are not attempting to completely mimic the same spectral physics between the benchmark and the parameterized code used in the GCM. So parameterized errors might include different choices in spectroscopy noted above. In the case of the aerosol IRE, since we are using the exact spectral aerosol properties generated by the GCM, there is no such possibility of similar uncertainty being introduced. Hence, the error shown for the aerosol IRE can be considered the true error in radiative transfer for that set of aerosol optical properties. For this reason, although we show the clean-sky error, most of the focus of this paper is on the aerosol IRE.

The computational burden of performing LBL calculations in combination with the DA method for representing the effects of scattering at this resolution of 10^{-4} cm^{-1} is still too great for global applications. However, sampling the gas optical depth every 1 cm^{-1} and running the radiation solver was found to be quite accurate in reproducing the absorbed flux in the atmosphere compared with this reference of 10^{-4} cm^{-1} (estimated to be within 0.2 W/m^2 for global mean shortwave absorption). This sampling interval makes it computationally feasible to perform the global calculations for a spectral range of 1–50,000 cm^{-1} to cover the entire solar spectrum. Those gaseous optical depths, along with aerosol properties (optical depth, single-scattering albedo and asymmetry parameter), and the surface albedo taken from the GCM's native grid are then inputted into the code containing the DA algorithm. For such clear-sky calculations, the flux results determined at 16 streams are found to agree quite closely to that determined at 32 streams. This procedure is repeated for the corresponding clean-sky (aerosol free) case. To obtain a daily mean result, these procedures are repeated for eight individual time steps for AM4 and 24 time steps for CESM, corresponding to the 3 and 1 h radiation time step interval in each model, respectively. These time steps are the standards used for both models in their experimental runs. However, the AM4 model does use a solar interpolation scheme to reduce the bias in the diurnal average fluxes between the 3 h case and what it would be if a 1 h time step was used instead.

Note that in this study the term “reference” is used to describe calculations done with this 1 cm^{-1} sampled solar LBL spectra consideration of the solar spectrum in conjunction with usage of a two-stream scattering technique. The term “benchmark” is used to describe the 1 cm^{-1} sampled LBL spectra in conjunction with a DA calculation done with gaseous absorption included.

Table 1
A Summary of the Global Flux Calculations Done in This Study for the CESM and AM4 Models, and the Attendant Gas Optical Properties, Radiation Solver, and Aerosol Spectral Resolution

Case number	Name	Gas optics	Radiation solver	Aerosol[a]
1	CESM PARAM	14 Band	PIFM	14 band
2	CESM REF	1 cm^{-1}	PIFM	14 band
3	CESM BENCH	1 cm^{-1}	DA	14 band
4	AM4 PARAM	18 Band	DE	18 band
5	AM4 REF	1 cm^{-1}	DE	18 band
6	AM4 BENCH	1 cm^{-1}	DA	18 band
7	AM4 PARAM (CESM input and 2 stream technique)	18 Band	PIFM	18 band
8	AM4 REF (CESM input and 2 stream technique)	1 cm^{-1}	PIFM	18 band

Note. The term "PARAM" refers to a parameterized calculation, "REF" refers to a reference two-stream calculation with the sampled 1 cm^{-1} LBL spectra, and "BENCH" refers to a benchmark DA calculation with the sampled 1 cm^{-1} LBL spectra. Also shown is a summary of the differences taken between cases and the appropriate figures in the text where they are referenced. Total Error for CESM = 1–3 (Figures 1–3 panels a and d). Spectral Error for CESM = 1–2 (Figures 1–3 panels b and e; Figure 9 through 14, and B2, panels a, b, and c). Solver Error For CESM = 2–3 (Figures 1–3 panels c and f). Total Error for AM4 = 4–6 (Figures 4–6 panels a and d). Spectral Error for AM4 = 4–5 (Figures 4–6 panels b and e). Solver Error for AM4 = 5–6 (Figures 4–6 panels c and f). Spectral Error for CESM Aerosol mapped onto AM4 18 band parameterization = 7–8 (Figure 9 through 14, and B2, panels d, e, and f).

[a]Note for the clean-sky calculation that the aerosol is ignored for all cases.

2. Results

A summary of the cases for the various global flux calculations done in this study and the attendant resolution for the specification of gas and aerosol optical properties, along with the type of radiation solver for each model, are presented in Table 1. Also shown for consideration is a summary of the differences taken between the cases that is the basis for the results presented in the figures.

For our results, the clear-sky error is separated into two main components: a) the clean-sky contribution (without aerosols present), and b) the aerosol IRE (clear-sky minus clean-sky) contribution. For each of these components, two contributions to their error are considered; that due to the two stream scattering solution utilized (the solver error) and that due to the spectral partitioning into bands (the spectral error). In Table 1, the global flux calculations done to ascertain these errors are summarized. In order to disentangle the separate effects of these two errors, a reference 1 cm^{-1} sampled LBL calculation is performed with the same two stream approximation (PIFM for CESM and DE for AM4) used in each model's radiation code (i.e., CESM REF and AM4 REF). This calculation is differenced from the parameterization calculation (i.e., CESM PARAM and AM4 PARAM) to estimate the spectral error. The solver error is then estimated as the difference between the reference two stream calculation (i.e., CESM REF and AM4 REF) and the benchmark DA calculation (i.e., CESM BENCH and AM4 BENCH). The total integrated spectral and solver errors are presented in subsections 2.1 and 2.2, and a decomposition of spectral error is the focus of the discussion in subsections 2.3 and 2.4. An examination of the error components for the clean-sky and aerosol IRE in this manner aids in understanding the contribution of gaseous absorption compared to the effect of aerosols, respectively, in the resultant clear-sky error. Note that the solver error can also be estimated by utilizing each model's parameterized result and differencing it from a DA calculation utilizing the same band spectra. However, we did not have the ability to run the DA method with the CESM parameterization. Further the two main benefits by the methodology used here is: a) that the spectral error represents the best accuracy possible for the GCM without improving the solver, while the solver error is the best result possible by using the sampled LBL spectra and b) the two errors sum up to the total. For AM4, there would be a slight increase in the magnitude of the aerosol IRE differences globally than what will be shown here, but the spatial patterns remain similar. Thus, we choose the method to determine the solver error highlighted here.

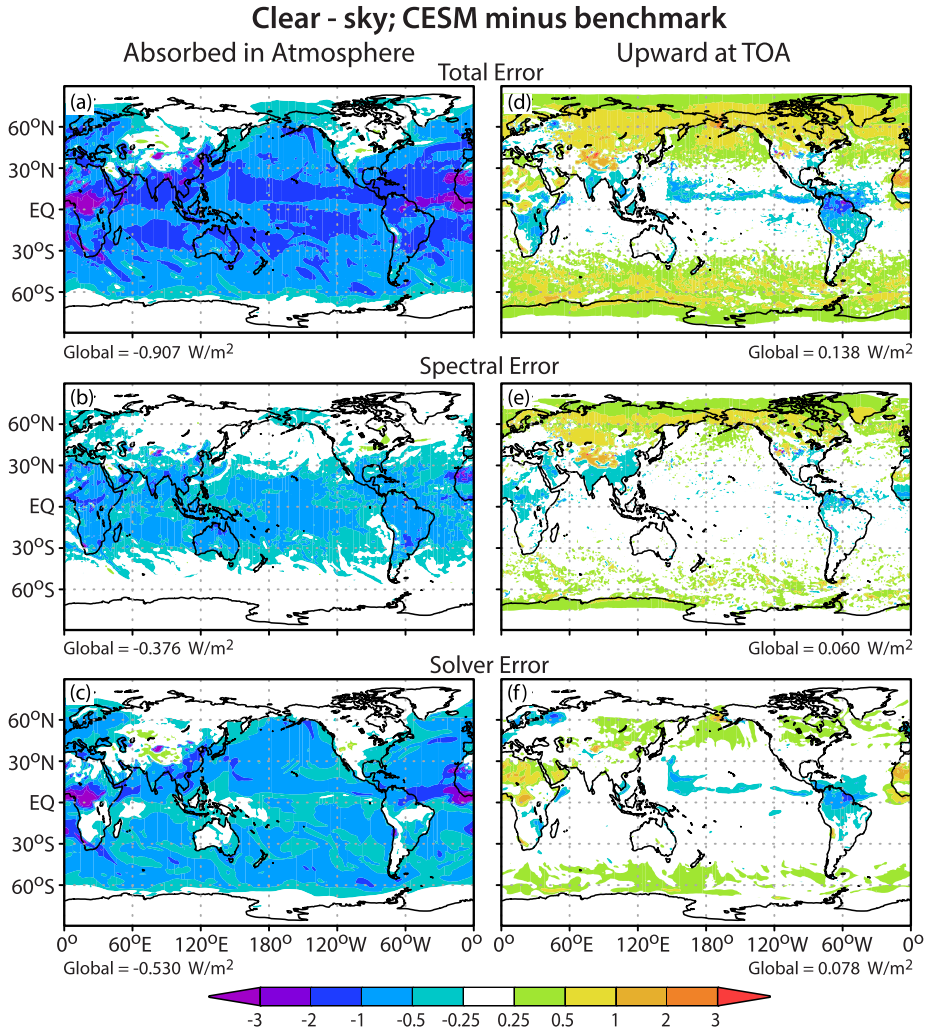

Figure 1. (a) The total error in the clear-sky absorbed flux in the atmosphere from the CESM parameterization as compared to the doubling-adding (DA) benchmark result and (b) as compared to the practical improved flux method (PIFM) reference result (spectral error). (c) The corresponding error in the PIFM reference result compared with the DA benchmark result (solver error). Panels d, e and f correspond to panels a, b and c for the upward flux at the TOA.

2.1. Components of the CESM Spectrally Integrated Parameterized Clear-Sky Error

Figure 1 panels a and d show the CESM parameterized global error in the clear-sky case for the absorbed flux in the atmosphere and the upward flux at the TOA, respectively. The pattern of the errors in the downward flux at the surface is nearly the same as that for the absorbed flux, but of opposite sign and is not shown here. It is seen that both the spectral error and solver error due to PIFM contribute towards the underestimate in the absorbed flux in the atmosphere and mostly an overestimate in the reflected flux at the TOA. The clear-sky error is separated into its clean-sky and aerosol IRE components in Figures 2 and 3, respectively.

As shown in Figure 2, CESM's underestimation of gaseous absorption mainly comes from the spectral error, which is discussed in detail in Appendix B. The spectral error is estimated to be -0.3 W/m^2. As we noted in the methodology section, differences in formulation of the H_2O continuum and other shortwave absorbers means that there is around $\sim$2 W/m^2 uncertainty in the true clean-sky benchmark value. The closeness of the value is most likely because we are using the same MTCKD 2.5 H_2O continuum formulation in our benchmark calculations as used in the CESM GCM. The closeness of values also indicates that the CESM

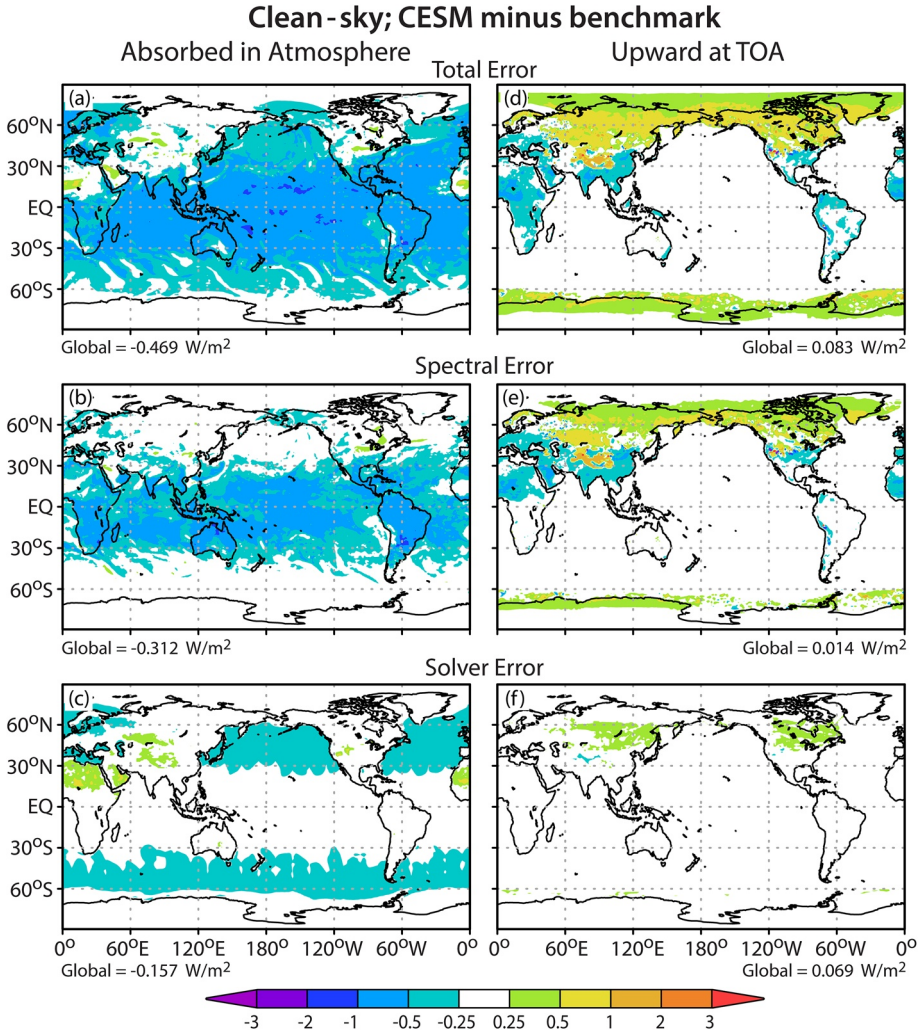

Figure 2. Same as Figure 1 for the clean-sky case.

radiation code is doing a decent job at representing the shortwave absorption by the atmospheric spectra it includes.

The clean-sky solver error is generally less over land than over the ocean, with the exception of land surfaces with higher surface albedo, such as the Sahara Desert. For the upward flux at the TOA, most of the significant errors are confined to land areas, indicating their association with higher surface albedos (note again we use the surface albedos predicted by the GCM) and the resultant increase in gaseous absorption and molecular scattering. As will be shown in subsections 2.3 and 2.4, the sign of this error is related to the wider band widths used by the CESM parameterization in the ultraviolet (UV) region where molecular scattering is more significant. Comparing panels d and e, the spatial pattern of the error is also dominated by the spectral error, but the presence of both underestimates and overestimates in the spectral error result in error cancellation and its global mean contribution (0.014 W/m^2) is less than the solver error due to PIFM (0.069 W/m^2). Hence, this suggests that at the TOA, for the clean-sky, the parameterized code produces results that are similar to the benchmark code.

For the aerosol IRE, Figure 3 panels a and d show that the effect of the aerosols is to further increase the underestimate in absorption and overestimate of reflection at the TOA seen for the clean-sky in Figure 2. Panels b and e show that some of this error that arises is due to the averaging of the overlap of gas and aerosol spectral characteristics. Although the spectral dependence of the aerosol properties is the same in the parameterized and reference calculations, the gas spectral specifications differ. This impact will be

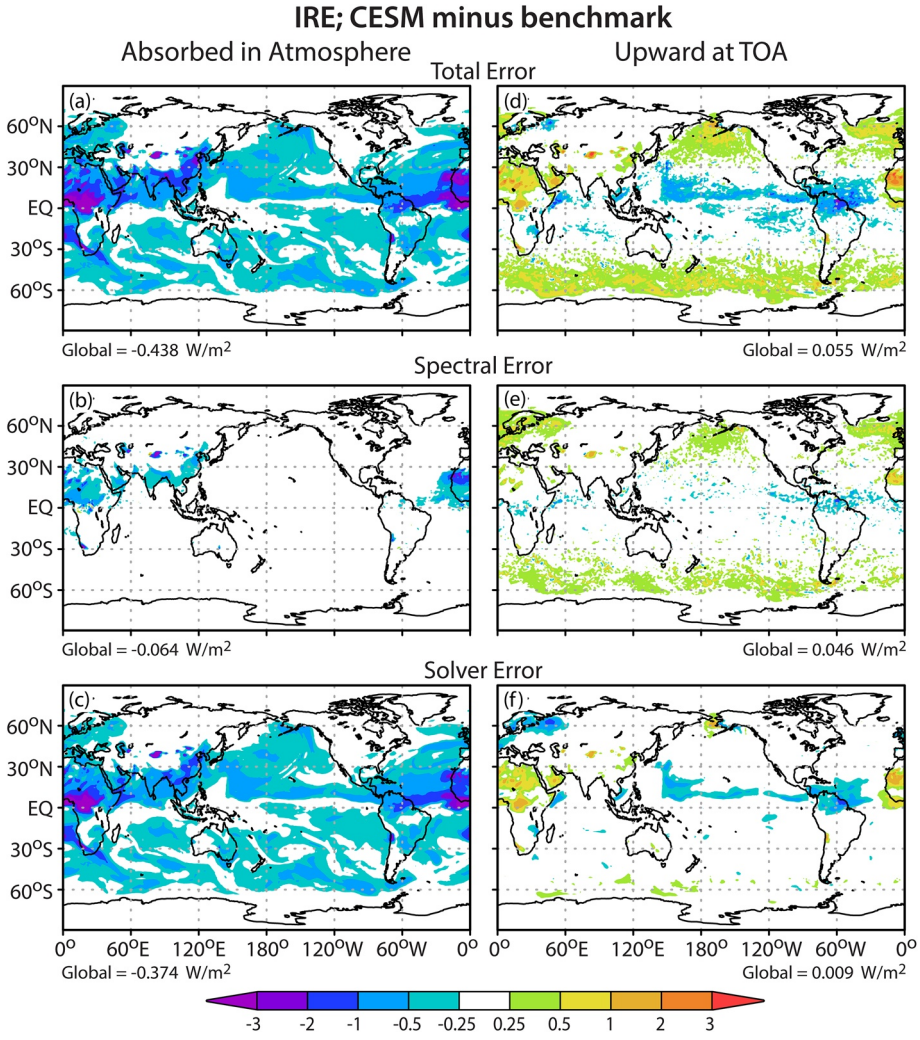

Figure 3. Same as Figure 1 for the aerosol IRE.

examined more closely in subsections 2.3 and 2.4. However, for the absorbed flux in the atmosphere, the majority of the error in the aerosol IRE is due to the use of PIFM two stream solver (Figure 3 panel c), with relatively negligible spectral error (Figure 3 panel b) in most regions except over equatorial Africa, the Saharan-Arabian desert region, and over southeast Asia. Hence, the majority of the underestimation of atmospheric absorption seen in Figure 1 panel a can be attributed in near equal weight to the spectral error underestimating absorption by greenhouse gases (Figure 2 panel b) and the solver error leading to an underestimate of the aerosol absorption (Figure 3 panel c).

For the upward flux at the TOA, the solver error (Figure 3 panel f) is again dominant where the largest biases occur such as the overestimates over the Sahara and equatorial Africa and the underestimates over the equatorial Pacific. Spectral error (Figure 3 panel e), on the other hand, leads to overestimation of reflection over the high latitude oceans.

2.2. Components of the AM4 Spectrally Integrated Parameterized Clear-Sky Error

The results of a similar analysis for the AM4 model are now presented. Again, in order to separate out the contributions of the spectral partitioning of bands, the approximations used for gaseous absorption, and use of the two-stream DE method to the aerosol IRE error, a reference calculation is performed using the DE method (see Table 1 for more details of the exact differences).

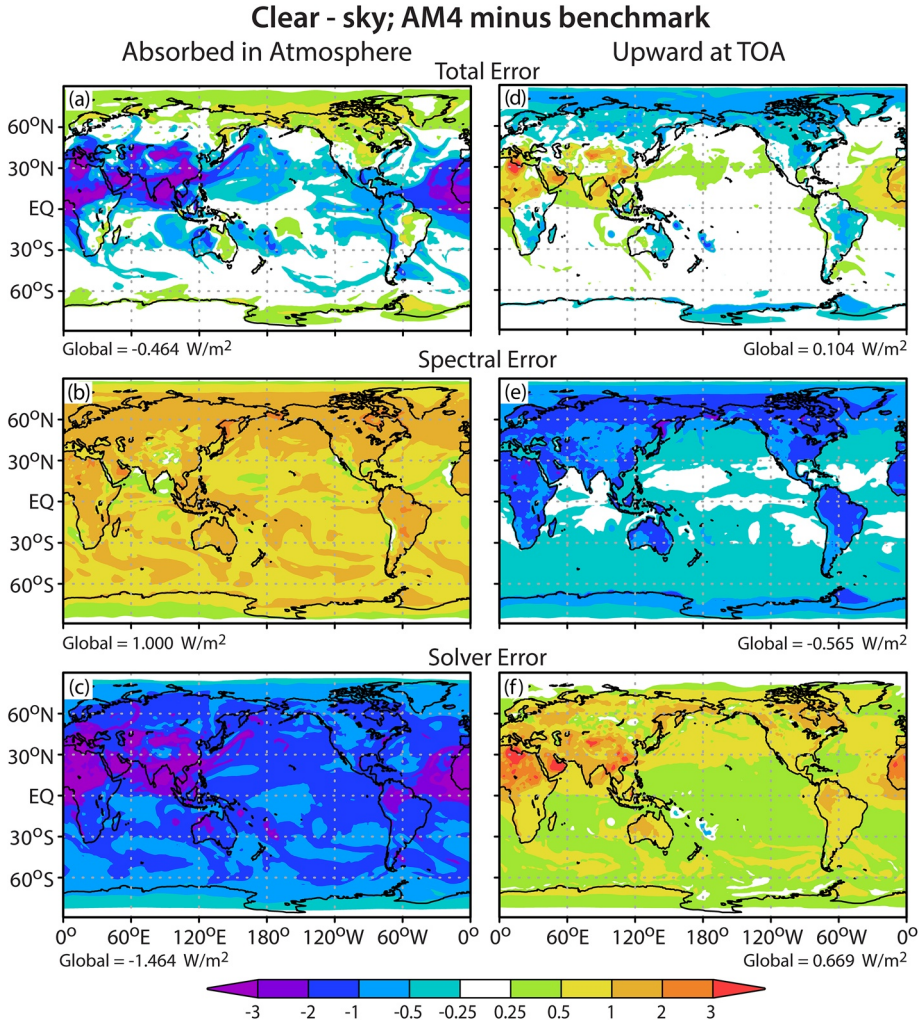

Figure 4. (a) The total error in the clear-sky absorbed flux in the atmosphere from the AM4 parameterization as compared to the doubling-adding (DA) benchmark result and (b) as compared to the Delta Eddington (DE) reference result (spectral error). (c) The corresponding error in the DE reference result compared with the DA benchmark result (solver error). Panels d, e and f correspond to panels a, b and c for the upward flux at the TOA.

The errors in the individual clear-sky and its clean-sky and aerosol IRE components are shown in Figures 4–6 respectively. Comparing Figure 4 with Figure 1, there is a similar underestimate in the clear-sky absorbed flux in the atmosphere and overestimate in the reflected flux at the TOA in AM4, but it is smaller than for CESM. In contrast to the results for the CESM model, however, these arise from compensation of larger magnitude errors among its spectral and solver components. From Figure 5, the clean-sky absorbed flux in the atmosphere is overestimated and the reflected flux at the TOA is underestimated, particularly over land areas; the opposite of the CESM results shown in Figure 2. Comparing panels b and e between Figures 4 and 5, virtually all the clear-sky spectral error is due to the clean-sky component. As noted in the methodology section, the GFDL AM4 GCM uses a stronger H_2O continuum than is used in our benchmark code. This would result in an estimated extra 0.70 W/m^{-2} global mean clean-sky absorption with all other factors being equal. Hence, the fact that GFDL GCM predicts an extra 1.0 W/m^{-2} more absorption than the benchmark code is mainly the result of this difference in parameterization. Similar to the CESM model, the closeness between the benchmark and GCM clean-sky calculations suggests that the GCM parameterization is able to capture the magnitude of clean-sky shortwave absorption within uncertainty of the spectroscopy. This is partially offset by the underestimate in the clean-sky absorption and overestimate in reflection due to the solver error from the DE method.

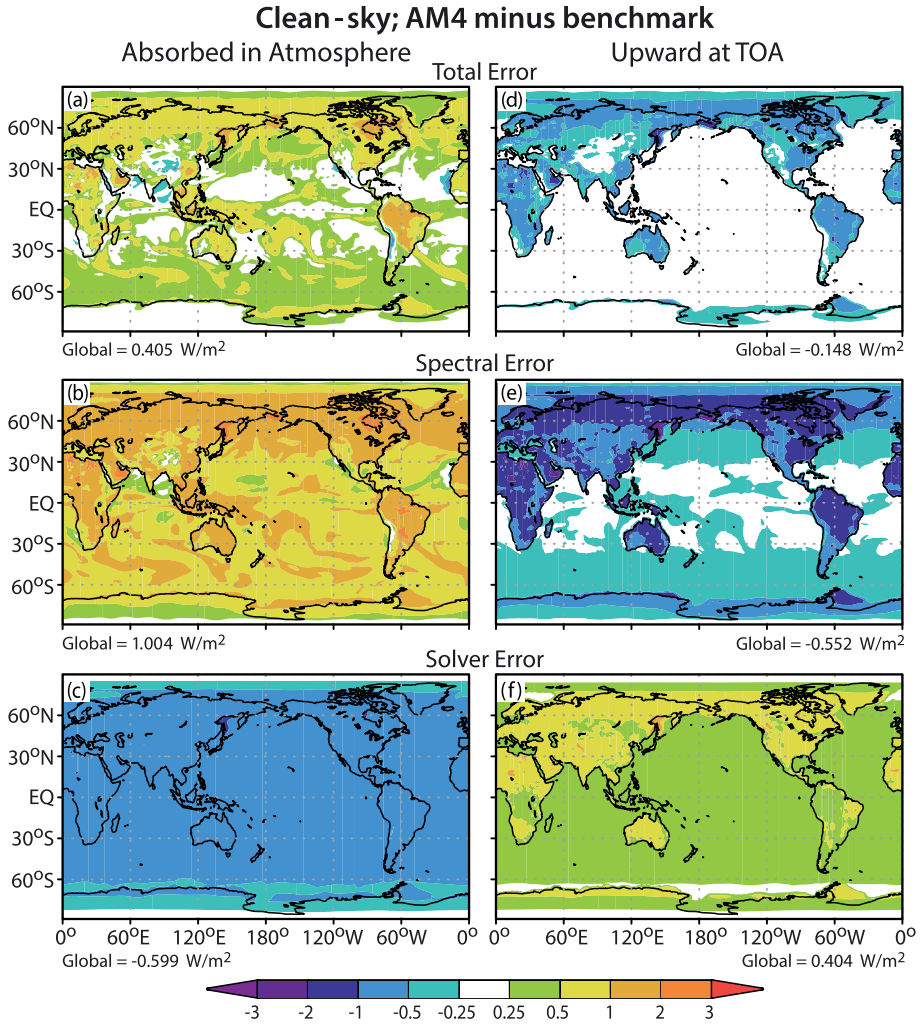

Figure 5. Same as Figure 4 for the clean-sky case.

For Figure 6, the AM4 aerosol IRE error is greater than CESM shown in Figure 3. Also, in contrast to CESM, there is a compensation of errors among the clean-sky and aerosol IRE components of the clear-sky. Comparing with CESM in Figure 3, it is seen that the solver error contributes more substantially towards the total aerosol IRE error, with the spectral error itself being negligible. This is due to the superiority of the 18 band AM4 parameterization in terms of its spectral partitioning for minimizing spectral error, and this finding will be examined more in depth later on in this paper. The DE method results in an underestimate of the aerosol IRE absorbed flux in the atmosphere and an overestimate of the reflected flux at the TOA, similar to the PIFM result seen for CESM aerosol in Figure 3, though the magnitudes are much larger. To understand how both solvers deal with the same aerosol we replace the DE solver with PIFM in AM4 (i.e., we use the radiation solver from CESM on the AM4 aerosol). This leads to a ~30% (0.241 W/m^2) reduction in global mean error in aerosol IRE absorption, and ~40% (0.108 W/m^2) reduction in aerosol IRE TOA reflection. This shows the improvement in the estimation of the effects of scattering on the flux disposition in the atmosphere that is afforded by PIFM. Upon further investigation, most of the difference between the two methods is in the diffusivity factor (defined as the reciprocal of the cosine of the zenith angle) (Räisänen, 2002), used in the determination of the diffuse beam layer reflection and transmission; for the DE method as stipulated in the AM4 parameterization, a value of 5/3 is assumed, while for PIFM the value is 2, resulting in more absorption.

A comparative summary of the global mean results for the two models for the total clear-sky error and all its components shown in Figures 1–6 is presented in Figure 7. For the absorbed flux in the atmosphere,

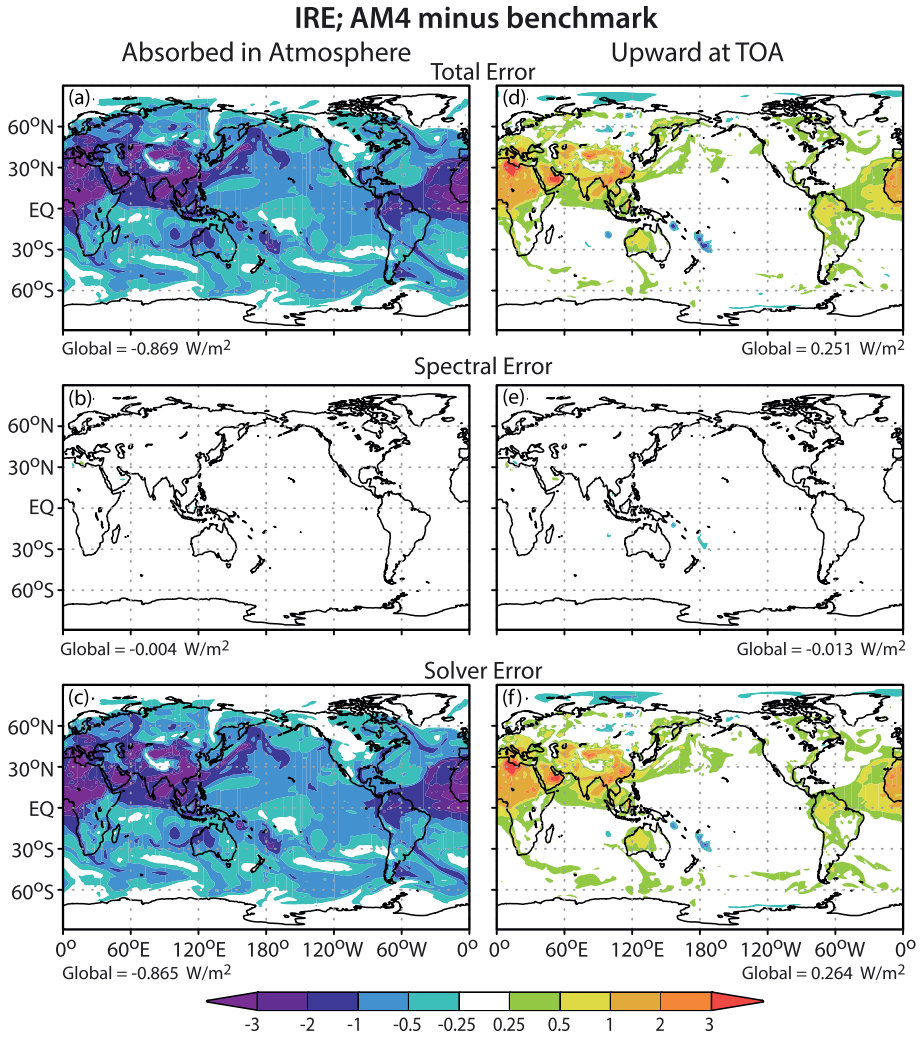

Figure 6. Same as Figure 4 for the aerosol IRE.

for CESM, both the spectral and solver error contribute significantly towards the overall underestimate in the total clear-sky error. Further, the total clean-sky and aerosol IRE components contribute nearly equally towards it. Consistent with the total scattering optical depth, the spectral error is more dominant towards the clean-sky contribution which has only molecular scattering, while the solver error is more dominant for the aerosol IRE contribution. Overall, all the individual components consist of underestimates. Meanwhile, for the AM4 model, the magnitude of the error in the individual components of the total clear-sky error are larger than CESM, but compensation of errors among them result in a smaller total clear-sky value. The underestimate in the solver error plays a greater role, due to both the DE method underestimating the absorption relative to PIFM and the larger aerosol amounts present in AM4. This highlights the greater importance of the accuracy of the scattering method used in a model's shortwave parameterization as the scattering optical depths increase. Similarly, compensating errors are noted among the clean-sky components, with the overestimate of gaseous absorption causing virtually all of the clear-sky spectral error. On the other hand, virtually all of the underestimate in the aerosol IRE error is due to the solver error.

The errors in the TOA reflection show a similar distribution among each component and between the two models to the atmospheric absorption except for the opposite sign. All of CESM's error components are of

Figure 7. The global mean error for the parameterized clear-sky, and its clean-sky and aerosol IRE components for the absorbed flux in the atmosphere with respect to the DA benchmark result for the (a) CESM and (b) AM4 models. In panels c and d, the corresponding results are presented for the upward flux at the TOA. Also shown are the spectral-error and the solver-error components of the total values.

the same sign, while for AM4 there are compensation of errors. A relatively small overestimate in the total clear-sky parameterized error is noted for both models, but the distribution of the error among the various components is again quite different, with all of CESM's error components being similar in sign, while for AM4 there are compensation of errors. For CESM, the solver error is slightly greater than the spectral error contribution towards the total clear-sky error. The errors in the clean-sky and aerosol IRE components are also relatively small. On the other hand, for AM4 the spectral and solver errors are much larger but of opposite signs, with the larger overestimate in the solver error again due to the DE method and larger aerosol amounts. However, the influence of higher surface albedos over land areas as seen in Figure 5 cause the clean-sky solver error to add an additional component to the overall error. The total error in the clean-sky and aerosol IRE are also relatively small, but are a result of compensating errors in the spectral error for the clean-sky, while the aerosol IRE error is due almost exclusively to the solver error. Thus, for the total aerosol IRE results shown in J17, the analysis done in these two subsections summarize the various components contributing towards it.

2.3. Effect of Spectral Partitioning on Differences Between the CESM and AM4 Parameterized Aerosol Instantaneous Radiative Effect Flux Error

Although the two stream techniques highlighted here play a dominant role in the parameterized aerosol IRE errors shown in J17, the contribution from the integrated spectral errors are shown to be notably different between the CESM and AM4 models. In this section we take a closer look at how the spectral error is related to the partitioning of spectral bands in these two shortwave radiation codes. Where the error in the scattering approximation is minimized, spectral error could become a more dominant source of error, and so this exercise highlights the importance of this consideration in the development of shortwave parameterizations.

Spectral partitioning was not an important consideration in the shortwave parameterizations used in the earlier GFDL SKYHI GCM (Fels, et al., 1980), which were based instead on a broadband framework. These included the parameterizations for H_2O vapor (Lacis & Hansen, 1974), and CO_2 (Sasamori et al., 1972). The coefficients in such formulations were later modified based on LBL calculations of gaseous absorption. The errors in the absorbed flux in the atmosphere were reduced from 10% and 20%, respectively, to 1%–2% (Freidenreich & Ramaswamy, 1993; Ramaswamy & Freidenreich, 1992). However, the presence of condensates introduce spectral dependencies in the overlap of scattering and absorption that are difficult to resolve in the broadband context. For instance, the SKYHI model followed the Lacis-Hansen prescription of assigning broadband values of drop reflectivity and transmissivity in handling the overlap of scattering and absorption. The errors incurred by this assumption were investigated in Ramaswamy and Freidenreich (1992), using LBL results as a guide. Although a method was derived to reduce the parameterized flux errors considering the overlap of drops and H_2O only in this broadband context, it became evident that partitioning of the solar spectrum into multiple bands would be needed to improve parameterization of the flux disposition in the atmosphere when both absorption and scattering are present.

In the development of a new multiple-band solar parameterization for use by the present generation of GFDL GCM's (Freidenreich & Ramaswamy, 1999), the availability of LBL results as a guide made it possible to examine the effect of spectral partitioning on the resultant parameterized flux errors in the presence of both absorption and scattering. The selection of the band delimiters took into consideration the spectral variations of H_2O and O_3, the principal gaseous absorbers, and cloud and molecular scattering. At the time of its development, the prescribed cloud single scattering properties used in LBL calculations were taken from the International Comparison of Radiation Codes for Climate Models (ICRCCM) project (Luther et al., 1988). Note that the effect of aerosols were not considered at the time of that parameterization's development. It was found that the magnitude of the parameterized error was not only affected by the number of band regions into which the spectrum was partitioned, but also by the specific placement of the delimiters demarcating the band regions. Since that study, some additional factors influencing the parameterized flux error have been deciphered. In Freidenreich and Ramaswamy (2005), revision of the original solar parameterization included a reduction of the number of bands from 25 to 18, due mainly to combining bands in the near-infrared, made possible from consideration of a lower spectral resolution than ICRCCM for specification of drop optical properties (Slingo, 1989) in LBL computations. Also, a revision in the LBL calculations included the use of a higher resolution (1 cm^{-1}) solar irradiance data-set (Fröhlich & Lean, 2004) for specifying the spectral dependence of incoming solar irradiance variations at the TOA instead of the relatively wide widths originally considered (Labs & Neckel, 1970). In considering the number of bands in the UV region, it was found that the parameterized flux errors were affected by the larger spectral variations in irradiance and its convolution with spectral variations in O_3 absorption, due to this revised data-set. The current formulation used by the AM4 model retains these band delimiters, but with modifications to increase gaseous absorption, including stratospheric absorption by H_2O vapor, improvements to the accounting of the H_2O continuum, as discussed previously in the methodology section, a more refined formulation for the effects of CO_2, and inclusion of absorption by CH_4 and N_2O (Zhao et al., 2018).

Table 2 summarizes the band delimiters chosen for the CESM RRTMG and GFDL AM4 parameterizations. The two models have quite different partitioning, with CESM having 14 spectral bands and most of them in the near infrared region, while AM4's 18 bands are more focussed on the visible + UV region. In Figure 8 panels a and b, the band delimiters from both schemes are superimposed upon the spectral distribution of the absorptive properties of the gaseous constituents, represented by the global mean column optical depth,

Table 2
Band Delimiters of the CESM RRTMG, and GFDL AM4 Shortwave Parameterizations

Band	CESM	GFDL
1	800–2,600 cm^{-1} (near-IR)	0–2,500 cm^{-1}(near-IR)
2	2,600–3,250 cm^{-1} (near-IR)	2,500–4,200 cm^{-1}(near-IR)
3	3,250–4,000 cm^{-1} (near-IR)	4,200–8,200 cm^{-1}(near-IR)
4	4,000–4,650 cm^{-1} (near-IR)	8,200–11,500 cm^{-1}(near-IR)
5	4,650–5,150 cm^{-1} (near-IR)	11,500–14,600cm^{-1}(near-IR)
6	5,150–6,150 cm^{-1} (near-IR)	14,600–16,700 cm^{-1}(vis)
7	6,150–7,700 cm^{-1} (near-IR)	16,700–20,000 cm^{-1}(vis)
8	7,700–8,050 cm^{-1} (near-IR)	20,000–22,300 cm^{-1}(vis)
9	8,050–12,850 cm^{-1} (near-IR)	22,300–24,600 cm^{-1}(vis)
10	12,850–16,000 cm^{-1} (near-IR,vis)	24,600–27,500 cm^{-1}(vis,UV)
11	16,000-22,650 cm^{-1} (vis)	27,500–32,400 cm^{-1}(UV)
12	22,650-29,000 cm^{-1} (vis,UV)	32,400–33,300 cm^{-1}(UV)
13	29,000-38,000 cm^{-1} (UV)	33,300–34,500 cm^{-1}(UV)
14	38,000-50,000 cm^{-1} (UV)	34,500–35,300 cm^{-1}(UV)
15	-	35,300–36,500 cm^{-1} (UV)
16	-	36,500–40,000 cm^{-1} (UV)
17	-	40,000–43,300 cm^{-1} (UV)
18	-	43,300–50,000 cm^{-1} (UV)

for the near-infrared and visible + UV regions, respectively. The high frequency variation of the gas optical depth mostly due to H_2O is seen in panel a, while in panel b its lower frequency variation mostly due to O_3 is depicted, and whose relative contribution increases with wavenumber as the H_2O contribution decreases. The CESM band widths in the near infrared for the first eight spectral bands (up to 8,050 cm^{-1}) are relatively narrow, while individual H_2O bands are combined in specifying the delimiters for bands 9 and 10, resulting in wider spectral widths, and even wider widths are present in the visible and UV regions. The GFDL AM4 band delimiters are comparatively much wider in the near-infrared, combining more individual H_2O bands, and much narrower in the visible and UV, especially for bands 12 to 15 where O_3 absorption is relatively strong and rapidly increasing with wavenumber.

In order to illustrate how the two different spectral partitioning schemes affect the band-by-band result and hence, the total parameterized aerosol IRE flux errors, an analysis is done in which the AM4 18-band radiation scheme uses the CESM climatology (i.e., meteorology and aerosol properties) and the same PIFM solver (Case 7 in Table 1). Further, the 14 band CESM specified aerosol properties are linearly extrapolated onto the 18 band AM4 delimiters. A similar sampled LBL 1 cm^{-1} reference calculation is performed (Case 8 in Table 1). Differencing these two cases (7–8) allows us to obtain an estimate of the spectral error for the AM4 parameterization, to compare more directly with the CESM result (1–2).

The total spectrally integrated results for the 14 band case were presented in Figure 3 panels b and e in discussing the spectral error contribution towards the aerosol IRE differences with the DA benchmark. They are repeated in Figures 9 and 10 (with a higher color bar resolution) to facilitate comparison with the separate contributions from the near-infrared and visible + UV regions and to compare with the corresponding results from the 18 band partitioning. The spectral delimiters comprising the near-infrared and visible + UV regions are defined to be the same as the GFDL band delimiters for bands 1–5 (0−14,600 cm^{-1}) and 6–18 respectively (14,600−50,000 cm^{-1}). For the 14 band results, the equivalent near-infrared and visible + UV contributions are derived, and their band flux values are incorporated into these spectral delimiters. For band 10, which straddles the near-IR and visible + UV regions (Table 2), the flux value in that band is partitioned between the two spectral regions based on the fraction of its band width which falls below or above 14,600 cm^{-1}.

The results for the absorbed flux by aerosol in the atmosphere are presented in Figure 9. The 14 band result shows an underestimate, most notably over equatorial Africa, the Saharan-Arabian desert region, and over southeast Asia. These occur in both spectral regions, but most of the contribution occurs in the visible + UV. However, the global mean values are quite small due to the limited geographical extent of where these occur. For the 18 band result, these negative biases in the visible + UV region over land are significantly reduced, but overestimates are seen to occur over oceanic regions mainly in the near-infrared. Overall, the global mean error is much less for both spectral regions.

In Figure 10 we do a similar comparison for the reflected flux at the TOA. For the 14 band case, the magnitude of the errors are larger over the globe than that seen for the absorbed flux, although the global mean biases remain small due to the cancellation between the low and high latitudes. This is especially true in the visible + UV region with the larger spatial errors. Again, the superiority of the 18 band partitioning is noted by the much smaller magnitude of the errors, which consists of mostly underestimates, mainly over the ocean, with some overestimates over the Sahara. Thus, these results confirm that the smaller aerosol IRE spectral error shown for the AM4 model in Figure 6 and panels b and e, compared with that shown for the CESM model in Figure 3 panels b and e is due mainly the superiority of the 18 band partitioning scheme present in AM4.

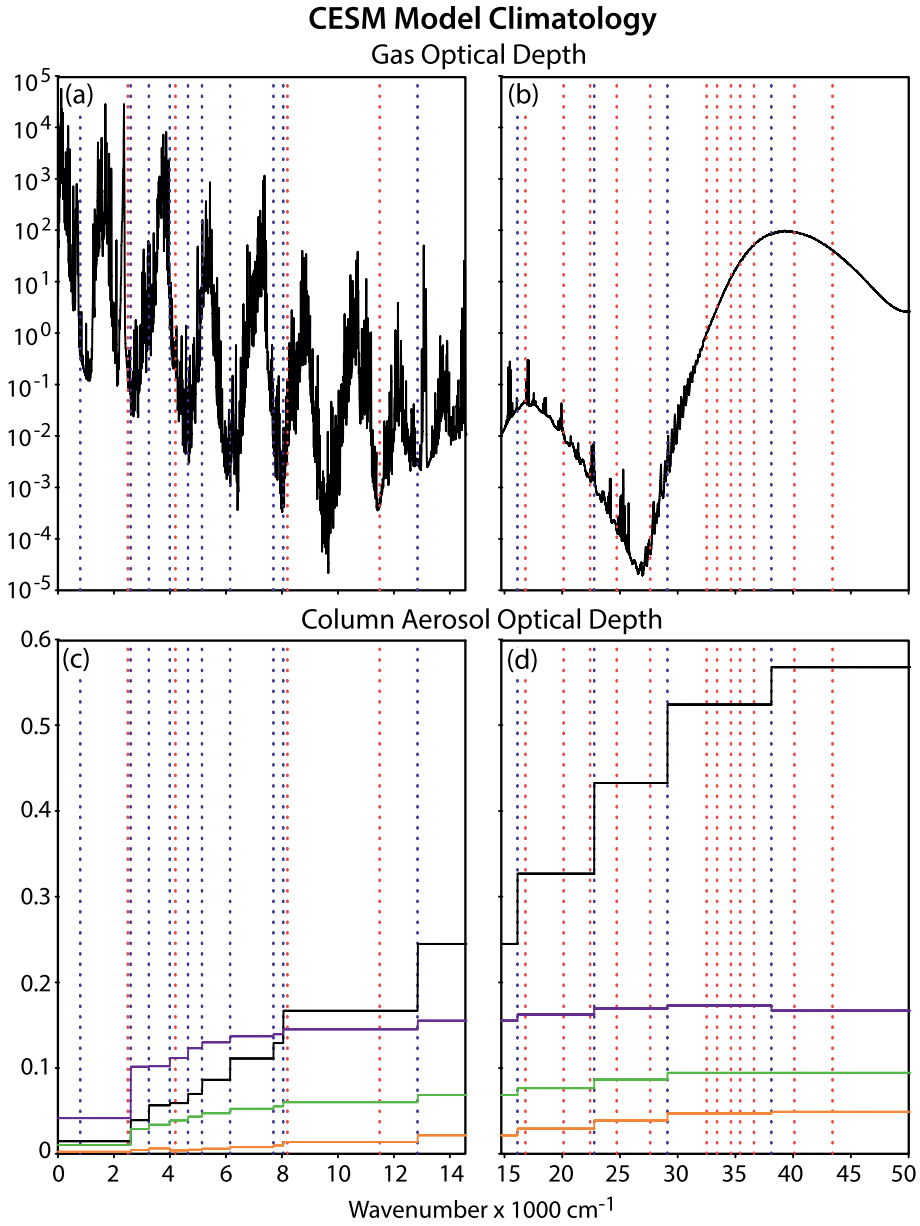

Figure 8. The globally averaged column total gas optical depth for (a) the near-IR and (b) visible + UV. Also shown is the column aerosol column extinction optical depth for four regions (as described in the text) highlighting the dominant effects of the four aerosol types: black carbon (black), dust (green) sulfate (orange) and sea-salt (violet) for (c) the near-IR and (d) visible + UV. The delimiters for the 14 band CESM model are highlighted in blue (dashed) and the 18 band GFDL AM4 model are highlighted in red (dashed).

2.4. Understanding the Impact of Different Aerosol Types on Spectral Error

These findings motivate an exploration of the factors causing the aerosol IRE errors to be generally less for the 18 band partitioning scheme, and particularly in the visible + UV region. This is done by examining the band-by-band dependence of the error differences over select geographical regions. Regions are chosen in which the individual aerosol types (dust, black carbon, sulfate and, sea-salt) are each the dominant contributor as well as having a relative maxima in the prescribed amount. This is done in order to highlight how variations in the aerosol type and the accompanying changes in the column gas amount and surface albedo impact the resultant aerosol IRE error between the two partitioning schemes. From an examination of the prescribed aerosol fields in the CESM model, it is relatively simple to locate regions where dust and sea-salt

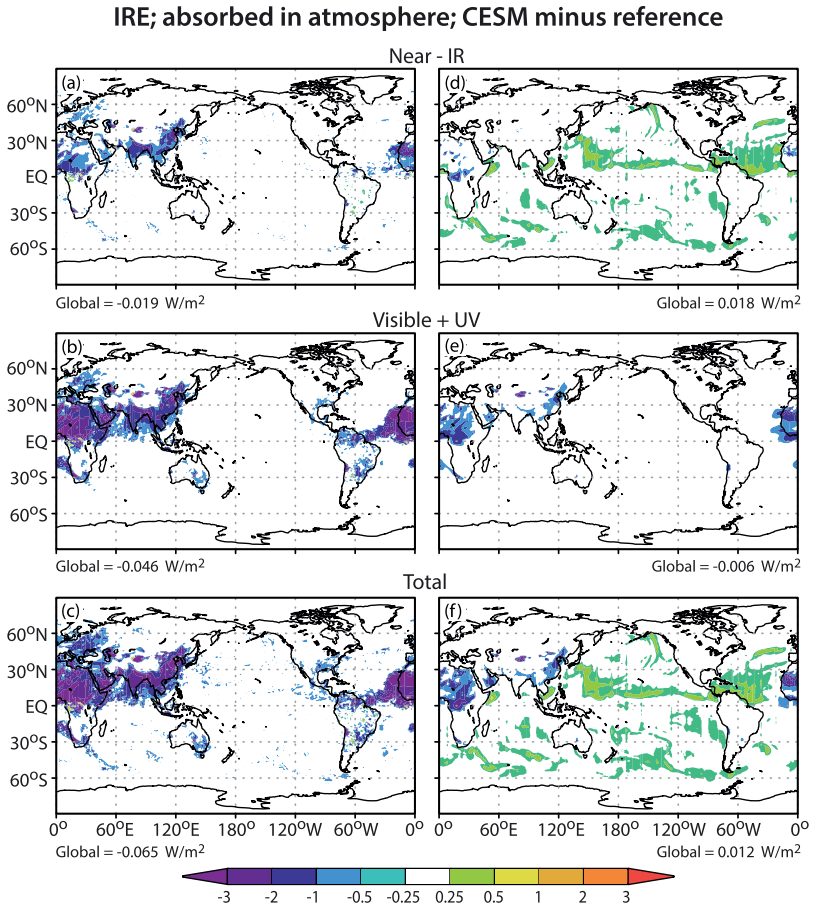

Figure 9. The error in the aerosol IRE absorbed flux in the atmosphere with respect to the corresponding IRE reference result for the 14 band CESM parameterization for (a) near-IR, (b) visible + UV and (c) spectral total, and for the 18 band AM4 parameterization for (d) near-IR, (e) visible + UV and (f) spectral total.

are dominant and independent of each other, but unfortunately, the regions where black carbon and sulfate dominate are generally coincidental. However, they have differing scattering properties, and in the visible part of the spectrum black carbon is strongly absorbing, while sulfate is highly reflecting. Thus, by calculating the combined single-scattering albedo for both using their corresponding extinction coefficients and column burdens, it is possible to differentiate geographical regions which are strongly absorbing and strongly scattering where both constituents simultaneously exist. Thus, the following geographical parameters are selected: 3°S-9°N and 15°E-25°E (black carbon), 10°N-30°N and 0°-60°E (dust), 36°N-42°N and 100°W-68°W (sulfate) and 3°N-23°N and 60°W-30°W (sea-salt).

In Figure 8 panels c and d, the band delimiters are superimposed upon the spectral distribution of these sector's spatial mean aerosol absorption optical depth from the CESM model. The spectral variation is a manifestation of the absorption characteristics for each aerosol type, and its magnitude depends on the aerosol amount in each sector. It is evident that there are differing spectral dependencies in the aerosol absorbing properties among them. Although the aerosol spectral characteristics are fixed over the relatively wide widths constituting the CESM bands, the effective spectral characteristics when they are combined with gaseous absorption are dictated by the gases' higher frequency variations. Thus, the simultaneous presentation of the gaseous and aerosol spectral characteristics in Figure 8 helps to illustrate in a broad sense the overlap that occurs between them in the band regions created by the different partitioning schemes. Note that the spectral dependence of each individual sector mean column gas optical depth is similar to that shown in Figure 8, which displays the global mean value, but their actual magnitude will differ slightly from this depending on the relative mean H_2O and O_3 amounts in that sector. Thus, in analyzing the spectral

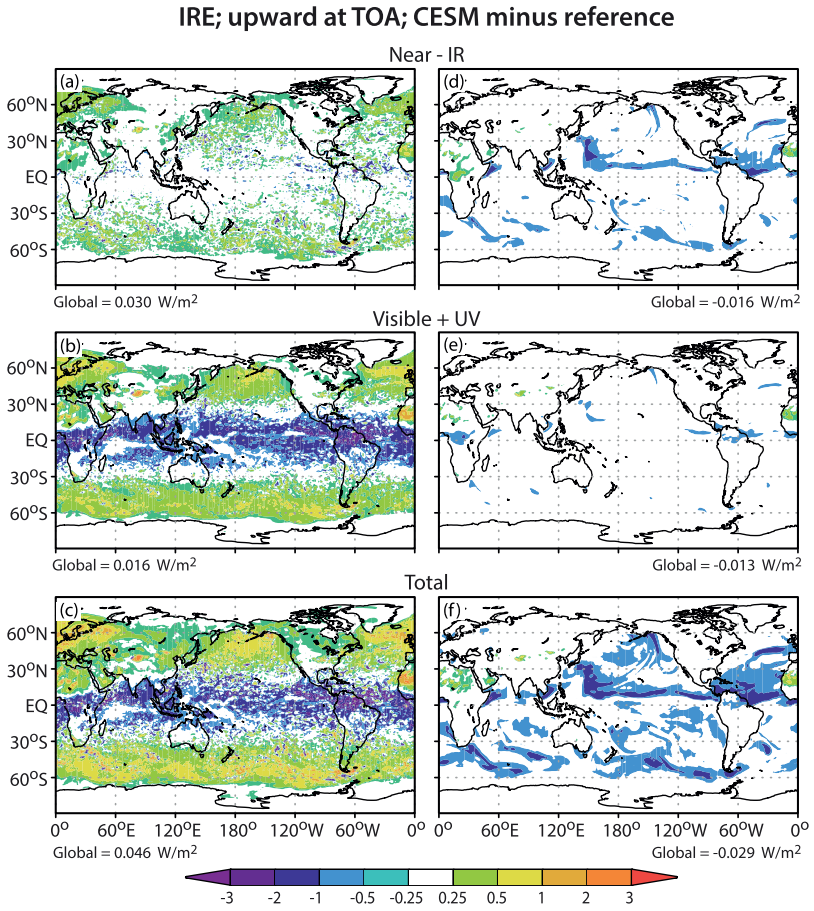

Figure 10. Same as Figure 9 except for upward flux at the TOA.

dependency of the aerosol IRE error in each sector, we also consider in a broad sense the influence of the accompanying changes in the column gas amounts in assessing the degree of overlap between them.

The band-by-band dependence of the aerosol IRE error between the partitioning schemes with 14 and 18 bands is investigated for these four geographical sectors. Besides the clear-sky aerosol IRE, equivalent clean-sky results are presented without gaseous absorption present (i.e., for an atmosphere with aerosol + molecular scattering minus molecular scattering only), which are referred to as the aerosol-only IRE, and require performing additional parameterized and reference calculations. Comparing the two error profiles highlight the contribution of the overlap of aerosol and gas spectral absorption characteristics to the additional error that arises in the clear-sky result. From Figures 8c and 8d, the overall magnitudes of the aerosol absorption optical depths among the four types are dependent on the aerosol amounts in those locations, but the band-by-band variability for any one type depends on the spectral dependence of its extinction properties. Besides aerosol type, two other factors that can influence the aerosol IRE error among the sectors are the gas amount in the column, as stated previously, and the surface albedo. The higher the column gas amount, the larger the contribution toward the parameterized flux error results from the effect of the overlap of gaseous absorption with the aerosol and molecular spectral characteristics. Similarly, the more reflective the surface is, the more that multiple reflections contribute towards the parameterized flux error due to overlap of the scattered beam with the spectral characteristics of the atmosphere above it. The column H_2O and O_3 amounts are presented in Table 3 as the averaged values over each sector as a basis for comparing results among them.

Table 3
The Time and Sector Averaged H_2O and O_3 Column Amounts (kg/m^2) at 10°N-30°N and 0°-60°E (Dust), 3°S-9°N and 15°E-25°E (Black Carbon), 36°N-42°N and 100°W-68°W (Sulfate) and 3°N-23°N and 60°W-30°W (Sea-Salt)

	H_2O	O_3
DUST	1.046	5.396×10^{-4}
BLACK CARBON	3.346	5.135×10^{-4}
SULFATE	1.365	6.979×10^{-4}
SEA-SALT	3.200	5.268×10^{-4}

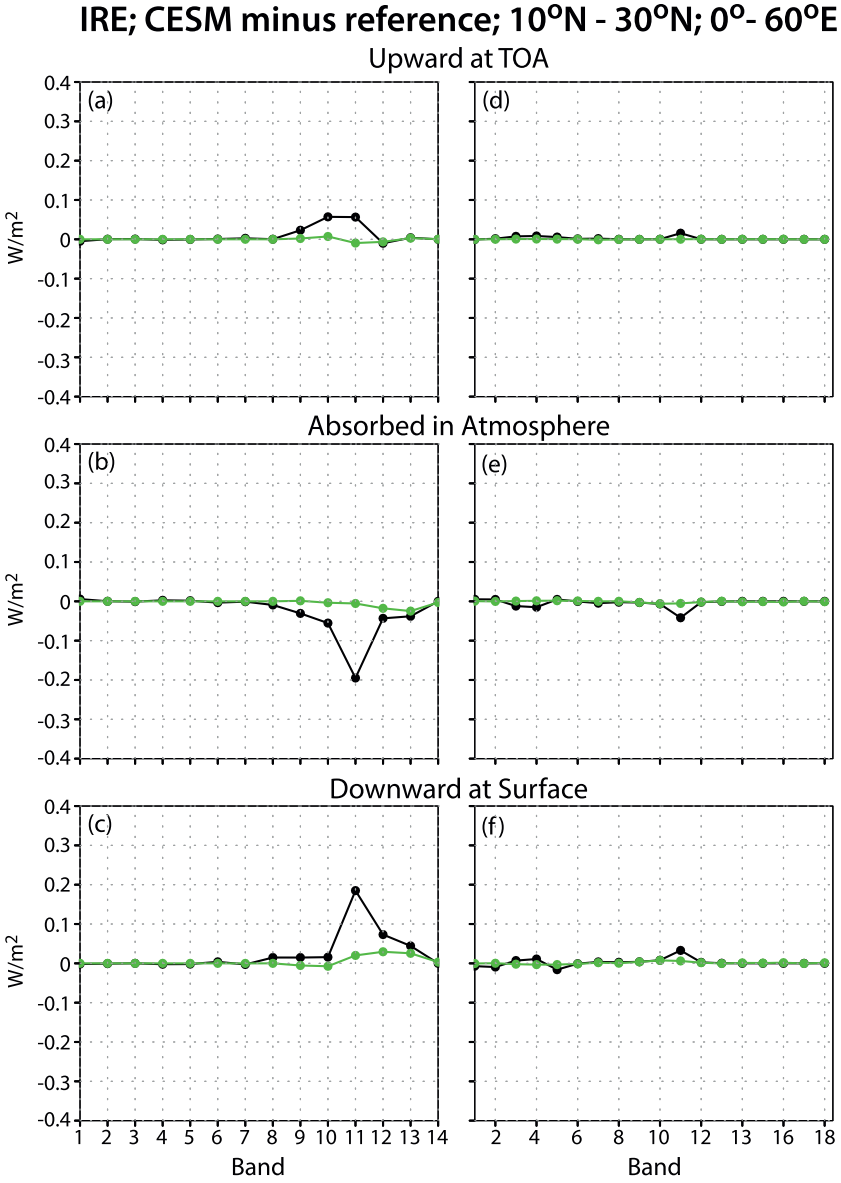

Figure 11. The sector average (10°N-30°N and 0°-60°E) band-by-band instantaneous radiative effect (IRE) error for the 14 band CESM parameterization compared to the corresponding practical improved flux method (PIFM) reference result highlighting the effect of dust aerosol for (a) the reflected flux at the TOA, (b) the absorbed flux in the atmosphere, and (c) the downward flux at the surface. Panels d, e and f show the corresponding errors with respect to the 18 band AM4 parameterization. Results are shown for the aerosol clear-sky IRE (black) and the aerosol-only clean-sky IRE (green).

First we show the results for the sector highlighting the effect of dust aerosol. Figure 11 shows the corresponding band-by-band aerosol IRE errors, while the spectral sum is shown in Table 4. Note that in Appendix B, a comparison between the 14 and 18-band partitioning schemes, for the corresponding errors in the individual flux components comprising these two quantities are examined, to highlight how spectral partitioning affects their errors, and to further reiterate the importance of partitioning in shortwave parametrization development in general. For the 14-band case, without gaseous absorption present (green-line in Figure 11), despite the errors that arise in the individual aerosol + molecular and molecular only cases in Figure B2 (see Appendix B), the aerosol-only IRE error itself is quite small. With the inclusion of

Table 4
The Spectral Sum of the Parameterized Aerosol-Only and Aerosol IRE

a)

	Upward		Absorbed		Downward	
	14	18	14	18	14	18
Aerosol-only IRE	−0.003	0.001	−0.053	−0.015	0.065	0.015
Aerosol IRE	0.129	0.042	−0.366	−0.073	0.342	0.039

b)

	Upward		Absorbed		Downward	
	14	18	14	18	14	18
Aerosol-only IRE	−0.204	−0.042	−0.198	−0.052	0.388	0.090
Aerosol IRE	−0.177	0.013	−0.471	0.222	0.809	0.227

c)

	Upward		Absorbed		Downward	
	14	18	14	18	14	18
Aerosol-only IRE	−0.037	−0.006	−0.006	−0.001	0.047	0.008
Aerosol IRE	0.105	−0.023	−0.043	0.006	−0.056	0.021

d)

	Upward		Absorbed		Downward	
	14	18	14	18	14	18
Aerosol-only IRE	−0.088	−0.008	−0.009	−0.004	0.108	0.011
Aerosol IRE	−0.141	−0.111	−0.092	0.091	0.305	0.044

Note. a) 10–30°N and 0–60°E (Dust), b) 3°S–9°N and 15–25°E (Black Carbon), c) 36–42°N and 100–68°W (Sulfate) and d) 3–23°N and 60–30°W (Sea-Salt) Between the 14 Band and 18 Band Partitioning Schemes.

gaseous absorption, the error in aerosol IRE becomes larger. This is particularly the case for the wider bands 10 through 13 as specified in the RRTMG code. For instance, in band 11, the effect due to overlap of gas and aerosol spectral characteristics in response to the overall weaker gas absorption causes a relatively larger aerosol IRE error. Another factor contributing towards the aerosol IRE error is the high surface albedo over this region which contains the Saharan and Arabian deserts. It is found by an individual grid point analysis that the largest errors occur over points with the highest surface albedo, which can exceed 0.4 in this region. So, despite the relatively low aerosol absorption optical depth (Figure 8) and dry atmosphere (Table 3) over this region, which should limit the error due to overlap of aerosol and gaseous spectral characteristics, the presence of the higher surface albedo increases the parameterized error.

The corresponding results using the 18 band specifications are shown in panels d, e and f. Due to the increased number of bands in the visible + UV (Table 2), the reduction of errors compared with the 14 band case are notable there, from the improved representation of the gaseous absorption overlap with the aerosol spectral characteristics. Only slightly larger errors are seen for the near-infrared bands compared to the 14 band case. Also the aerosol-only IRE error, though small for the 14 band case, is further reduced as well. Thus, Table 4 shows a significant reduction in the total spectral error. This is shown in Figures 9 and 10 more generally over the Saharan-Arabian desert region.

The second sector (3°S-9°N and 15°E-25°E) based over equatorial Africa highlights the effect of black carbon. From Figures 8c and 8d, the aerosol is not only significantly more absorbing than dust in all bands, but that the absorption increases with wavenumber at a much greater rate than for dust. As a result, in Figure 12 and Table 4 the corresponding band-by-band aerosol IRE is larger than for the dust region. This increase of absorption occurs in conjunction with the increased molecular scattering. Hence, we see an enhanced aerosol only IRE (green line) compared to the dust case over bands 11–13. The inclusion of overlap of gases (black line) further enhances the error. As O_3 is the predominant absorber in this spectral region, it suggests that the width of the bands is not fully resolving the spectral convolution with O_3.

The corresponding results using the 18 band specifications are shown in panels d, e and f. Overall, the degree of the reduction of the errors for both the aerosol-only and aerosol IRE are even larger, compared with the 14 band case. However, the greater aerosol absorption also causes larger errors in the near infrared bands due to its overlap with gaseous absorption. This again highlights the importance of proper partitioning to limit the error from overlapping of absorption and scattering. As a result, Figures 9 and 10 show a reduction of underestimates more generally in this region for the 18 band case.

It is of interest to frame these errors in the IRE fluxes in the context of uncertainty in black carbon optical properties. Wang et al. (2018) discuss that uncertainty in the black carbon aging process can lead to a global mean 24% uncertainty in radiative forcing. While we don't perform runs for individual aerosol, for AM4 the region over equatorial Africa between about 5°N and about 20°S contain the highest percentage (70%–80%) for which black carbon contributes toward the total column aerosol optical depth. Using this sector as a basis for analysis in both models, the parameterized IRE sector-mean fractional error in the upward flux corresponding to Figure 3 panel d, is −3.2%, while for AM4, the fractional error corresponding to Figure 6 panel d is +15.5%. Hence, the errors in radiative transfer are comparable to those brought about uncertainty in the optical properties of aerosol.

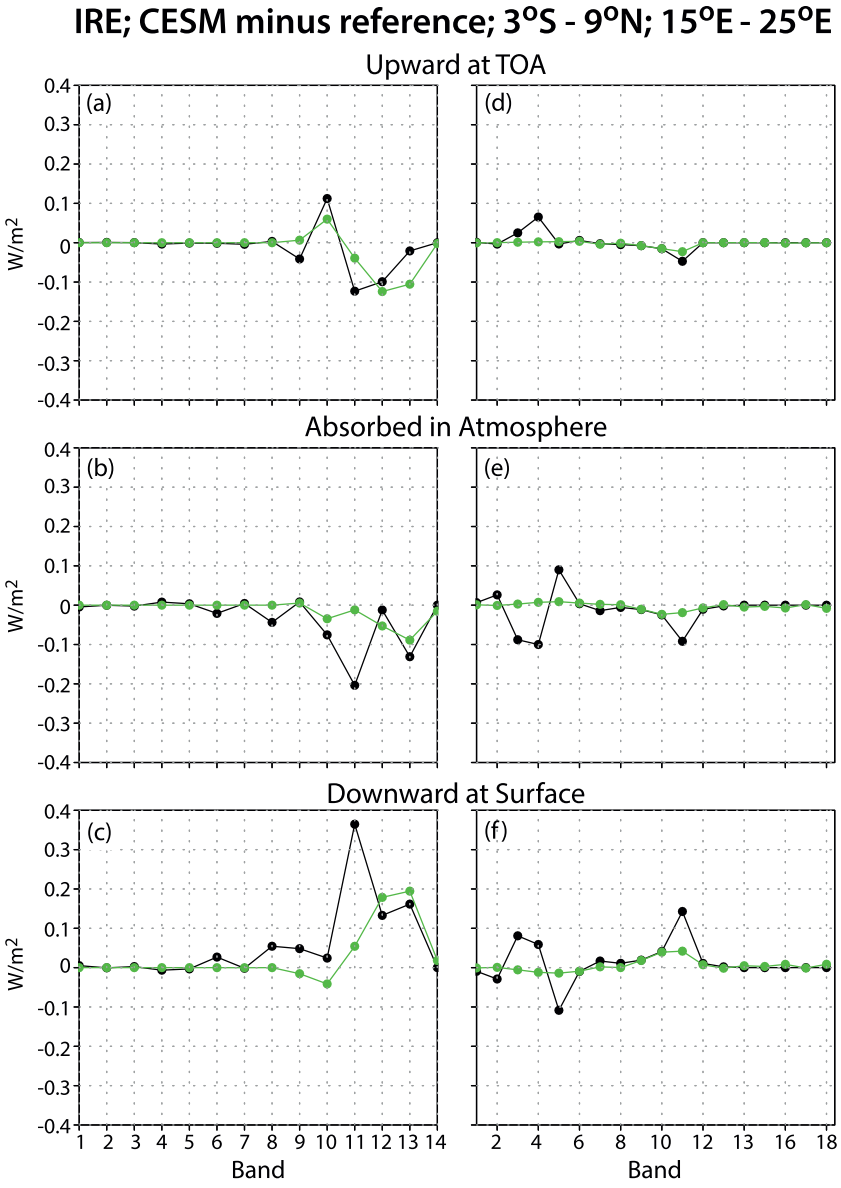

Figure 12. Same as Figure 11 except for the (3°S-9°N and 15°E-25°E) sector average highlighting the effect of black carbon aerosol.

The third sector (36°N-42°N and 100°W-68°W) over eastern North America is dominated by sulfate aerosol. From Figures 8c and 8d, the sulfate absorption is the weakest among the four types, and from Table 3 has relatively smaller H_2O vapor in the column, but has the largest O_3. As a result it has the smallest errors overall for the 14 band case, but these are again further reduced in the 18 band case. The larger O_3 amount, which would increase the error in the UV region due to overlap is not significant due to the weak aerosol absorption.

Finally, the corresponding results for the sector where sea-salt is the dominant aerosol type (3°N-23°N and 60°W-30°W) off the northeastern coast of South America is shown in Figure 14. Compared with sulfate, from Figure 8 the aerosol absorption is stronger and hence the aerosol-only IRE error increases somewhat. Also being a tropical location, from Table 3 the H_2O vapor in the column is also larger, and the overlap effect of the two causes greater errors compared with sulfate for the 14 band case, especially in bands 11 and 12.

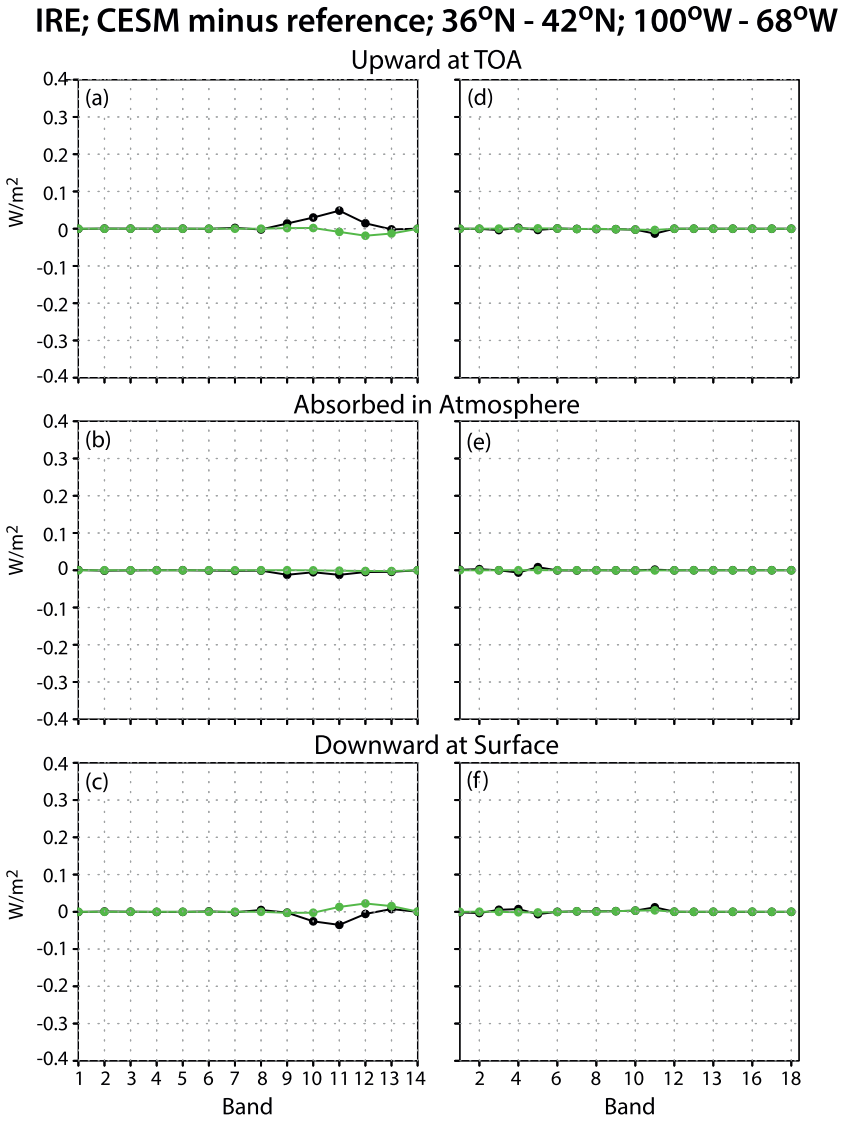

Figure 13. Same as Figure 11 except for the (36°N-42°N and 100°W-68°W) sector average highlighting the effect of sulfate aerosol.

For the 18 band case the aerosol-only IRE is again significantly reduced, but the aerosol and gaseous absorption overlap causes larger errors in the near-infrared bands. Thus from Table 4, comparing the 18 with the 14 band case, the underestimate error is only marginally reduced for the aerosol IRE for the upward flux at the TOA, and the error for the absorbed flux in the atmosphere is similar but is an overestimate. Again these trends are noted in Figures 9 and 10 over this region.

Overall, the spectral distribution and magnitude of the parameterized aerosol IRE flux errors is dependent on the relative strengths of gas and aerosol absorption, which also depends on the type and amount of aerosol present. Improvement in parameterizing the aerosol IRE solely due to spectral partitioning is achieved by better accounting for the gas and aerosol overlap, through the usage of more bands and a judicious choice of the band delimiters, especially in the visible + UV region. Using too wide spectral widths there can lead to greater errors in the aerosol IRE. Overall, the increased resolution of the 18 band specifications in the visible + UV for dust, black carbon and sulfate aerosols leads to a notable reduction of the errors seen where they are predominant in the global results in Figures 9 and 10, but the degraded resolution in the near-infrared leads to only marginal improvements compared to the 14 band case for sea-salt over oceanic regions.

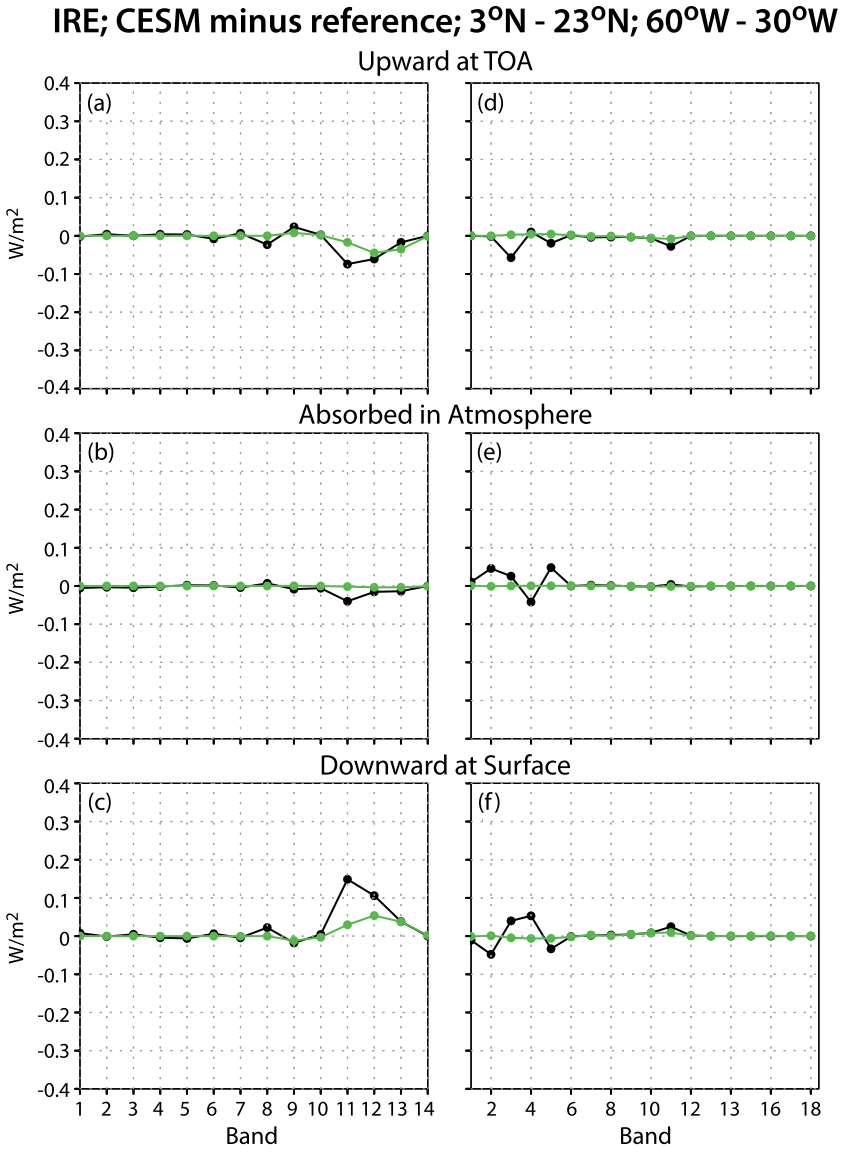

Figure 14. Same as Figure 11 except for the (3°N-23°N and 60°W-30°W) sector average highlighting the effect of sea-salt aerosol.

3. Conclusions

The RFMIP-IRF protocol was designed to develop native model error diagnostics to uncover the dominant sources of error in how participating models calculate radiative forcing. In this work, we have shown how these diagnostics can reveal the sources of model error in instantaneous aerosol forcing, allowing for judicious approaches aimed at reducing it.

From a collection of global scale reference calculations and through the RFMIP diagnostics, in subsections 2.1 and 2.2 the shortwave radiative flux biases have been decomposed into their clean-sky and aerosol IRE components for the AM4 and CESM models. These have been further separated into errors due to their individual band parametrizations and those resulting from the use of their two stream solver techniques. The first source arises from how the spectral dependencies of absorption and scattering are resolved through partitioning of the solar spectrum into bands, and the second source arises from a comparison of a reference calculation performed with each model's two stream technique with its equivalent DA benchmark solution. Overall, the global patterns of the error are quite different between the two models. Considering the global

mean values, for the CESM model, the total clear-sky flux is underestimated for atmospheric absorption and overestimated for TOA reflection, and both the clean-sky and aerosol IRE components of these are of the same sign. Further, both the spectral and solver error components among all three quantities are of the same sign. For AM4, the total clear-sky flux is also underestimated for atmospheric absorption and overestimated for TOA reflection and of smaller magnitude than CESM. However, this is due to compensation of errors among the clean-sky and aerosol IRE components. Further, the spectral and solver errors also exhibit compensation of errors, and both are larger in magnitude than seen in CESM. Of note is the fact that the spectral error in the aerosol IRE in AM4 is quite small.

In subsections 2.3 and 2.4 we have investigated why this spectral error in the aerosol IRE in AM4 is much smaller than CESM. We show that from the spectral decomposition, significant errors in the parameterized aerosol IRE fluxes can arise from the overlap of gas and aerosol absorption if the spectral band widths sampling them are too wide. Further, by examining various geographical sectors where individual aerosol types are dominant, we have shown that the degree and distribution of the errors spectrally depend on the relative strengths of gas absorption and aerosol absorption to each other. For the 18 band partitioning case used in the GFDL AM4 model, the judicious choice of the band delimiters generally minimizes the magnitude of parameterized spectral errors and it is generally smaller than for the 14 bands that CESM uses. While the band delimiter error diagnosis is outside of the scope of the RFMIP-IRF protocol, since additional analysis techniques were used, it shows that spectral partitioning in the presence of scattering and absorption can be an important factor in aerosol radiative forcing error. Thus, in summary, AM4 exhibits a smaller spectral error while CESM exhibits a smaller solver error. However, the magnitude of reduction of the solver error due to CESM's more sophisticated scattering technique is greater than AM4's reduction in the spectral error due to the more judicious choice of band delimiters. Future model improvement for both models should take into account both these factors. To make this point more evident, from investigations currently ongoing, the role of spectral partitioning becomes more important in cloudy atmospheres considering their much thicker scattering optical depths, possibly requiring even further partitioning than the 18 bands used in AM4.

As part of the RFMIP experimental protocol for CMIP6, the work presented here provides a concrete example of how native model error diagnostics provide an exacting level of detail to quantifying uncertainties in radiative forcing due to aerosols. We hope that other models will partake in this endeavor besides the two presented here. Participating models can expect to achieve this level of detail so that they can better understand aerosol forcing and ultimately the Earth system's response to that forcing. Additionally, future availability of global benchmark results for other modeling centers should provide a source for them to seek ways to improve their individual model shortwave parameterizations.

Appendix A: Correction of CESM Benchmark Aerosol IRE Results

Since the publication of J17, the global results for the CESM benchmark case shown in Figure 4 panels a, c and e of that paper were discovered to be incorrect, due to an anomaly in the aerosol radiative properties that were read into the benchmark code. The CESM parameterized results that were provided (Dan Feldman, personal communication) were done correctly, thus the differences with the benchmark result shown in J17 Figure 4 panels b,d and f were also affected. Specifically, the aerosol properties were inadvertently placed one layer below their intended position, so that the actual values in the lowest layer were ignored. So the benchmark code did not see the same profile of aerosol optical depths that the GCM radiation code did. This resulted in an underestimation of the total column aerosol optical depth by the amount that should have been specified there. The increase in the column optical depth resulting from correcting for this is found to be generally less than 0.004 except over the Taklamakan desert where it is up to 0.3. As a result, the most notable change geographically is to further increase the positive aerosol IRE for the absorbed flux in the atmosphere by up to 5 W/m^2 generally over equatorial Africa and the southeastern Asian region, with even greater enhancement over parts of the Sahara and over the Taklamakan desert where large dust loadings are present. For the downward flux at the surface, there is a similar enhancement in the negative aerosol IRE. The most notable change in the upward flux at the TOA is to further decrease the negative aerosol IRE over the Saharan and east Asian regions in general by up to 1 W/m^2. In summary, the resulting global mean aerosol IRE increases by about 0.2 W/m^2 for the upward flux at the TOA and absorbed flux in

the atmosphere, and increases the magnitude of the underestimate (more negative) by about 0.5 W/m^2 for the downward flux at the surface.

The revised CESM parameterized global error in the aerosol IRE for the absorbed flux in the atmosphere and the upward flux at the TOA were shown in Figure 3 panels a and d respectively. The CESM error shows an overall reduction in the overestimate for the upward flux at the TOA and an increase in the underestimate for the absorbed flux in the atmosphere compared with Figure 4, panels b and d in J17. There is also a change in sign to an overall overestimate in the downward flux at the surface (not shown here). In fact, underestimates in the absorbed flux and overestimates in the downward flux occur virtually over the entire globe, and this is more consistent with the GFDL AM4 model aerosol IRE comparisons with its reference results shown in Figure 3 panels d and f of J17.

The motivation for undertaking a closer examination was the appearance of overestimate biases in the absorbed flux aerosol IRE for CESM in Figure 4 panel d in J17 while negative biases were present globally for AM4 in Figure 3 panel d. It was assumed at the time that this was due to the differing aerosol scattering characteristics and climatology present in CESM. In this closer examination, it was surmised that perhaps the broadband specification of the CESM aerosol properties were possibly a contributing factor, so one experiment done was to test the effect of interpolating its band properties spectrally to produce a high resolution specification and then redoing the benchmark calculation. However, this had little impact on the results. It was not until the algorithm characterizing the CESM parameterization was tested with the erroneous input file and compared with the provided flux results that the aforementioned source was identified. In a similar way, future work for RFMIP will involve being provided both the parameterized flux results and the meteorological input files used to generate them; thus it is important that both the provided input climatology is used correctly and the provided parameterized flux results are done correctly, so that determined error between the parameterized and benchmark computations is also correct.

Appendix B: Effect of Spectral Partitioning on Differences Between the CESM and AM4 Individual Component Parameterized Flux Errors

To analyze more closely the factors involved in the differing aerosol IRE errors between the parameterized and reference results due to the two partitioning schemes, how spectral partitioning and band averaging affect the corresponding band-by-band errors in the individual clear-sky and clean-sky flux components comprising it are examined here. The sector 10°N-30°N and 0°-60°E is chosen to illustrate this, to further elucidate the reduction in the aerosol IRE over this region in the 18 band case shown in Figure 11. This exercise helps to further clarify why the choice of the band delimiters is an important consideration in shortwave parameterization development.

Besides the overlap of gaseous absorption and aerosol spectral characteristics considered already, the spectral characteristics of molecular scattering and the solar irradiance at the TOA also influence the parameterized error in the flux calculation. Figure B1 shows the band delimiters superimposed upon the spectral variations of the Rayleigh column optical depth in panels a and b and solar irradiance at the TOA in panels c and d. The Rayleigh optical depth assumes greater importance as wavenumber increases in the visible and particularly in the UV region. In the parameterized computations, the value for a given band is determined at the midpoint wavenumber of the band. The solar irradiance used in the reference computations increases rapidly with wavenumber, then becomes nearly invariant in the near-infrared beyond about 6,000 cm^{-1}, while it decreases in the visible and more so in the UV. In the parameterized computations, the solar irradiance value assigned in a band is its spectral total in the band. Thus, especially in the UV, the choice of the band delimiters and the resulting spectral widths assume greater importance towards errors that arise due to the convolution of spectral variations in the solar irradiance with variations in aerosol and molecular scattering and gaseous absorption. To illustrate to what degree averaging the solar irradiance variation in a band alone leads to flux errors, the PIFM 14 and 18 band reference calculations (Cases 2 and 8, respectively, in Table 1) are repeated using a fixed solar flux per wavenumber value in each band, by dividing the band total solar irradiance by its spectral width, to mimic the use of a single parameterized band value. The resultant band-by-band flux differences that arise compared to the nominal cases are also presented.

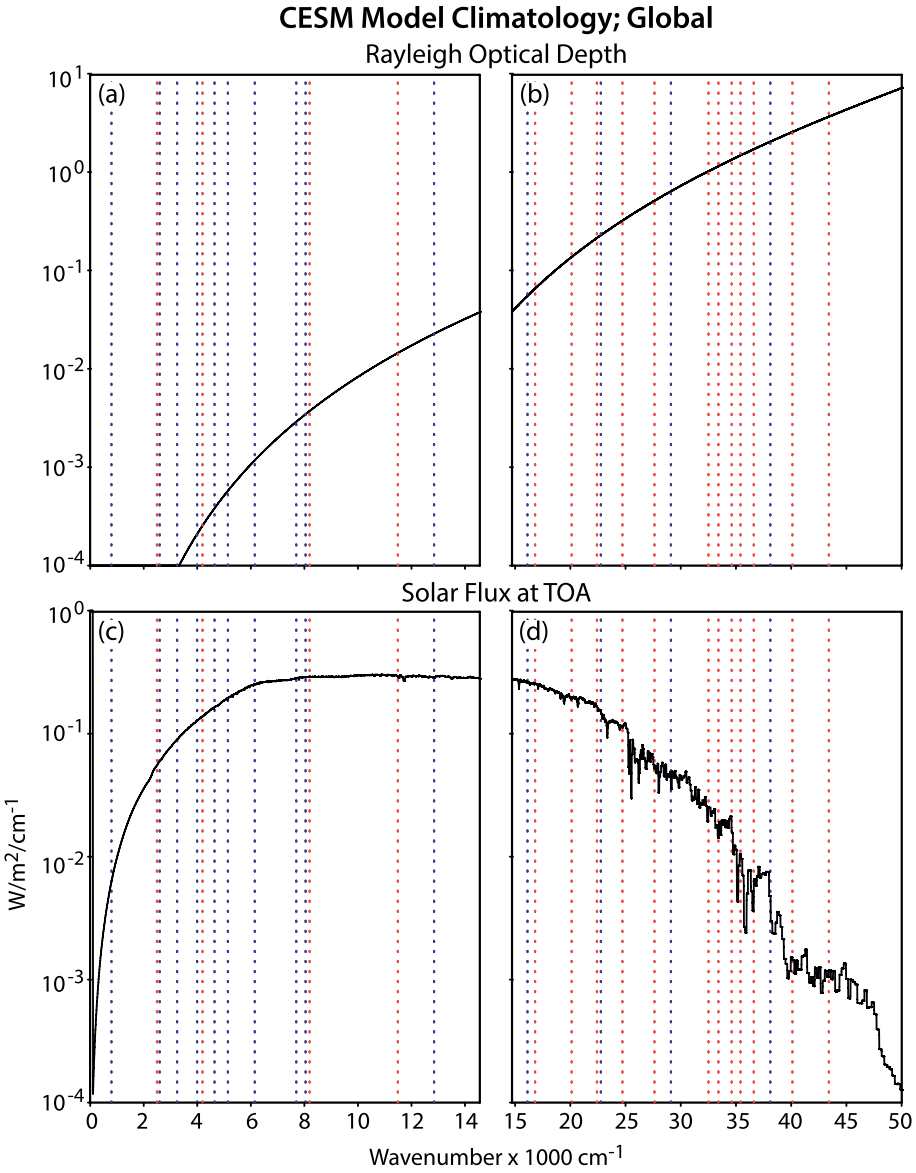

Figure B1. The global mean column total Rayleigh optical depth for (a) the near-IR and (b) visible + UV, and the global mean incident solar flux at the TOA in W/m^2 for (c) the near-IR and (d) the visible + UV. The delimiters for the 14 band CESM model are highlighted in blue (dashed) and the 18 band GFDL AM4 model are highlighted in red (dashed).

Figure B2 shows the band-by-band errors from the 14 band CESM and the 18 band AM4 parameterizations compared to the PIFM 14 and 18 band reference calculations, respectively, used in generating the results in Figure 11. Results are shown for the reflected flux at the TOA, the absorbed flux in the atmosphere, and the downward flux at the surface. For the 14 band case, shown in panels a, b and c, the errors in the first eight bands are dominated by the parameterization of gaseous absorption as evident in the results for the absorbed flux in the atmosphere and the downward flux at the surface for the clean-sky (red) and clear-sky (black) cases. From Figure 8, the strong H_2O absorption and the narrow band widths mitigate the effect of its overlap with aerosol spectral characteristics; thus they are nearly the same resulting in minimal errors in the resultant aerosol IRE (Figure 11). For bands 9 and 10, the errors are still mainly confined to the clean and clear-sky cases, but with somewhat larger difference between them, causing larger errors in the aerosol IRE. This is indicative of a greater effect due to overlap of gas and aerosol spectral characteristics in response to the overall weaker gas absorption. A relatively large underestimate errors in absorption is

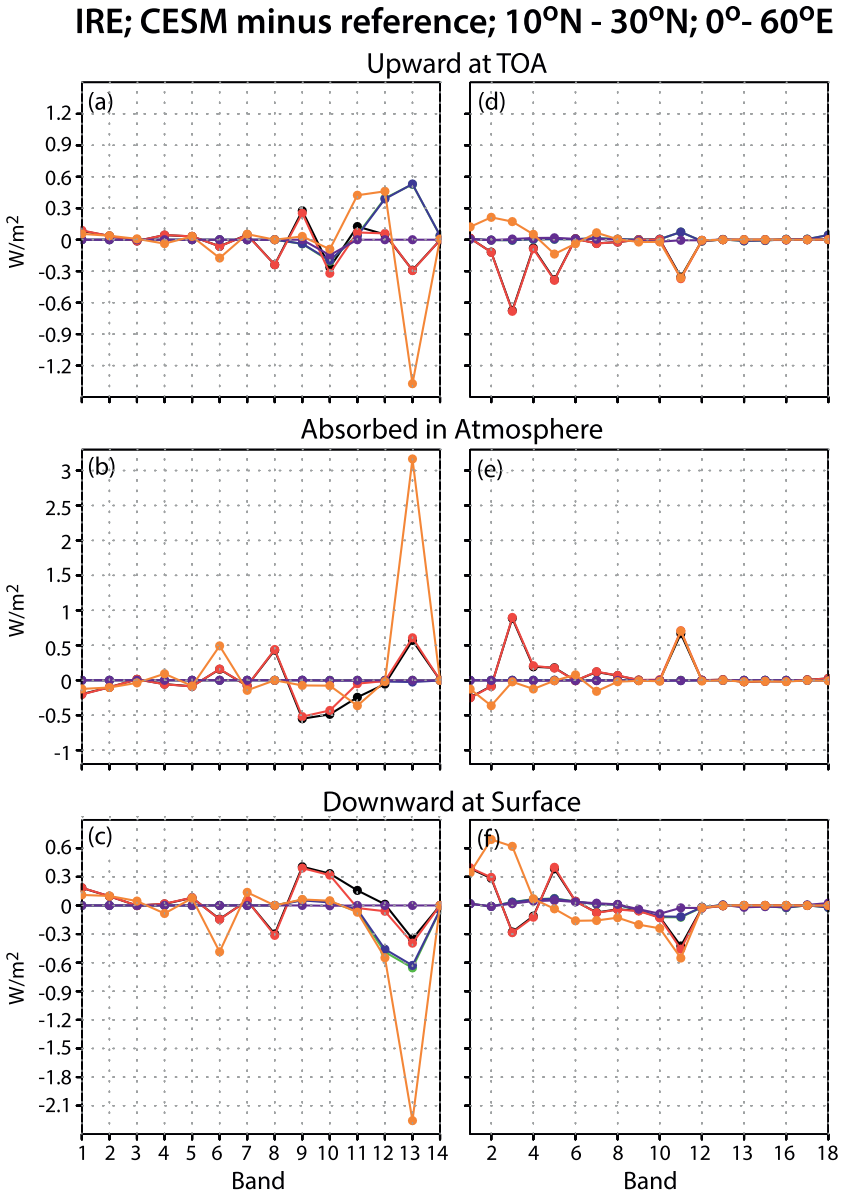

Figure B2. The sector mean (10°N-30°N and 0°-60°E) band-by-band error in for the 14 band CESM parameterization compared to the corresponding practical improved flux method reference result highlighting the effect of dust aerosol for (a) the reflected flux at the TOA, (b) the absorbed flux in the atmosphere, and (c) the downward flux at the surface. Panels d, e and f show the corresponding errors with respect to the 18 band AM4 parameterization. Results are shown for the clear-sky (black), clean-sky (red), aerosol + molecular (blue), and molecular only (green). Also shown (orange) is the band-by-band differences between using a fixed solar irradiance per wavenumber at the TOA in the PIFM reference calculation and the nominal case.

noted in bands 9 and 10, and this may explain much of the underestimate globally in the spectral error seen in panel b of Figures 1 and 2. For bands 11 to 13, the wider spectral widths and increased overlap among the various constituents play more of a part in the resulting errors that occur. For bands 12 and 13, where gaseous absorption is relatively weaker, the larger rate of increase in the Rayleigh optical depth (Figure B1 panel b) influence the errors for the molecular effect only (green), causing an overestimate in the reflected flux at the TOA and underestimate at the surface for the molecular effect only. These are mitigated slightly by the presence of aerosols (blue). From the results shown in Figure 1e for the clean-sky case, the differing sign of the errors globally over land areas with higher surface albedo are dictated by the overlap of gaseous absorption and aerosol and molecular scattering with the wide band widths in bands 12 and 13. The ten-

dency for overestimates occur in regions of weaker gaseous absorption, where the error due to averaging the overlap of aerosol and molecular scattering is more dominant, whereas underestimates occur in the regions of stronger gaseous absorption where errors due to averaging the overlap of its effect over a higher reflecting surface is more dominant, as noted in this sector. Also, the averaging of the solar irradiance variations come into play, as the relatively large irradiances start to decrease with wavenumber (Figure B1 panel d). The largest bias due to averaging the solar irradiance occurs in band 13, where it varies by about an order of magnitude. This contributes towards an overestimate in the absorbed flux in the atmosphere, accompanied by underestimates in the reflected flux at the TOA and downward flux at the surface. The sign of the errors in the clean and clear sky fluxes are related to this due to the convolution of the gas absorption due to O_3 increasing by several orders of magnitude over the width of the band (Figure 8 panel b) along with the corresponding decrease in the solar irradiance (Figure B1 panel d). This highlights how averaging of solar irradiance variations contributes toward parameterized flux errors and requires that the spectral widths chosen aren't too wide in this spectral region. For band 14, despite the wide spectral width of the band, the relatively small solar irradiance limits the parameterized flux errors.

The corresponding results for the 18 band partitioning scheme are shown in panels d, e and f in Figure B2. For the clean-sky and clear-sky cases, the largest overestimate error due to gaseous absorption occurs in band 3, which is relatively wide (Table 2) and encompasses several H_2O absorption bands (Figure 8 panel a). Also, band 11 is noted for a relatively larger overestimate error in absorption there. This is due to a rapidly increasing O_3 optical depth with wavenumber (Figure 8 panel b) along with a rapidly decreasing solar irradiance (Figure B1 panel d), much like what occurred for band 13 for the CESM delimiters. These are the two principal factors contributing to the corresponding overestimate globally seen for the AM4 model in panel b for Figures 4 and 5. However, the effect of averaging the aerosol spectral characteristics in the near-infrared region where they are more weakly absorbing along with the increased number of bands in the visible + UV region where their absorption strengths are more significantly varying with wavenumber (Figure 8 panel d) minimizes the parameterization errors due to the overlap of gaseous absorption and aerosol spectral characteristics contributing to a small aerosol IRE error (Figure 11). Also, the parameterization flux errors due to the effect of averaging the Rayleigh optical depth and the solar irradiance at the TOA are minimized due to the narrower spectral widths in the visible + UV compared to the 14 band case. This again reiterates the importance of carefully selecting the band delimiters in this region.

Acknowledgments
The author would like to acknowledge the many helpful suggestions from the various coauthors as well as those from two internal reviewers. A. L. Jones is funded by DOE award DE-SC0012313, "Collaborative Research: Centralized activities in support of the Radiative Forcing Model Intercomparison Project" and Princeton University's Cooperative Institute for Climate Studies (CICS). The material from co-authors Feldman and Collins is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research and Regional and Global Model Analysis under contract number DE-AC02-05CH11231. This report was prepared by Pu Lin under award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

Data Availability Statement

The data used in this paper along with readme files are permanently hosted on the GFDL data portal (https://data1.gfdl.noaa.gov/) and can be accessed at ftp://data1.gfdl.noaa.gov/users/David.Paynter/2019JD032323/.

References

Dudhia, A. (2016). The Reference Forward Model (RFM). *Journal of Quantitative Spectroscopy and Radiative Transfer*, *186*, 243–253. https://doi.org/10.1016/j.jqsrt.2016.06.018

Fels, S. B., Mahlman, J. D., Schwarzkopf, M. D., & Sinclair, R. W. (1980). Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamical response. *Journal of the Atmospheric Sciences*, *37*, 2265–2297. https://doi.org/10.1175/1520-0469(1980)037<2265:sstpio>2.0.co;2

Freidenreich, S. M., & Ramaswamy, V. (1993). Solar radiation absorption by CO_2, overlap with H_2O, and a parameterization for general circulation models. *Journal of Geophysical Research*, *98*(D4), 7255–7264. https://doi.org/10.1029/92jd02887

Freidenreich, S. M., & Ramaswamy, V. (1999). A new multiple-band solar radiative parameterization for general circulation models. *Journal of Geophysical Research*, *104*(D24), 31389–31409. https://doi.org/10.1029/1999jd900456

Freidenreich, S. M., & Ramaswamy, V. (2005). Refinement of the geophysical fluid dynamics laboratory solar benchmark computations and an improved parameterization for climate models. *Journal of Geophysical Research*, *110*, D17105. https://doi.org/10.1029/2004JD005471

Fröhlich, C., & Lean, J., (2004). Solar radiative output and its variability: Evidence and mechanisms. *Astronomy and Astrophysics Review*, *12*, 273–320. https://doi.org/10.1007/s00159-004-0024-1

Hunt, G. E., & Grant, I. P. (1969). Discrete space theory of radiative transfer and its application to problems in planetary atmospheres. *Journal of the Atmospheric Sciences*, *26*, 963–972. https://doi.org/10.1175/1520-0469(1969)026<0963:dstort>2.0.co;2

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The community Earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1

Jones, A. L., Feldman, D. R., Freidenreich, S., Paynter, D., Ramaswamy, V., Collins, W. D., & Pincus, R. (2017). A new paradigm for diagnosing contributions to model aerosol forcing error. *Geophysical Research Letters*, *44*(23). 12004–12012. https://doi.org/10.1002/2017GL075933

Joseph, J. H., Wiscombe, W. J., & Weinman, J. A. (1976). The delta-Eddington approximation for radiative flux transfer. *Journal of the Atmospheric Sciences*, *33*, 2452–2459. https://doi.org/10.1175/1520-0469(1976)033<2452:tdeafr>2.0.co;2

Labs, D., & Neckel, H. (1970). Transformation of the absolute solar radiation data into the International practical temperature scale of 1968? *Solar Physics*, *15*, 79–87. https://doi.org/10.1007/bf00149474

Lacis, A. A., & Hansen, J. (1974). A parameterization for the absorption of solar radiation in the Earth's atmosphere. *Journal of the Atmospheric Sciences*, *31*, 118–133. https://doi.org/10.1175/1520-0469(1974)031<0118:apftao>2.0.co;2

Luther, F. M., Ellingson, R. G., Fouquart, Y., Fels, S., Scott, N. A., & Wiscombe, W. J. (1988). Intercomparison of radiation codes in climate models (ICRCCM): Longwave clear sky results: A work-shop summary. *Bulletin of the American Meteorological Society*, *69*, 46–48. https://doi.org/10.1175/1520-0477-69.1.40

Mlawer, E. J., Brown, S. J. P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, *102*, 16663–16682. https://doi.org/10.1029/97jd00237

Oreopoulos, L., & Mlawer, E. (2010). MODELING: The continual intercomparison of radiation codes (CIRC). *Bulletin of the American Meteorological Society*, *91*, 305–310. https://doi.org/10.1175/2009BAMS2732.1

Paynter, D., Ramaswamy, V., & September (2014). Investigating the impact of the shortwave water vapor continuum upon climate simulations using GFDL global models. *Journal of Geophysical Research: Atmospheres*, *119*(8), 10720–10737. https://doi.org/10.1002/2014JD021881

Pincus, R., Forster, P. M., & Stevens, B. (2016). The radiative forcing model intercomparison project (RFMIP): Experimental protocol for CMIP6. *Geoscientific Model Development*, *9*, 3447–3460. https://doi.org/10.5194/gmd-9-3447-2016

Pincus, R., Mlawer, E. J., Oreopoulos, L., Ackerman, A. S., Baek, S., Brath, M., et al. (2015). Radiative flux and forcing parameterization error in aerosol-free clear skies. *Geophysical Research Letters*, *42*, 5485–5492. https://doi.org/10.1002/2015GL064291

Räisänen, P. (2002). Two-stream approximations revisited: A new improvement and tests with GCM data. *Quarterly Journal of the Royal Meteorological Society*, *128*, 2397–2416. https://doi.org/10.1256/qj.01.161

Ramaswamy, V., & Freidenreich, S. M. (1992). A study of broadband parameterizations of the solar radiative interactions with water vapor and water drops. *Journal of Geophysical Research*, *97*(D11), 11487–11512. https://doi.org/10.1029/92jd00932

Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P. F., et al. (2013). The HITRAN2012 molecular spectroscopic database. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *130*, 4–50.

Sasamori, T., London, J., & Hoyt, D. V. (1972). Radiation budget of the Southern Hemisphere. *Meteorological Monographs*, *13*, 9–23. https://doi.org/10.1007/978-1-935704-33-1_2

Slingo, A. (1989). A GCM parameterization of the shortwave radiative properties of water clouds, *Journal of the Atmospheric Sciences*, *46*, 1419–1427. https://doi.org/10.1175/1520-0469(1989)046<1419:AGPFTS>2.0.CO;2

Soden, B. J., Collins, W. D., & Feldman, D. R. (2018). Reducing uncertainties in climate models. *Science*, *361*(6400), 326–327. https://doi.org/10.1126/science.aau1864

Stouffer, R. J., Eyring, V., Meehl, G. A., Bony, S., Senior, C., Stevens, B., & Taylor, K. E. (2017). CMIP5 Scientific gaps and recommendations for CMIP6. *Bulletin of the American Meteorological Society*, *98*(1), 95–105. https://doi.org/10.1175/BAMS-D-15-00013.1

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Wang, Y., Ma, P. L., Peng, J., Zhang, R., Jiang, J. H., Easter, R. C., & Yung, Y. L. (2018). Constraining aging processes of black carbon in the community atmosphere model using environmental chamber measurements. *Journal of Advances in Modeling Earth Systems*, *10*(10), 2514–2526. https://doi.org/10.1029/2018ms001387

Zdunkowski, W. G., Welch, R. M., & Korb, G. (1980). An investigation of the structure of typical two-stream methods for the calculation of solar fluxes and heating rates in clouds. *Beiträge zur Physik der Atmosphäre*, *53*, 147–166, 1980.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018). The Global atmosphere and land model AM4.0/LM4.0 - Part II: Model description, sensitivity studies, and tuning strategies. *Journal of Advances in Modeling Earth Systems*, *10*(3). https://doi.org/10.1002/2017MS001209