Understanding Top-of-Atmosphere Flux Bias in the AeroCom Phase III Models: A Clear-Sky Perspective

Journal Article

Author(s):
Su, Wenying; Liang, Lusheng; Myhre, Gunnar; Thorsen, Tyler J.; Loeb, Norman G.; Schuster, Gregory L.; Ginoux, Paul; Paulot, Fabien; Neubauer, David; Checa-Garcia, Ramiro; Matsui, Hitoshi; Tsigeridis, Kostas; Skeie, Ragnhild B.; Takemura, Toshihiko; Bauer, Susanne E.; Schulz, Michael

Publication date:
2021-09

Permanent link:
https://doi.org/10.3929/ethz-b-000507982

Rights / license:
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

Originally published in:
Journal of Advances in Modeling Earth Systems 13(9), https://doi.org/10.1029/2021MS002584
Understanding Top-of-Atmosphere Flux Bias in the AeroCom Phase III Models: A Clear-Sky Perspective

Wenying Su 1, Lusheng Liang 2, Gunnar Myhre 3, Tyler J. Thorsen 1, Norman G. Loeb 4, Gregory L. Schuster 4, Paul Ginoux 5, Fabien Paulot 6, David Neubauer 6, Ramiro Checa-Garcia 6, Hitoshi Matsui 7, Kostas Tsigaridis 8, Ragnhild B. Skeie 9, Toshihiko Takemura 10, Susanne E. Bauer 11, and Michael Schulz 11

1 NASA Langley Research Center, Hampton, VA, USA, 2 Science Systems & Applications Inc., Hampton, VA, USA, 3 CICERO Center for International Climate Research, Oslo, Norway, 4 NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, 5 Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland, 6 Laboratoire des Sciences du Climat et de l’Environnement, IPSL, Gif-sur-Yvette, France, 7 Graduate School of Environmental Studies, Nagoya University, Nagoya, Japan, 8 Center for Climate Systems Research, Columbia University, New York, NY, USA, 9 NASA Goddard Institute for Space Studies, New York, NY, USA, 10 Research Institute for Applied Mechanics, Kyushu University, Fukuoka, Japan, 11 Research Department, Norwegian Meteorological Institute, Oslo, Norway

Abstract Biases in aerosol optical depths (AOD) and land surface albedos in the AeroCom models are manifested in the top-of-atmosphere (TOA) clear-sky reflected shortwave (SW) fluxes. Biases in the SW fluxes from AeroCom models are quantitatively related to biases in AOD and land surface albedo by using their radiative kernels. Over ocean, AOD contributes about 25% to the 60°S–60°N mean SW flux bias for the multi-model mean (MMM) result. Over land, AOD and land surface albedo biases contribute about 40% and 30% to the 60°S–60°N mean SW flux bias for the MMM result. Over ocean, aerosol optical depths (AOD) contribute about 25% to the 60°S–60°N mean SW flux bias for the multi-model mean (MMM) result. Over land, AOD and land surface albedo biases contribute about 40% and 30%, respectively, to the 60°S–60°N mean SW flux bias for the MMM result. Furthermore, the spatial patterns of the SW flux biases derived from the radiative kernels are very similar to those between models and CERES observation, with the correlation coefficient of 0.6 over ocean and 0.76 over land for MMM data using data of 2010. Satellite data used in this evaluation are derived independently from each other, consistencies in their bias patterns when compared with model simulations suggest that these patterns are robust. This highlights the importance of evaluating related variables in a synergistic manner to provide an unambiguous assessment of the models, as results from single parameter assessments are often confounded by measurement uncertainty. Model biases in land surface albedos can and must be corrected to accurately calculate TOA flux. We also compare the AOD trend from three models with the observation-based counterpart. These models reproduce all notable trends in AOD except the decreasing trend over eastern China and the adjacent oceanic regions due to limitations in the emission data set.

Plain Language Summary Aerosol optical depths (AOD) from satellite retrievals have been used to evaluate the AeroCom models. However, these evaluations are often non-conclusive due to uncertainties in the retrievals and the differences among many products. In this study, biases in top-of-atmosphere (TOA) reflected SW fluxes are linked to biases in AOD and surface albedo by using their respective radiative kernels. We found that biases in AOD and land surface albedo can explain significant amount of the SW flux biases on both global and regional scales. This study highlights the importance of evaluating related variables in a synergistic manner to provide an unambiguous assessment of the models, as results from single parameter assessments are often confounded by measurement uncertainty. This study clearly demonstrates the deficiency of land surface albedo used by the AeroCom models and usage of observation-based land surface albedo should be required for all models participating in AeroCom experiments.

1. Introduction

Atmospheric aerosols play important roles in the climate systems directly by scattering and absorbing the solar and terrestrial radiation, and indirectly by modifying the cloud properties. The direct aerosol radiative effect (DARE) and direct aerosol radiative forcing (DARF) are often used to quantify aerosols’ impact on climate. DARE is defined as the mean radiative flux perturbation due to the presence of aerosols (both natural and anthropogenic), while DARF is the anthropogenic component of DARE. Many studies used
satellite measurements to estimate the DARE (e.g., Loeb & Manalo-Smith, 2005; Remer & Kaufman, 2006; Su et al., 2013; Yu et al., 2006). However, determining DARE from satellite measurements is more challenging because current satellite sensors cannot discriminate anthropogenic aerosols from natural aerosols. Our current understanding of DARE relies mostly on the AeroCom model simulations (Myhre et al., 2013; Schulz et al., 2006), with a few studies estimating observational constrained DARE (Bellouin et al., 2005, 2008; Paulot et al., 2018; Su et al., 2013).

As the AeroCom models are vital in advancing our understanding on how aerosols are affecting the climate and its future projection, many aspects of the models have been examined under prescribed conditions. Stier et al. (2013) assessed the host model uncertainties on DARE by prescribing identical aerosol radiative properties in all models and found significant model diversity in simulated clear- and all-sky DARE. One of the variables that contributed to this diversity was surface albedo which had a global mean inter-model relative standard deviation of 4% and significantly larger variability on regional scale. Randles et al. (2013) examined the performance of radiative transfer schemes used in the models and found that diversity among models in the top-of-atmosphere (TOA) DARE depended on the solar zenith angle (SZA) and was the largest for purely scattering aerosols at low SZAs (~20%). They also noted that models overestimated the TOA clear-sky flux under aerosol-free conditions by about 1.3%–3.5% depending on the atmospheric profiles and SZAs.

Aerosol properties from the AeroCom models have been compared against satellite retrievals and AERosol RObotic NETwork (AERONET; Holben et al., 2001) measurements extensively (e.g., Gliß et al., 2021; Kinne et al., 2006). However, radiative fluxes and land surface albedo from the AeroCom models have not been compared with the satellite observations. In this study, TOA fluxes from AeroCom phase III 2019 control experiment are compared with fluxes from the Clouds and the Earth’s Radiant Energy System (CERES; Loeb et al., 2016; Wielicki et al., 1996). We will focus on clear-sky TOA shortwave (SW) flux comparison, because of the large inter-model differences in cloud fraction (Stier et al., 2013). In order to understand the global and regional flux differences between CERES and the AeroCom models, differences in aerosol optical depth (AOD) and land surface albedo (ω) are also presented, as these two variables contribute the most to TOA SW flux under clear-sky conditions. The observational-based SW flux is independently derived from AOD and land surface albedo data sets used in this study based on different satellite instruments and algorithms. When the AOD and surface albedo differences can explain the SW flux differences, it is more likely that these differences are robust. Using different observations to evaluate the models synergistically can therefore be beneficial over using a single variable and provide more reliable diagnostics for model evaluation. Furthermore, relying on radiative kernels for AOD and surface albedo, one can tie the AOD and land surface albedo differences between models and observations to the flux differences. The AeroCom models included in this study are briefly described in Section 2. Satellite observations and data sets are in Section 3, and the radiative kernels are described in Section 4. Results on global and regional comparisons are presented in Section 5, and trend comparisons are in Section 6. Discussions and conclusions are in Section 7.

2. AeroCom Models

This study uses the AeroCom phase III 2019 control experiment (https://wiki.met.no/aerocom/phase3-experiments). For this experiment, models use harmonized anthropogenic and biomass burning emissions from the Community Emission Data System (CEDS; Hoesly et al., 2018) for Coupled Model Intercomparison Project Phase 6 (CMIP6). Modeling centers are required to submit simulation results for at least 2010 and 1850, using 2010 meteorology and prescribed sea surface temperature from input4MIPS (Durack & Taylor, 2018). Among the models that participated in this control experiment, nine models provided all necessary variables (i.e., AOD, TOA reflected SW flux under clear-sky conditions, surface upwelling and downwelling SW flux), and are included in this study. Table 1 lists these models, along with their spatial resolution and references describing the details of each model. All model outputs are linearly interpolated to a 1° × 1° latitude-longitude grid to facilitate comparisons with satellite observational data. The global means calculated using the interpolated grid differ less than 0.01% from using the original spatial resolutions. A brief description of each model is given in Appendix A.
EBAF data has been widely used to evaluate global general circulation models (e.g., Bauer et al., 2020; Paulot et al., 2018; Pincus et al., 2008; Su et al., 2010; H. Wang & Su, 2013; Wild et al., 2013).

### 3. Satellite Observations

#### 3.1. TOA Reflected Shortwave Flux

Clear-sky TOA radiative fluxes for 1° × 1° latitude-longitude regions are from the CERES Energy Balanced and Filled (EBAF) Ed4.1 product (Loeb et al., 2018). EBAF data takes advantage of the many algorithm improvements that have been made in the Edition 4 CERES Level 1–3 data products and it is the only global data set that can be used to study the variations of Earth radiation budget over a range of temporal and spatial scales. The Earth’s energy imbalance in the CERES EBAF is constrained to be consistent with ocean heat content (Johnson et al., 2016) by using an objective constraint algorithm to adjust SW and longwave (LW) TOA fluxes within their ranges of uncertainty to remove the inconsistency between average global net TOA flux and heat storage in the Earth-atmosphere system (Loeb et al., 2009, 2018). Additionally, because of the relatively coarse spatial resolution of the CERES instruments (~20-km at nadir), the standard CERES Level-3 data products have many spatial gaps in monthly mean clear-sky TOA flux maps due to the absence of cloud-free areas occurring at the CERES footprint scale in some 1° × 1° regions. In EBAF, this problem is mitigated by inferring clear-sky fluxes from both CERES and Moderate Resolution Imaging Spectroradiometer (MODIS) measurements to produce a new clear-sky TOA flux climatology for every 1° × 1° grid box every month. The uncertainty for TOA clear-sky SW flux is estimated as 5 W m⁻² on the grid box level (Loeb et al., 2018). EBAF data has been widely used to evaluate global general circulation models (e.g., Bauer et al., 2020; Loeb et al., 2020; Paulot et al., 2018; Pincus et al., 2008; Su et al., 2010; H. Wang & Su, 2013; Wild et al., 2013).

#### 3.2. Aerosol Optical Depth

AODs retrieved from MODIS and Multi-Angle Imaging Spectroradiometer (MISR) are used in this study to compare with the model simulations and to help interpreting the SW flux biases. The Aqua MODIS collection 6.1 monthly gridded Dark Target and Deep Blue merged AOD product (Sayer et al., 2014) combines AODs retrieved from Dark Target over water algorithm with Dark Target and Deep Blue over land algorithms to provide more complete AOD spatial coverage over snow-ice-free surfaces. Validations against AERONET and Maritime Aerosol Network (MAN) data indicate that MODIS dark target retrievals agree well with AERONET over land (high correlation and low bias; Levy et al., 2013). Over ocean, Aqua MODIS AODs are also highly correlated with those from AERONET, but biased high at low AODs and the scatter for high AODs is significantly larger than the expected error (Levy et al., 2013; Schutgens et al., 2020). Validations of the deep blue retrievals over land also indicate very low bias (Sayer et al., 2014; Schutgens et al., 2020).

MISR Level-3 gridded AOD product at 0.5° × 0.5° spatial resolution is also used in this study. The MISR Level-3 product is aggregated from higher spatial resolution version 23 (V23) level 2 data (Garay et al., 2020). Compared to the V22 AOD retrieval, V23 implemented many changes which resulted in significant reduction in AOD over ocean. The 16-year mean AOD is reduced from 0.157 in V22 to 0.114 in V23 over

---

**Table 1**

*List of AeroCom Models Used in This Study*

<table>
<thead>
<tr>
<th>Long name</th>
<th>Short name</th>
<th>Lat × long</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM5-ATRAS</td>
<td>CAM5</td>
<td>1.9° × 2.5°</td>
<td>Matsui (2017)</td>
</tr>
<tr>
<td>ECHAM6.3-HAM2.3-met2010</td>
<td>ECHAM</td>
<td>1.875° × 1.875°</td>
<td>Tegen et al. (2019)</td>
</tr>
<tr>
<td>GFDDLAM4-met2010</td>
<td>GFDDLm</td>
<td>1.00° × 1.25°</td>
<td>Zhao et al. (2018)</td>
</tr>
<tr>
<td>GFDDLAM4-fSST</td>
<td>GFDLf</td>
<td>1.00° × 1.25°</td>
<td>Zhao et al. (2018)</td>
</tr>
<tr>
<td>GISS-ModelE2.1.1-OMA</td>
<td>GISS OMA</td>
<td>2.0° × 2.5°</td>
<td>Bauer et al. (2020); Tsigaridis et al. (2013)</td>
</tr>
<tr>
<td>GISS-ModelE2.1.1-MATRIX</td>
<td>GISS MATRIX</td>
<td>2.0° × 2.5°</td>
<td>Bauer et al. (2008)</td>
</tr>
<tr>
<td>INCA</td>
<td>INCA</td>
<td>2.25° × 2.50°</td>
<td>Balkanski et al. (2004); Schulz et al. (2009)</td>
</tr>
<tr>
<td>MIROC-SPRINTARS</td>
<td>SPRINTARS</td>
<td>0.55° × 0.55°</td>
<td>Takemura et al. (2005)</td>
</tr>
<tr>
<td>OsloCTM3v1.01-met2010</td>
<td>Oslo</td>
<td>2.25° × 2.25°</td>
<td>Lund et al. (2018); Sevde et al. (2012)</td>
</tr>
</tbody>
</table>
ocean. However, the AODs over land agree well between V22 and V23. Validations against the AERONET measurements show modest improvement in V23 retrievals over land in comparison to V22, with the absolute bias decreases from $-0.004$ to $-0.002$ and the percent of retrievals that fall within the error envelope, defined as $\pm 0.03 + 10\%\text{AOD}$, increases from 59.7% to 66.1%. However, V23 retrievals tend to overestimate AOD at low AODs (like MODIS, mentioned above) and underestimate AOD at high AODs. Validations against MAN AOD derived from handheld Microtops Sun photometers show significant improvement in V23, the absolute bias is reduce from 0.037 to 0.0 and the percent of retrievals that fall within the error envelope increases from 61% to 87% in comparison to V22.

### 3.3. Surface Albedo

The Terra and Aqua combined MODIS Bidirectional Reflectance Distribution Function (BRDF)/albedo product (MCD43C1, Version 6) provides the daily weighting parameters to calculate the directional hemispherical reflectance (black-sky albedo) and bihemispherical reflectance (white-sky albedo) at a spatial resolution of 0.05° over land. These parameters are used in polynomial albedo representations to estimate the black-sky albedo at any given SZA and the white-sky albedo (Li & Strahler, 1992; Lucht et al., 2000). The MODIS broadband surface albedos agree well with the in situ ground-based and airborne measurements, the root mean square errors are less than 0.020 (~7%) for forest during the dormant periods and 0.025 (~15%) during the snow-covered periods, less than 0.030 (~9%) for agriculture and grassland during the dormant periods and 0.050 (~4%) during the snow-covered periods, and less than 0.047 (~0.5%) for the snow-covered tundra (Cescatti et al., 2012; Roman et al., 2013; Z. Wang et al., 2012, 2014). The MODIS broadband surface albedos also agree with other satellite products to within 0.01 or ~5% (Carrer et al., 2010; Taberner et al., 2010).

The actual surface albedo is a combination of black-sky albedo and white-sky albedo from the following (Schaf et al., 2002):

$$
\alpha(\theta_0, \text{AOD}, \phi) = \alpha_{ws} \times f(\theta_0, \text{AOD}, \phi) + \alpha_{bs}(\theta_0) \times (1 - f(\theta_0, \text{AOD}, \phi))
$$

(1)

where $\alpha_{ws}$ is the white-sky albedo and $\alpha_{bs}$ is the black-sky albedo, $f$ is the fraction of diffuse light and depends on the SZA ($\theta_0$), AOD, and aerosol types ($\phi$). A look-up table of $f$ (available at https://www.umb.edu/spectralmass/terra_aqua_modis/modis_user_tools) is provided for continental and maritime aerosol types with AODs ranging from 0 to 0.98 for SZAs between 0° and 89°.

For a given month, actual surface albedo is calculated every 10 min to take into account that black-sky albedo is a function of SZA and the value of $f$ is determined using monthly MODIS Dark Target and Deep Blue merged AOD assuming continental aerosol type. At every time step within every $1° \times 1°$ grid, surface albedo calculated at 0.05° resolution are converted to surface upwelling SW fluxes by using the clear-sky surface downwelling SW fluxes from CERES Edition 4 synoptic daily hourly product (SYN1deg-1Hour; Doelling et al., 2013), then average into the $1° \times 1°$ grid. These upwelling SW fluxes are averaged over the month, then divided by the monthly mean downwelling SW flux to produce the monthly mean surface albedo.

### 4. Radiative Kernels

Radiative kernels are partial derivatives of TOA SW flux derived by imposing small systematic perturbations to the base-state values. Thorsen et al. (2020) derived aerosol radiative kernels using 1 year (2007) of 3-hourly MERRA-2 data. Radiative kernels of AOD, single-scattering albedo (SSA), aerosol asymmetry factor, and surface albedo are derived for each month. These kernels are able to reproduce the aerosol direct radiative effect to within 0.3 Wm$^{-2}$ when compared to the true aerosol direct radiative effect calculation in MERRA-2. These kernel calculations are very computational costly, thus kernels are only calculated using 2007 data. Although it would be most accurate to apply these kernels to the same year of measurements and simulations, the impact of interannual variability on these kernels is expected to be small (Thorsen et al., 2020) and they are applied to the AeroCom simulations of 2010 (the output year selected by the AeroCom phase III experiment).

Figure 1 shows the AOD and surface albedo kernels for April and October (differ from the annual mean results presented in Thorsen et al., 2020). Kernels are expressed in units of watts per meter squared per
unit change in the respective variables. Figures 1a and 1b show that the TOA SW flux would increase by more than 30 Wm\(^{-2}\) over the oceans off the west coast of Africa if the AOD were increased by 1, whereas the increment is less than 10 Wm\(^{-2}\) over the Sahara desert because TOA SW flux has smaller sensitivity to AOD changes over bright surface than over dark surface. Aerosol composition also affects the AOD kernels. For example, the AOD radiative kernel over South America is about 25 Wm\(^{-2}\) in April and is reduced to 15 Wm\(^{-2}\) in October (see Figure 1b) during the biomass burning season. Figures 1c and 1d show the regional TOA SW flux changes for unit increase in surface albedo, which are mostly sensitive to solar insolation and are also sensitive to the initial surface albedo. In this study, we use these kernels to assess the contributions of AOD and surface albedo biases to TOA SW flux biases in the AeroCom models.

The clear-sky TOA reflected SW flux bias in the model can be expressed as:

\[
\Delta F = F_m - F_o = \Delta F_{\text{AOD}} + \Delta F_{\text{AF}} + \Delta F_{\text{X}} \tag{2}
\]

where \(F_m\) and \(F_o\) are the TOA reflected SW flux from models and CERES EBAF, \(\Delta F_{\text{AOD}}\) and \(\Delta F_{\text{AF}}\) are the flux biases caused by biases in AOD and surface albedo in models relative to MODIS retrievals calculated using their respective radiative kernels, and \(\Delta F_{\text{X}}\) is the residual in flux bias, which can be attributed to differences in aerosol composition and atmospheric state (i.e., water vapor, ozone, etc.), and also to uncertainties in radiative transfer calculations, satellite retrievals, and kernel calculations.

5. Comparisons Between Models and Observations

Figure 2 shows the seasonal cycle of monthly mean clear-sky SW flux calculated over ocean and land between 60°S and 60°N (60°S–60°N) from CERES EBAF (black line), nine AeroCom models, and the multi-model mean (MMM, black dashed line) result for 2010. Figure 3 shows the seasonal cycle of monthly mean AOD over 60°S–60°N ocean and land from MODIS, MISR, nine AeroCom models, and the MMM. Over ocean, the seasonal cycle of SW flux is very similar to that of solar insolation as the AODs remain relatively constant throughout the year (Figure 3a). Clear-sky SW fluxes from INCA and ECHAM-HAM models are greater than those from CERES EBAF throughout the year, and AODs from these two models are also
on the high end when compared with the satellite retrievals, consistent with the AOD assessment from Gliß et al. (2021). SW flux from Oslo agrees with EBAF almost perfectly, despite its AOD is on the low end when compared with satellite retrievals. Monthly mean clear-sky SW fluxes from all other models are smaller than EBAF by about 2–3 W m⁻², and their AODs are generally smaller than the MODIS retrievals as well, consistent with the findings of Gliß et al. (2021). It is worth noting that even though AODs from SPRINTARS are much lower than all other models, its SW fluxes are actually greater than a couple of other models because its aerosols are almost non-absorbing (see Figure 10). The MMM fluxes are about 1–2 W m⁻² smaller than CERES EBAF and the MMM AODs lie in between MODIS and MISR AODs. Over land, the GFDL models produce greater clear-sky SW fluxes than EBAF, GISS models agree with EBAF fairly well, and all other models have low biases. This results in MMM being 1–2 W m⁻² smaller than CERES EBAF. AODs from all models are outside the boundaries of satellite retrievals. Gliß et al. (2021) also found that almost all AeroCom models underestimate AOD when compared with AERONET and other satellite retrievals. Thus the MMM AODs are smaller than both MODIS and MISR AODs except during the boreal summer months. There are no correspondences between flux biases and AOD biases over land, as land surface albedo biases from these models also contribute to the flux biases.

The seasonal cycles of AODs from MODIS and MISR are very similar to each other over both ocean and land (Figure 3). However, the AOD seasonal cycles from the AeroCom models differ significantly from the observations and from each other, especially over land. The MISR AODs are smaller than the MODIS AODs. As mentioned in Section 3.2, the mean AOD over ocean in the recently released MISR V23 is smaller than V22 by about 0.04, while the mean AOD over land is nearly unchanged. Schutgens et al. (2020)

---

**Figure 2.** Top-of-atmosphere (TOA) clear-sky reflected shortwave (SW) flux comparisons between Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF), individual AeroCom models, and the multi-model mean (MMM) over ocean (a) and land (b) between 60°S and 60°N.

**Figure 3.** Aerosol optical depth (AOD) comparisons between satellite retrievals (Moderate Resolution Imaging Spectroradiometer [MODIS] and Multi-Angle Imaging Spectroradiometer [MISR]), individual AeroCom models, and the multi-model mean (MMM) over ocean (a) and land (b) between 60°S and 60°N.
intercompared AOD retrievals from 14 satellite products (MISR was not included) and evaluated them against the AERONET and MAN results. When collocating each individual satellite product with either AERONET or MAN, they found that over land Aqua MODIS AODs show good agreement with AERONET (high correlation and low bias); Aqua MODIS AODs over ocean are also highly correlated with those from AERONET and MAN, but are biased high. They also argued that the satellite retrieval diversity can be used as a proxy for retrieval uncertainty. Judging from their analysis, it is reasonable to assume Aqua MODIS retrievals represent the high end of AODs whereas MISR retrievals represent the low end over ocean. It is also notable that monthly gridded AODs from MISR often suffer spatial gaps and incoherent features because of MISR’s narrow swath, and the sampling difference also likely contributed to the global mean AOD difference seen between MODIS and MISR. Hence, in the following discussion, we only present results using MODIS AODs, but the regional difference features with respect to models remain the same when MISR AODs are used instead (see Appendix B).

To understand the contributions of AOD and land surface albedo biases to TOA SW flux biases, $\Delta F_{AOD}$ is calculated over ocean and land and $\Delta F_a$ is calculated over land using their respective radiative kernels. Figure 4a shows the monthly mean $\Delta F$ for MMM over $60^\circ S$–$60^\circ N$ and tropical ($30^\circ S$–$30^\circ N$) ocean (solid lines), and the flux residuals ($\Delta F_i$) after accounting for the contribution of AOD biases to $\Delta F$ (dashed lines). Flux biases are reduced by up to 0.6 W m$^{-2}$ over $60^\circ S$–$60^\circ N$ ocean and by up to 1 W m$^{-2}$ over tropical ocean after accounting for $\Delta F_{AOD}$. Figure 4b shows the monthly mean $\Delta F$ for MMM over $60^\circ S$–$60^\circ N$ and tropical land (solid line), and after accounting for the contribution of $\Delta F_{AOD}$ (solid lines with circles) and after accounting for both $\Delta F_{AOD}$ and $\Delta F_a$ (dashed lines). Flux biases are reduced by up to 3 W m$^{-2}$ after accounting for biases in AOD, and are reduced further by accounting for biases in surface albedo.

Table 2 summarizes the annual mean $\Delta F$ for MMM, and how much AOD and surface albedo (only over land) biases contribute to $\Delta F$. Over ocean, accounting for the AOD bias reduces the MMM flux bias by about 25%. Over land, accounting for the AOD and surface albedo biases reduces the MMM flux bias by about 70%. The residual in flux bias can be explained by differences in aerosol composition, radiative transfer calculation uncertainties of the models (Randles et al., 2013), and uncertainties in the radiative kernels and in satellite retrievals.

The monthly mean $\Delta F$ for AeroCom models and their respective $\Delta F_x$ are show in Figure 5 for all models listed in Table 1. Over ocean, accounting for the biases in AOD reduces flux biases for almost all models, with larger impact over tropical oceans than over global oceans. Over land, accounting for the biases in AOD and surface albedo not only reduces the flux biases but also minimizes the seasonal dependence for Oslo.

![Clear-sky SW flux difference: Ocean](image)

**Figure 4.** Monthly mean multi-model mean (MMM) shortwave (SW) flux biases relative to Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) (solid lines) for 2010 over $60^\circ S$–$60^\circ N$ and tropical ($30^\circ S$–$30^\circ N$) ocean (a) and land (b). Over ocean, the dashed lines are the flux bias residuals after accounting for the contributions of aerosol optical depth (AOD) biases to the flux biases. Over land, lines with circles represent ($\Delta F - \Delta F_a$) and dashed lines are the flux bias residuals after accounting for the contributions of both AOD and surface albedo biases to the flux biases.

<table>
<thead>
<tr>
<th>Land</th>
<th>Ocean</th>
<th>Global</th>
<th>Tropical</th>
<th>Global</th>
<th>Tropical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta F$</td>
<td>$\Delta F_{AOD}$</td>
<td>$\Delta F_a$</td>
<td>$\Delta F_x$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-1.0$</td>
<td>$-0.3$</td>
<td>$-0.4$</td>
<td>$-1.7$</td>
<td>$-0.9$</td>
</tr>
<tr>
<td></td>
<td>$-0.4$</td>
<td>$-0.4$</td>
<td>$-0.4$</td>
<td>$-0.4$</td>
<td>$-0.8$</td>
</tr>
<tr>
<td></td>
<td>$-0.3$</td>
<td>$-0.4$</td>
<td>$-0.4$</td>
<td>$-1.3$</td>
<td>$-2.0$</td>
</tr>
</tbody>
</table>

*Note.* AOD, aerosol optical depth; CERES, Clouds and the Earth’s Radiant Energy System; EBAF, Energy Balanced and Filled; SW, shortwave.
SPRINTARS, INCA and CAM5 models. For the GFDL models, accounting for the biases in AOD and surface albedo reverses the models from overestimating to underestimating the flux of about the same magnitude. For ECHAM model, $\Delta F_{\Delta F}$ is slightly more negative than $\Delta F$. For the two GISS models, accounting for the biases in AOD and surface albedo increases the flux biases by about 5 Wm$^{-2}$. This is largely due to the flux biases having little correspondence with the AOD and albedo biases, particularly over the Amazon, central Africa, and the Tibetan Plateau (see Figure 8).

On a 1° × 1° latitude-longitude grid box level, the flux biases due to AOD and surface albedo biases derived from using the radiative kernels correlate very well with $\Delta F$ for MMM. Figure 6a shows the relationship between $\Delta F$ and $\Delta F_{\Delta AOD}$ over ocean using all grid boxes between 50°S–40°N of the entire year of 2010 (approximately 255,000 data points used), and the correlation coefficient is 0.6. MODIS retrievals produce some spuriously large AODs over high latitude oceans (possibly due to the presence of sea ice) and are therefore excluded in the correlation analysis. Figure 6b shows the relationship between $\Delta F$ and $\Delta F_{\Delta AOD} + \Delta F_{\Delta A}$ over land using all grid boxes between 60°S–60°N of the entire year of 2010 (approximately 142,000 data points used), and the correlation coefficient is 0.76. Table 3 lists the correlation coefficients for MMM and nine AeroCom models between $\Delta F$ and $\Delta F_{\Delta AOD}$ over ocean and between $\Delta F$ and $\Delta F_{\Delta AOD} + \Delta F_{\Delta a}$ over land for four seasonal months and the entire year. Over land, MMM has the highest correlation in January and the lowest in October; Over ocean, MMM has the highest correlation in April and the lowest correlation in October. Correlation coefficients for individual models vary across the seasons, and it is not clear what causes these seasonal variations. Over land, the yearly correlation coefficient is the highest for SPRINT model ($r = 0.94$) and lowest for GISS models ($r = 0.78$). Over ocean, the yearly correlation coefficient is the highest for CAM5 and ECHAM-HAM models ($r = 0.78$) and lowest for GISS MATRIX model ($r = 0.52$).

Figure 7 shows the MMM regional TOA SW flux biases (a), SW flux biases due to biases in AOD and surface albedo calculated from their radiative kernels (b), AOD biases (c), and land surface albedo biases(d) for
April 2010. For this month the 60°S–60°N mean $\Delta F_{EF}$ is $-1.8 \text{ Wm}^{-2}$ and $-1.2 \text{ Wm}^{-2}$ over ocean and land. The spatial patterns of flux bias derived from kernels correspond well with $\Delta F$. Correlation coefficient is 0.66 between $\Delta F_{EF}$ and $\Delta F_{AOD}$ over ocean, and is 0.75 between $\Delta F_{EF}$ and $\Delta F_{AOD}$ over land. After accounting for the contributions of AOD and land surface albedo to $\Delta F_{EF}$, the 60°S–60°N mean flux bias ($\Delta F_{EF}$) is reduced to $-1.3 \text{ Wm}^{-2}$ and $0.1 \text{ Wm}^{-2}$ over ocean and land.

Good correspondence between $\Delta F_{EF}$ and flux biases derived from radiative kernels is also observed for all models. Figure 8 shows the regional $\Delta F$ (left panels), and SW flux biases due to AOD and land surface albedo biases calculated from their radiative kernels (right panels) for April 2010. The spatial distributions between the two GFDL models are very similar; hence, only the GFDLf (observed SST without wind nudging) simulation is shown. The GISS MATRIX model performs better than GISS OMA model in terms of simulating sulfate aerosols and AOD (Bauer et al., 2020). We also find that GISS MATRIX model agrees

Table 3

<table>
<thead>
<tr>
<th>Model</th>
<th>Land (60°S–60°N)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Ocean (50°S–40°N)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Jan</td>
<td>Apr</td>
<td>July</td>
<td>Oct</td>
<td>Year</td>
<td>Jan</td>
<td>Apr</td>
<td>July</td>
<td>Oct</td>
<td>Year</td>
</tr>
<tr>
<td>MMM</td>
<td>0.82</td>
<td>0.75</td>
<td>0.73</td>
<td>0.72</td>
<td>0.76</td>
<td>0.65</td>
<td>0.66</td>
<td>0.53</td>
<td>0.48</td>
<td>0.60</td>
</tr>
<tr>
<td>CAM5</td>
<td>0.88</td>
<td>0.81</td>
<td>0.73</td>
<td>0.70</td>
<td>0.81</td>
<td>0.78</td>
<td>0.77</td>
<td>0.79</td>
<td>0.70</td>
<td>0.78</td>
</tr>
<tr>
<td>ECHAM</td>
<td>0.84</td>
<td>0.86</td>
<td>0.69</td>
<td>0.71</td>
<td>0.81</td>
<td>0.78</td>
<td>0.78</td>
<td>0.72</td>
<td>0.68</td>
<td>0.78</td>
</tr>
<tr>
<td>GFDLm</td>
<td>0.83</td>
<td>0.83</td>
<td>0.72</td>
<td>0.71</td>
<td>0.79</td>
<td>0.62</td>
<td>0.71</td>
<td>0.53</td>
<td>0.54</td>
<td>0.61</td>
</tr>
<tr>
<td>GFDLf</td>
<td>0.89</td>
<td>0.86</td>
<td>0.73</td>
<td>0.70</td>
<td>0.84</td>
<td>0.58</td>
<td>0.76</td>
<td>0.53</td>
<td>0.69</td>
<td>0.68</td>
</tr>
<tr>
<td>GISS OMA</td>
<td>0.81</td>
<td>0.78</td>
<td>0.75</td>
<td>0.75</td>
<td>0.78</td>
<td>0.74</td>
<td>0.69</td>
<td>0.62</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>GISS MATRIX</td>
<td>0.80</td>
<td>0.78</td>
<td>0.76</td>
<td>0.75</td>
<td>0.78</td>
<td>0.62</td>
<td>0.52</td>
<td>0.44</td>
<td>0.45</td>
<td>0.52</td>
</tr>
<tr>
<td>INCA</td>
<td>0.85</td>
<td>0.87</td>
<td>0.84</td>
<td>0.85</td>
<td>0.85</td>
<td>0.70</td>
<td>0.72</td>
<td>0.65</td>
<td>0.59</td>
<td>0.67</td>
</tr>
<tr>
<td>SPRINTARS</td>
<td>0.93</td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
<td>0.94</td>
<td>0.55</td>
<td>0.66</td>
<td>0.56</td>
<td>0.54</td>
<td>0.59</td>
</tr>
<tr>
<td>Oslo</td>
<td>0.86</td>
<td>0.88</td>
<td>0.85</td>
<td>0.84</td>
<td>0.85</td>
<td>0.74</td>
<td>0.69</td>
<td>0.54</td>
<td>0.51</td>
<td>0.64</td>
</tr>
</tbody>
</table>
better than GISS OMA when compared to MODIS and MISR AODs, especially over land (Figure 3). Here only results from GISS MATRIX model are shown.

Over ocean, the spatial distribution of $\Delta F$ corresponds well with $\Delta F_{AOD}$ for all models. For April, ECHAM-HAM model has the highest correlation coefficient of 0.78 and GISS MATRIX model has the lowest correlation coefficient of 0.52 (Table 3). A common feature for all models is that they underestimate the dust outflow off the west coast of Africa (Figure 9), thus also underestimate the TOA SW flux. Another common feature is the underestimation of aerosols over tropical ocean, which is consistent with the finding of Kinne et al. (2006) using early AeroCom experiment outputs. All models except CAM5 and SPRINTARS overestimate AOD and SW flux over the Southern Ocean, with INCA model has the largest overestimation. It is worth pointing out that several models (i.e., ECHAM-HAM, GISS MATRIX, and SPRINTARS) significantly overestimate AOD and flux over New Caledonia and surroundings. Over land, the spatial distribution of $\Delta F$ and $\Delta F_{AOD} + \Delta F_{alb}$ are highly correlated, despite that there are large diversities among models in terms of AOD and land surface albedo (Figure 9). The highest correlation coefficient is 0.94 for SPRINTARS model and the lowest is 0.78 for GISS models for April. The high correlations between $\Delta F$ and kernel-based flux biases indicate that AOD and land surface albedo biases can explain most of the SW flux biases. This finding is significant as the MODIS AOD and surface albedo, and the CERES clear-sky flux are derived independently from each other (i.e., they rely upon different measurements and algorithms). Consistencies in their biases indicate that the AOD biases shown in this study are robust and constraining the modeled AODs by satellite observations and correcting the land surface albedo used in the models will improve the SW flux agreement between models and CERES EBAF.

The consistency in spatial distributions between AOD/surface albedo biases and SW flux biases demonstrates that independently derived satellite products are valuable in diagnosing model deficiencies when used jointly. Consistency in regional features of these variables can be helpful in identifying the particular processes and/or parameterizations that are responsible for these biases. For example, when comparing AODs off the west coast of Africa from MODIS with AODs from the models, it is unclear if the MODIS retrieval overestimates the dust outflow from Africa or if the models underestimate the dust outflow. Including comparisons of clear-sky TOA SW fluxes between CERES and models confirms that the models
Figure 8. Top-of-atmosphere (TOA) clear-sky reflected shortwave (SW) flux biases of AeroCom models relative to Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) for April 2010 (left panels). Right panels use radiative kernel calculations to determine the flux biases associated with aerosol optical depth (AOD) and surface albedo biases.
Figure 9. Aerosol optical depth (AOD) biases (left panels) and land surface albedo biases (right panels) of AeroCom models relative to Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals for April 2010.
indeed underestimate the dust outflow there which result in low biases of SW flux. One may argue that overestimating dust aerosol absorption can also contribute to the flux biases. However, to explain the magnitude of flux bias shown in Figure 8, biases of dust aerosol SSA need to be on the order of 0.25 based on the SSA radiative kernels from Thorsen et al. (2020). We thus conclude that the flux biases over the west coast of Africa are mostly due to AOD biases. Additionally, accounting for the impact of AOD and surface albedo biases on SW flux using radiative kernels of these variables significantly improves the regional flux agreement between models and CERES EBAF. The agreements of monthly global and tropical means over ocean and land are also improved for all models except for the GISS models.

As mentioned before, differences in aerosol composition, radiative transfer calculation uncertainties of the models (Randles et al., 2013), and uncertainties in the radiative kernels and in satellite retrievals can all contribute to $\Delta F_s$. The radiative kernels that we use in this study are based on MERRA-2 reanalysis, whose base-state aerosols are different from the AeroCom models. Over ocean, aerosols in MERRA-2 are generally more absorbing than the AeroCom models. Over land, single-scattering albedo (SSA) from many AeroCom models agree reasonably well with that from MERRA-2 (global mean SSA difference is about 0.02–0.03), except that GISS MATRIX model has much lower SSA than MERRA-2 (global mean SSA difference exceeds 0.06) whereas SPRINTARS model has much higher SSA than MERRA-2 (global mean SSA difference is about 0.04). The differences in aerosol composition affects the magnitude of kernel derived flux biases from AOD and surface albedo biases and are part of $\Delta F_s$. Although accounting for the biases in AOD and surface albedo between AeroCom models and satellite retrievals does not entirely eliminate the TOA SW flux bias, it certainly reduces the global mean biases (except for GISS models over land potentially due to compensating errors) and mitigates large regional biases for all models. Currently no aerosol composition observations on the global scales are available for constraining the model simulations, but correcting the aerosol loading and land surface albedo in the models then adjusting the aerosol composition, size distribution, effective refractive indices, and aerosol hygroscopic

Figure 10. Single-scattering albedo for (a) MERRA-2, (b) CAM5, (c) ECHAM, (d) GFDL, (e) GISS MATRIX, (f) INCA, (g) SPRINTARS, (h) Oslo, and (i) multimodel mean (MMM) for April. MERRA-2 result is for 2007, and all others are for 2010.
growth to minimize the residual fluxes can further improve the TOA SW flux agreement between models and observation. Radiative kernels of each model can be developed to achieve maximum consistency to guide the model development.

6. Regional Trends

Quantifying how aerosols and aerosol radiative effect changed over the past and providing future projections are also main AeroCom objectives beside simulating the global aerosol distributions. In this section, trends derived from the model simulations are compared against those derived from the observations. In some regions, land surface albedo has changed significantly in the last decade. To separate the effects of land surface albedo change from aerosol change on TOA clear-sky fluxes, we examine the trends of DARE rather than SW flux, as changes in land surface albedo has a much larger impact on SW flux than on aerosol radiative effect.

Historical simulations from ECHAM-HAM, GFDL (AMIP run), and Oslo models are used to calculate the regional trends in aerosol distributions. Here we compare the regional trends in AODs derived from these models with those from Aqua MODIS retrievals using the time period (July 2002 to December 2017) common to both. All three models use emissions from CEDS until 2014, but use different emissions between 2015 and 2017. Figure 11 shows the AOD trend per decade calculated from Aqua MODIS and from model simulations, and areas with stipplings indicate the trend is significant at the 95% confidence interval. All three models reproduce the decreasing aerosol trends over Europe, eastern United States and the Atlantic Ocean, and the increasing trends over India, Indian Ocean, Arabian Peninsula, and Central Africa. The most notable difference among the models is over eastern China and the adjacent oceanic regions where MODIS indicates a decreasing trend. The vastly different trends among the models are due to different emission
data used between 2015 and 2017 for the three models, and the fact that some pathway scenarios significantly underestimate the recent decline in anthropogenic aerosol emissions over China (Z. Wang et al., 2021).

The DARE is the difference between the TOA SW flux for a pristine atmosphere with no aerosols and the actual atmosphere with aerosols (Loeb et al., 2021):

$$DARE = F(0, \alpha, X) - F(AOD, \alpha, X)$$  
(3)

where \(\alpha\) is the surface albedo and \(X\) represents other atmospheric variables that influence TOA SW flux. The deseasonalized anomaly in DARE is:

$$\delta DARE = \delta F(0, \alpha, X) - \delta F(AOD, \alpha, X)$$  
(4)

Deseasonalized DARE anomaly and trend can be readily calculated from model outputs. However, it is more challenging to derive DARE from observations as pristine conditions cannot be observed directly. Assuming the contribution of other atmospheric variables to \(\delta F(0, \alpha, X)\) is negligible, then \(\delta DARE\) can be expressed as follows:

$$\delta DARE = \delta F(\alpha) - \delta F(AOD, \alpha, X)$$  
(5)

here, \(\delta F(\alpha)\) is the TOA flux anomaly due to surface albedo change, and can be derived from surface albedo anomaly and radiative kernel (Paulot et al., 2018). Over land, surface albedo is based on MCD43C1 described in Section 3.3. Over ocean, surface albedo is specified from a look-up table based on the Coupled Ocean Atmosphere Radiation Transfer model (Jin et al., 2004; Rutan et al., 2009). We can derive \(\delta F(AOD, \alpha, X)\) directly from EBAF data.

Figure 12 shows DARE trend (Wm\(^{-2}\) per decade) calculated using Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) (a) and using ECHAM-HAM (b), GFDL (c), and Oslo (d) model simulations. Data from July 2002 to December 2017 are used here.
ECHAM-HAM model shows a strong increasing trend over eastern China, whereas trends from EBAF, GFDL, and Oslo are very muted. The observation-based DARE trends are more pronounced over water than land, while the opposite is true for AOD trends. Loeb et al. (2021) concluded that this is due to a strong water-land contrast in DARE efficiency over eastern China that is likely enhanced by strong absorption by aerosols.

Figure 13 further exams the deseasonalized anomalies of SW DARE over three land regions listed in Table 4. Over eastern China, DARE from EBAF shows a small increasing trend (Table 4), whereas model simulations show very different trends. The strong increasing trend from ECHAM-HAM is partly due to the large emission reduction introduced by SSP3-7.0 starting in 2015. Over India, both EBAF and model simulations show decreasing trends, with EBAF shows the sharpest decline. This could be partly due to the decreasing trend in surface albedo over India that results in larger radiative efficiency. Large DARE variabilities from ECHAM-HAM over China and India are due to large dust emission variabilities used by the model. Over eastern USA, anomalies and trends from EBAF and model simulations show excellent agreement.
7. Conclusions

AeroCom models have played an essential role in advancing our understanding of DARF, though large diversity still exists among the models. To improve the model performance and to understand the root causes of the large diversity among them, the models have been evaluated against in-site and satellite observations. To date, evaluations have been mainly focusing on AOD. In this study, we evaluate TOA clear-sky reflected SW fluxes from the AeroCom models against the SW fluxes from CERES EBAF. Additionally, AODs and land surface albedo from AeroCom models are also evaluated against satellite retrievals in order to explain the SW flux biases.

To quantify how much the SW flux biases can be explained by the biases in AOD and land surface albedo, we use the radiative kernels of AOD and land surface albedo developed from MERRA-2 reanalysis to attribute their contributions to SW flux biases. Over ocean, the 60°S–60°N mean TOA SW flux bias in MMM is reduced by 25% after accounting for the contribution from AOD bias. Over land, the 60°S–60°N mean TOA SW flux bias in MMM is reduced by 70% after accounting for the contribution from biases in AOD and land surface albedo. Furthermore, the spatial patterns of the flux bias derived from the radiative kernels are very similar to those between models and CERES EBAF, with the correlation coefficient of 0.6 over ocean and 0.76 over land for MMM using data of 2010. The correlation coefficients for all models considered in this study are also high, ranging from 0.52 to 0.78 over ocean and from 0.78 to 0.95 over land. The high correlation indicates that most of the SW flux biases can be explained by the biases in AOD and surface albedo between models and observations. Given that the CERES EBAF TOA SW flux is independently derived from MODIS AOD and land surface albedo, consistencies in their bias patterns when compared with model simulations suggest that these features are robust. In addition, the regional patterns of flux bias are unique to each model, which point to the deficiency in each model in simulating the specific aerosols in different source regions.

The AOD and DARE trends from ECHAM-HAM, GFDL, and Oslo model are compared with the observation-based counterparts. All three models reproduce the decreasing trends in MODIS AOD over Europe, eastern United States and the Atlantic Ocean, and central South America, and the increasing trends over India, the Indian Ocean, Arabian Peninsula, and Central Africa. The models fail to reproduce the decreasing trend in AOD over eastern China and the adjacent oceanic regions due to limitations in the emission data set.

Using independently derived satellite data sets (TOA reflected SW flux, AOD, and surface albedo) to assess the AeroCom models provide an opportunity to evaluate related variables in a synergistic manner, thus provide an unambiguous assessment of the model performance and point to ways that can improve the aerosol simulations. Regional bias patterns in these variables, when they corroborate each other, offer a more convincing assessment of the model performance and possibly the cause of the differences. Radiative kernels provide a convenient way to link the AOD and surface albedo biases to TOA SW flux biases, which can be used as a diagnostic tool for model development. All models should correct their land surface albedo by using satellite-derived product as inputs and constrain AODs using satellite retrievals and the AeroCom community should make this a requirement for all models participating in the AeroCom experiments. Implementing these changes will improve the global and regional SW flux agreement between models and satellite observations, and reduce the diversity among the models.

Table 4

| Aerosol Direct Radiative Effect Trends (W m⁻² per Decade) Over China, India, and USA |
|----------------------------------|----------------------------------|----------------------------------|
|                                  | China                            | India                            | USA                              |
| EBAF                            | 0.4 ± 0.3                        | −1.9 ± 0.3                        | 1.2 ± 0.3                        |
| ECHAM                           | 1.8 ± 0.8                        | −0.7 ± 0.3                        | 1.6 ± 0.3                        |
| GFDL                            | 0.0 ± 0.3                        | −0.8 ± 0.2                        | 1.3 ± 0.2                        |
| Oslo                            | 0.2 ± 0.2                        | −1.1 ± 0.3                        | 1.0 ± 0.2                        |

Note. Only land areas in the indicated latitude and longitude ranges are included in the regional trend calculation.
Appendix A

A brief description of each model used in this study is provided below.

A1. CAM5-ATRAS

The Community Atmosphere Model version 5 (CAM5) with the Aerosol Two-dimensional bin module for foRmation and Aging Simulation (ATRAS) uses a two-dimensional sectional aerosol representation with 12 particle size bins (from 1 to 10,000 nm in dry diameter) and eight black carbon mixing state bins (Matsui, 2017). The CAM5-ATRAS model considers the following atmospheric aerosol processes: emissions, new particle formation, condensation of sulfate, nitrate, and organic aerosols, coagulation, activation, aqueous-phase formation, dry and wet deposition, and aerosol-radiation-cloud interactions. Aerosol optical properties are calculated based on the Mie theory code (Bohren & Huffman, 1998), and radiative transfer for shortwave (SW) and longwave (LW) is calculated by the Rapid Radiative Transfer Method for GCMs (Iacono et al., 2008). CAM5-ATRAS aerosol simulations have been evaluated by surface, aircraft, and satellite observations in our previous studies (e.g., Matsui & Mahowald, 2017; Matsui & Moteki, 2020; Matsui et al., 2018).

A2. ECHAM6.3

ECHAM6.3-HAM2.3 is the combination of the global climate model ECHAM6.3 (Mauritsen et al., 2019) and the Hamburg Aerosol Module (HAM2.3; Tegen et al., 2019). It uses the PSRad (Pincus & Stevens, 2013) two-stream radiative transfer scheme with 16 SW and 14 LW bands. The Monte Carlo independent column approximation is used for cloud overlap. Aerosol radiative properties are pre-computed using Mie theory and read from a look-up table based on Mie size-parameter and volume-weighted real and imaginary refractive index for seven aerosol modes containing up to five aerosol species (sulfate, black carbon, particulate organic matter, sea salt and dust) and aerosol water. Aerosol water uptake is based on kappa-Koehler theory (O’Donnell et al., 2011). In subtropical oceans where shallow convective clouds are prevalent, AOD is overestimated likely because precipitation from shallow convective clouds is only allowed if the clouds reach a certain thickness (Muench & Lohmann, 2020). Furthermore, black carbon and organic carbon concentrations are underestimated to some extent, which may be due to underestimated biomass burning emissions and cause to low AOD in biomass burning regions (Tegen et al., 2019).

A3. GFDL

The Geophysical Fluid Dynamics Laboratory Atmospheric Model version 4 (AM4) has cubed-sphere topology with 96 × 96 grid boxes per cube face (C96; approximately 100 km grid size) and 33 levels in the vertical, contains an aerosol bulk model that generates mass concentration of aerosol fields (sulfate, carbonaceous aerosols, sea salt and dust) from emissions and a “light” chemistry mechanism designed to support the aerosol model but with prescribed ozone and radicals (Zhao et al., 2018). Simulations up to the year 2014 are driven by time-varying boundary conditions, and natural and anthropogenic forcings developed in support of CMIP6 (Eyring et al., 2016), except for ship emission of SOx (black carbon ship emission is included). For the following simulated years, the anthropogenic emissions for 2014 are repeated. The dust emission is driven by the simulated winds from constant sources with their erodibility expressed as a function of surrounding topography (Ginoux et al., 2001). The sea salt emissions are based on Martensson et al. (2003) and Monahan et al. (1986) for fine and coarse mode particles, respectively. Aerosols are externally mixed except for black carbon, which is internally mixed with sulfate. The optical properties of the mixture are calculated by volume weighting of their refractive indices using a Mie code. The GFDL-AM4-met2010 (GFDLm) and GFDL-AM4-fSST (GFDLf) models are run with observed sea surface temperature and sea-ice distribution. In addition for GFDLm, the wind components are nudged, with a 6-h relaxation time, toward the NCEP-NCAR reanalysis (Kalnay et al., 1996). The diagnostics are projected from the C96 cubed-sphere to equally spaced 1° latitude and 1.25° longitude grid using first order conservative method. In GFDL model, the aerosol effect is estimated by calling the radiative transfer scheme twice, with and without aerosols in the absence of clouds. The radiative time step is 1 h for SW and 3 h for LW. The SW code is an update of
the 18-band formulation of Freidenreich and Ramaswamy (2005). These updates are described in detail by Zhao et al. (2018). They are related to H$_2$O, CO$_2$, and O$_3$ formulations, and SW absorption by CH$_4$ and N$_2$O. In addition, the effects of the SW water self-continuum and the O$_2$ and N$_2$ continua have been updated.

A4. GISS

The GISS model hosts two aerosol schemes, the GISS One-Moment Aerosol (OMA) and the Multiconfiguration Aerosol TRacker of miXing state (MATRIX) models use the same aerosol emissions. Sea salt, dimethyl sulfide, isoprene, and dust emission fluxes are calculated interactively. Anthropogenic dust sources are not represented in ModelE2.1. Dust emissions vary spatially and temporally only with the evolution of climate variables like wind speed and soil moisture (Miller et al., 2006). OMA is a mass-based scheme including sea salt, dust, sulfate, nitrate, ammonium, carbonaceous aerosols (BC and OC) aerosols (Tsigaridis et al., 2013). Aerosols are externally mixed and assumed to have a prescribed constant size distribution. Aerosol hydration in OMA is calculated in the radiation code following Tang and Munkelwitz (1994). MATRIX (Bauer et al., 2008) is an aerosol microphysics scheme that tracks aerosol mixing state, based on the quadrature method of moments, in which the amount of water in aerosol is calculated with the aerosol thermodynamics module EQSAM (Metzger et al., 2002), using the phase state of an ammonia-sulfate-nitrate-water inorganic aerosol (OA) in thermodynamic equilibrium for metastable aerosols, except for sea salt where the Lewis parameterization is used (Lewis & Schwartz, 2013). As such, hygroscopic swelling of aerosol is already considered and does not need to be recalculated during the radiative calculations.

A5. INCA

I Nteraction with Chemistry and Aerosols (INCA) is a chemistry-aerosol model coupled to a land surface and a dynamical model. INCA simulates dust, sea salt, black carbon (BC), NO$_3$, SO$_4$, SO$_2$, and organic aerosol (OA) with a combination of accumulation, coarse, and super-coarse modes, as well as soluble and insoluble components (Schulz et al., 2009). Because of the simplified chemistry scheme, DMS emissions are prescribed and not interactively calculated, and the secondary organic aerosols are not simulated therefore this specific run is underestimating the OA. In the current version BC soluble mode is internally mixed with sulfate (R. Wang et al., 2016), for which the refractive index is estimated using the Maxwell-Garnett method, improving the accuracy of the BC optical absorption properties.

The radiative transfer model for the calculations with aerosols relies on the RRTM model as implemented by the European Centre for Medium-Range Weather Forecasts, a model that we used for SW and LW calculations. The number of spectral bands used for aerosols is 6 for SW and 16 for LW spectrum. The spectral dependence of optical properties of each aerosol species has been estimated with Mie theory of spherical particles with log-normal distribution.

A6. MIROC-SPRINTARS

An aerosol climate model, Spectral Radiation Transport Model for Aerosol Species (SPRINTARS; Takemura et al., 2005, 2009), is incorporated into a coupled atmosphere-ocean general circulation model, MIROC6 (Tatebe et al., 2019). The horizontal and vertical resolutions are T213 (~0.5625° x 0.5625° in longitude and latitude) and L56, respectively. SPRINTARS calculates the aerosol-radiation and aerosol-cloud interactions by coupling the radiation and cloud-precipitation schemes, respectively as well as aerosol transport processes. The radiative transfer scheme, mstrnX, adopt a two-stream discrete-ordinate method with a correlated k-distribution method (Sekiguchi & Nakajima, 2008). Scattering and absorption of solar and terrestrial radiation by aerosols are calculated assuming the Mie theory with refractive indices of dry aerosols and water from d’Almeida et al. (1991). The volume-weighted refractive indices are assumed for internally mixed particles between black carbon and organic aerosols as well as aerosols and water.

A7. Oslo

The OsloCTM3 is a global, offline chemical transport model (CTM) driven by 3-hourly meteorological data from the European Centre for Medium-Range Weather Forecast (ECMWF) Integrated Forecast System
(IFS) model (Lund et al., 2018; Sovde et al., 2012). The model is run in a 2.25° × 2.25° horizontal resolution, with 60 vertical levels (the uppermost centered at 0.1 hPa). The treatment of transport and scavenging, as well as individual aerosol modules, is described in detail in Lund et al. (2018) and references therein. The aerosol optical properties in OsloCTM3 are described in Myhre et al. (2007) with some recent updates, where the BC mass absorption coefficient (MAC) is following the formula in Zanatta et al. (2016) and a weak absorption implemented for OA (Lund et al., 2018).

**Appendix B**

Figure B1 shows the regional AOD biases of the AeroCom models relative to MISR retrievals (left panels) and the regional SW flux biases due to AOD biases (relative to MISR retrievals) and land surface albedo biases (relative to MODIS retrievals) calculated from their radiative kernels (right panels) for April 2010. Many of the regional AOD bias patterns shown here are very similar to the AOD biases shown in Figure 9. The SW flux biases calculated from the radiative kernels using MISR AODs also resemble those shown in Figure 8. However, the biases over the tropical oceans are much muted when MISR AOD is used. The correlation coefficients between ΔF and ΔF_{AOD} + ΔF_{s} range from 0.79 to 0.94 over land, which is very similar to those derived when MODIS AOD is used. The correlation coefficients between ΔF and ΔF_{AOD} range from 0.26 to 0.63 over ocean, not as high as when MODIS AOD is used. The reduced correlation over ocean is partly due to retrieval differences between MODIS and MISR, but largely due to MISR sampling issue as evident in the stripping features of the AOD bias plots.
Figure B1. The monthly mean aerosol optical depth (AOD) biases of AeroCom models relative to Multi-Angle Imaging Spectroradiometer (MISR) retrieval (left panels), top-of-atmosphere (TOA) shortwave (SW) flux biases due to AOD biases (models—MISR) and land surface albedo biases (models—Moderate Resolution Imaging Spectroradiometer [MODIS]) calculated from their respective radiative kernels (right panels) for April 2010.
Data Availability Statement

All AeroCom simulations are available at the Norwegian Meteorological Institute. The CERES EBAF Ed4.1 data were obtained from https://ceres.larc.nasa.gov/data/. MODIS 495 MYD08_M3_0_6_1 550 nm AOD Dark Target+Deep Blue Combined data were obtained from the Giovanni online data system, developed and maintained by the NASA CES DISC. The V6 MODIS Bidirectional Reflectance Distribution Function (BRDF)/albedo products (MCD43C1) were obtained from the Land Processes Distributed Active Archive Center (LP DAAC) through https://lpdaac.usgs.gov/products/mcd43c1v006/.

References


