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 [1], Gregory L. Schuster [1], Paul Ginoux [4], Fabien Paulot [4], David Neubauer [5], Ramiro Checa-Garcia [6], Hitoshi Matsui [7], Kostas Tsigaridis [8,9], Ragnhild B. Skeie [3], Toshihiko Takemura [10], Susanne E. Bauer[9], Michael Schulz [11]

[1]NASA Langley Research Center, Hampton, Virginia, USA
[2]Science Systems & Applications Inc., Hampton, Virginia, USA
[3]CICERO Center for International Climate Research, Oslo, Norway
[4]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, 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

Key Points:

- Model biases in clear-sky top-of-atmosphere reflected shortwave fluxes are linked to biases in aerosol optical depth and surface albedo
- Over ocean, aerosol optical depth bias contributes about 25% to the 60°S-60°N mean SW flux bias for the multi-model mean (MMM) result
- Over land, aerosol optical depth 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

Corresponding author: Wenying Su, `Wenying.Su-1@nasa.gov`

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 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 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. We also compare the AOD trend from three models with the observation-based counterpart. These models reproduce all notable trends in AOD (i.e. decreasing trend over eastern United States and increasing trend over India) except the decreasing trend over eastern China and the adjacent oceanic regions due to limitations in the emission dataset.

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 reflected shortwave fluxes are linked to biases in aerosol optical depth and surface albedo by using their respective 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 contribute about 40% and 30%, respectively, to the 60°S-60°N mean SW flux bias for the MMM result. The spatial patterns of the SW flux biases derived from the radiative kernels are very similar to the model flux bias relative to satellite observations, with the correlation coefficient of 0.6 over ocean and 0.76 over land for MMM using data of 2010. 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.

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; Yu et al., 2006; Su et al., 2013). However, determining DARF from satellite measurements is more challenging because current satellite sensors cannot discriminate anthropogenic aerosols from natural aerosols. Our current understanding of DARF relies mostly on the AeroCom model simulations (Schulz et al., 2006; Myhre et al., 2013), with a few studies estimating observational constrained DARF (Bellouin et al., 2005, 2008; Su et al., 2013; Paulot et al., 2018).

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 DARF by prescribing identical aerosol radiative properties in all models and found significant model diversity in simulated clear- and all-sky DARF. 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) DARF 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 to 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., Kinne et al., 2006; Gliß et al., 2021). However, radiative fluxes 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) (Wielicki et al., 1996; Loeb et al., 2016). 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 (α) 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 datasets 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 datasets 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, 9 models provided all necessary variables (i.e. aerosol optical depth, 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 $1^\circ \times 1^\circ$ 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.

Table 1. List of AeroCom models used in this study.

Long name	Short name	lat × long	Reference
CAM5-ATRAS	CAM5	1.9° × 2.5°	Matsui2017
ECHAM6.3-HAM2.3-met2010	ECHAM	1.875° × 1.875°	Tegen2019
GFDL-AM4-met2010	GFDLm	1.00° × 1.25°	Zhao2018
GFDL-AM4-fSST	GFDLf	1.00° × 1.25°	Zhao2018
GISS-ModelE2.1.1-OMA	GISS OMA	2.0° × 2.5°	Tsigaridis2013, Bauer2020
GISS-ModelE2.1.1-MATRIX	GISS MATRIX	2.0° × 2.5°	Bauer2008
INCA	INCA	2.25° × 2.50°	Balkanski2004, Schulz2009
MIROC-SPRINTARS	SPRINTARS	0.55° × 0.55°	Takemura2005
OsloCTM3v1.01-met2010	Oslo	2.25° × 2.25°	Lund2018, Sovde2012

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 dataset 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 constrainment 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 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 Wm^{-2} on the grid box level (Loeb et al., 2018). EBAF data has been widely used to evaluate global general circulation models (e.g., Pincus et al., 2008; Su et al., 2010; H. Wang & Su, 2013; Wild et al., 2013; Paulot et al., 2018; S. E. Bauer et al., 2020; Loeb et al., 2020).

3.2 Aerosol Optical Depth

Aerosol optical depths (AOD) retrieved from Moderate Resolution Imaging Spectroradiometer (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 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 ±(0.03+10%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 for forest during the dormant periods and 0.025 during the snow-covered periods, less than 0.030 for agriculture and grassland during the dormant periods and 0.050 during the snow-covered periods, and less than 0.047 for the snow covered tundra (Z. Wang et al., 2012, 2014; Cescatti et al., 2012; Roman et al., 2013). The MODIS broadband surface albedos also agree with other satellite products to within 0.01 (Taberner et al., 2010; Carrer et al., 2010).

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

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

where α_{ws} is the white-sky albedo and α_{bs} is the black-sky albedo, f is the fraction of diffuse light and depends on the SZA (θ_0), AOD, and aerosol types (ϕ). 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 minutes 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°×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°×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.

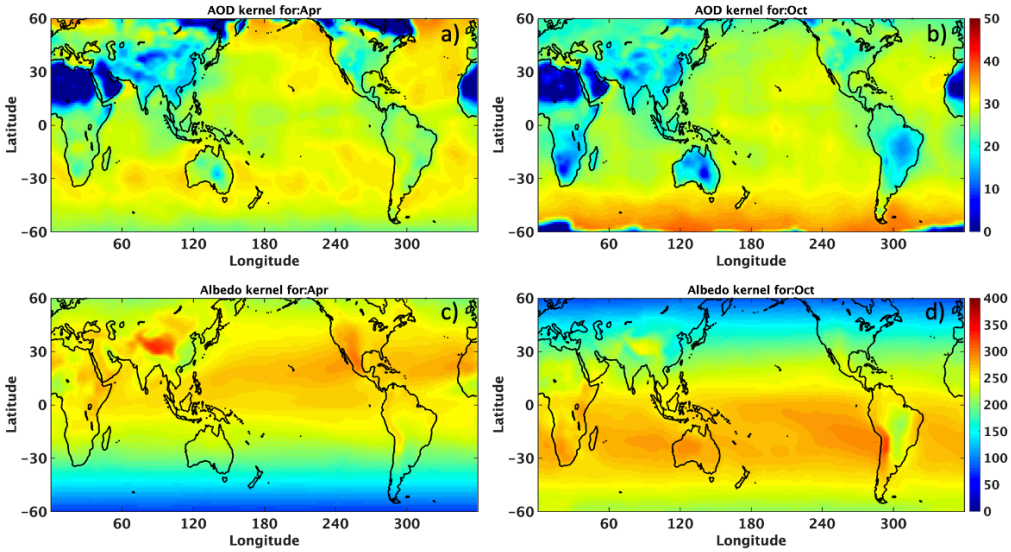

Figure 1. The monthly mean TOA shortwave flux radiative kernels (in Wm^{-2}) for aerosol optical depth (a) April, (b) October, and surface albedo (c) April, (d) October.

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 aerosol optical depth, 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. Figure 1c and 1d shows 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:

$$\Delta F = F_m - F_o = \Delta F_{AOD} + \Delta F_\alpha + \Delta F_\chi, \tag{2}$$

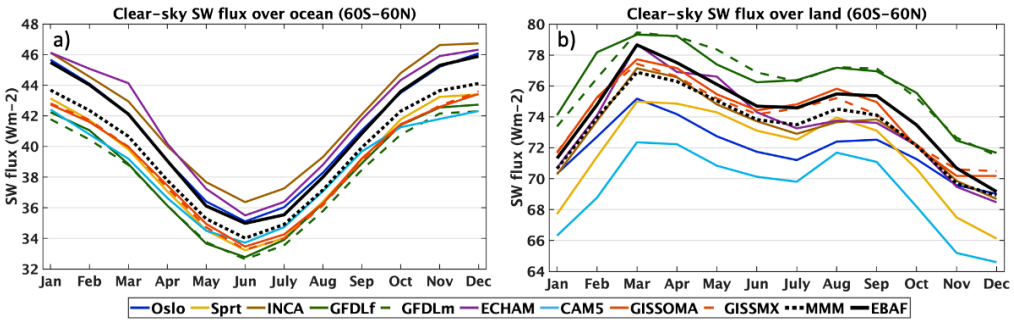

Figure 2. TOA clear-sky reflected SW flux comparisons between CERES EBAF, individual AeroCom models, and the multi-model mean (MMM) over ocean (a) and land (b) between 60°S and 60°N.

where F_m and F_o are the TOA reflected SW flux from models and CERES EBAF, ΔF_{AOD} and ΔF_α 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 ΔF_χ 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 (60S-60N) 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 60S-60N 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 Wm^{-2}, and their AODs are generally smaller than the MODIS retrievals as well. It is worth noting that even though AODs from SPRINT-ARS 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 Wm^{-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 Wm^{-2} smaller than CERES EBAF. AODs from all models are outside the boundaries of 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 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

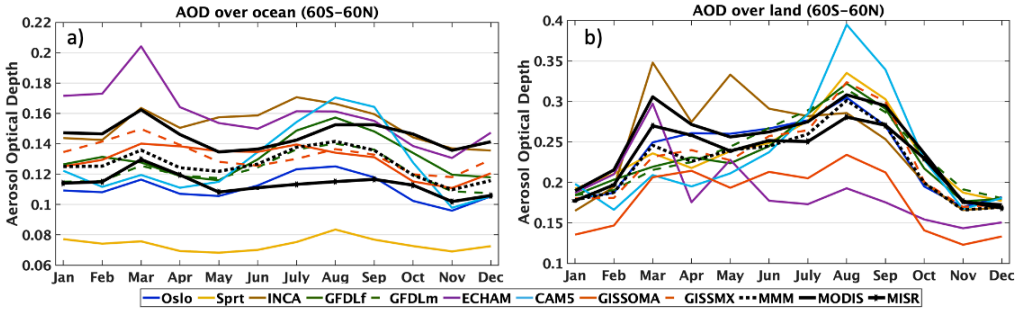

Figure 3. Aerosol optical depth comparisons between satellite retrievals (MODIS and MISR), individual AeroCom models, and the multi-model mean (MMM) over ocean (a) and land (b)between 60°S and 60°N.

Table 2. Annual mean SW flux biases of multi-model mean relative to CERES EBAF (ΔF) over land and ocean, flux biases due to AOD biases ΔF_{AOD} and land surface albedo biases ΔF_{α}, and the flux bias residues ΔF_{χ}.

	Land		Ocean	
	Global	Tropical	Global	Tropical
ΔF	-1.0	-1.7	-1.7	-2.8
ΔF_{AOD}	-0.4	-0.4	-0.4	-0.8
ΔF_{α}	-0.3	-0.9	—	—
ΔF_{χ}	-0.3	-0.4	-1.3	-2.0

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) 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, ΔF_{AOD} is calculated over ocean and land and ΔF_{α} is calculated over land using their respective radiative kernels. Figure 4a shows the monthly mean ΔF for MMM over 60S-60N and tropical (30°S-30°N) ocean (solid lines), and the flux residuals (ΔF_{χ}) after considering the contribution of AOD biases to ΔF (dashed lines). Flux biases are reduced by up to 0.6 Wm^{-2} over 60S-60N ocean and by up to 1 Wm^{-2} over tropical ocean after accounting for ΔF_{AOD}. Figure 4b shows the monthly mean ΔF for MMM over 60S-60N and tropical land (solid line), and after accounting for the contri-

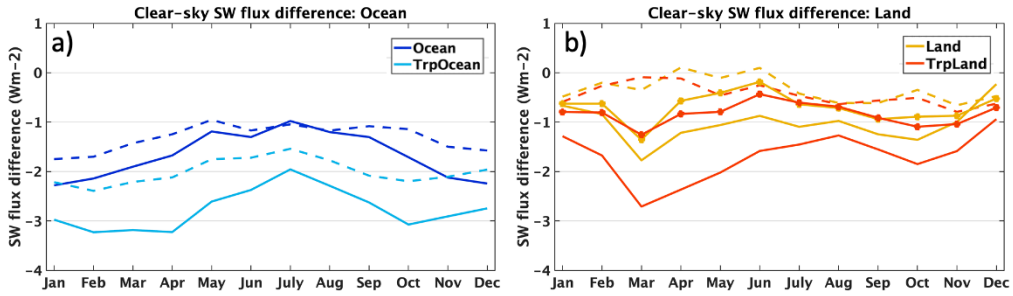

Figure 4. Monthly mean multi-model mean SW flux biases relative to CERES EBAF (solid lines) for 2010 over 60°S–60°N and tropical (30°S–30°N) ocean (a) and land (b). Over ocean, the dashed lines are the flux bias residuals (ΔF_χ) after accounting for the contributions of AOD biases to the flux biases. Over land, lines with circles represent($\Delta F \quad - \quad \Delta F_\alpha$) and dashed lines are the flux bias residuals after accounting for the contributions of both AOD and surface albedo biases to the flux biases.

bution of ΔF_{AOD} (solid lines with circles) and after accounting for both ΔF_{AOD} and ΔF_α (dashed lines). Flux biases are reduced by up to 3 Wm^{-2} after accounting for biases in AOD, and are reduced further by accounting for biases in surface albedo. Table 2 summarizes the annual mean ΔF for MMM, and how much AOD and surface albedo (only over land) biases contribute to ΔF. 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 ΔF for AeroCom models and their respective ΔF_χ 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, 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, ΔF_χ is slightly more negative than ΔF. 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 ΔF for MMM. Figure 6a shows the relationship between ΔF and ΔF_{AOD} over ocean using all gird 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 ΔF and ΔF_{AOD}+ΔF_α over land using all gird boxes between 60°S–60°N for 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 ΔF and ΔF_{AOD} over ocean and between ΔF and ΔF_{AOD}+ΔF_α over land for four seasonal months and the entire year. Over land, MMM has the highest cor-

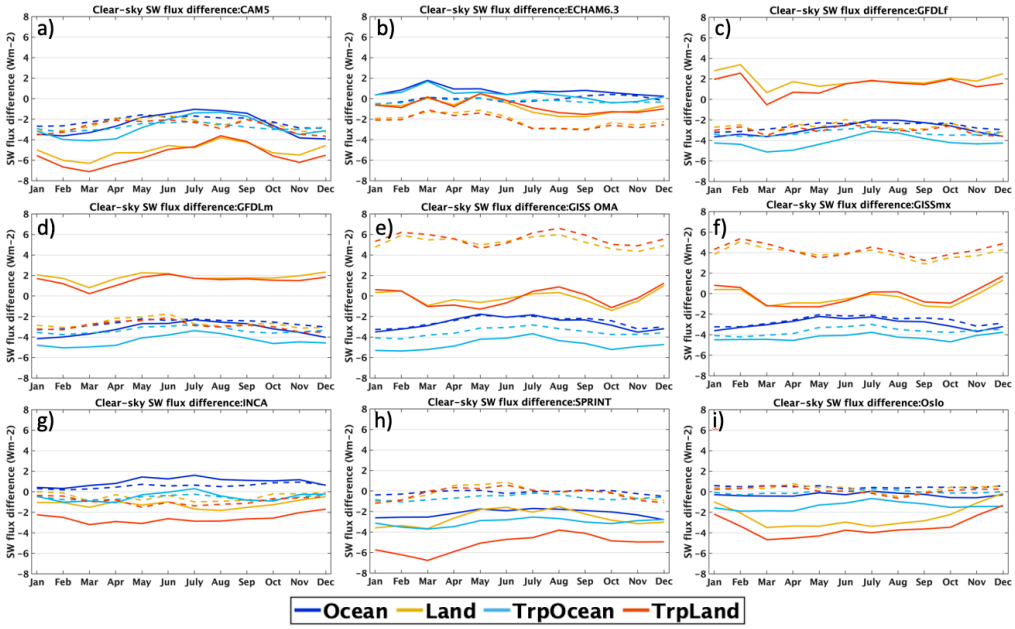

Figure 5. Monthly mean SW flux biases of the AeroCom models for 2010 over 60°S–60°N ocean (solid blue lines) and land (solid orange lines), and over tropical (30°S–30°N) ocean (solid cyan lines) and land (solid red lines). The dashed lines are the flux bias residuals (ΔF_χ) for the corresponding regions. (a) CAM5, (b) ECHAM, (c) GFDLf, (d) GFDLm, (e) GISS OMA, (f) GISS MATRIX, (g) INCA, (h) SPRINTARS, and (i) Oslo.

relation 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. 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 60S-60N mean ΔF is -1.8 Wm^{-2} and -1.2 Wm^{-2} over ocean and land. The spatial patterns of flux bias derived from kernels correspond well with ΔF. Correlation coefficient is 0.66 between ΔF and ΔF_{AOD} over ocean, and is 0.75 between ΔF and ΔF_{AOD}+ΔF_α over land. After accounting for the contributions of AOD and land surface albedo to ΔF, the 60S-60N mean flux bias (ΔF_χ) is reduced to -1.3 Wm^{-2} and 0.1 Wm^{-2} over ocean and land.

Good correspondence between ΔF and flux biases derived from radiative kernels is also observed for all models. Figure 8 shows the regional ΔF (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 (S. E. Bauer et al., 2020). We also find that GISS MATRIX model agrees better than GISS OMA when compared to MODIS and MISR AODs. Here only results from GISS MATRIX model are shown.

Table 3. Correlation coefficients between ΔF and $\Delta F_{AOD}+\Delta F_{\alpha}$ over land, and between ΔF and ΔF_{AOD} over ocean for January, April, July, October and the entire year of 2010 for multi-model mean (MMM) result and the nine AeroCom models.

Model	Land (60°S–60°N)					Ocean (50°S–40°N)				
	Jan	Apr	July	Oct	Year	Jan	Apr	July	Oct	Year
MMM	0.82	0.75	0.73	0.72	0.76	0.65	0.66	0.53	0.48	0.60
CAM5	0.88	0.81	0.73	0.70	0.81	0.78	0.77	0.79	0.70	0.78
ECHAM	0.84	0.86	0.69	0.71	0.81	0.78	0.78	0.72	0.68	0.78
GFDLm	0.83	0.83	0.72	0.71	0.79	0.62	0.71	0.53	0.54	0.61
GFDLf	0.89	0.86	0.73	0.70	0.84	0.58	0.76	0.53	0.69	0.68
GISS OMA	0.81	0.78	0.75	0.75	0.78	0.74	0.69	0.62	0.68	0.68
GISS MATRIX	0.80	0.78	0.76	0.75	0.78	0.62	0.52	0.44	0.45	0.52
INCA	0.85	0.87	0.84	0.85	0.85	0.70	0.72	0.65	0.59	0.67
SPRINTARS	0.93	0.94	0.95	0.94	0.94	0.55	0.66	0.56	0.54	0.59
Oslo	0.86	0.88	0.85	0.84	0.85	0.74	0.69	0.54	0.51	0.64

height

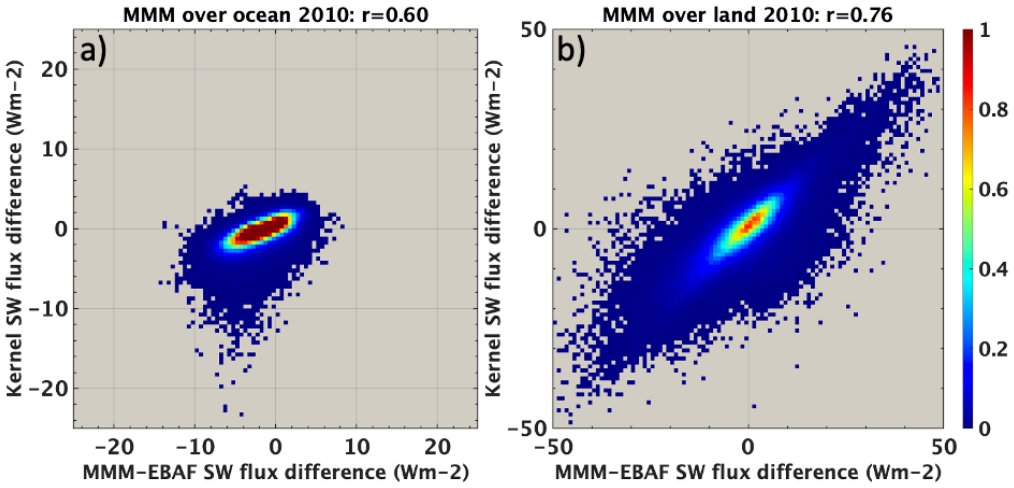

Figure 6. Relationship between SW flux biases of multi-model mean relative to CERES EBAF and kernel based flux biases using all grid boxes over ocean between 50°S–40°N using all 12 months of 2010 (a), and using all grid boxes over land between 60°S–60°N using all 12 months of 2010 (b). Over land, kernel SW flux biases are $\Delta F_{AOD} + \Delta F_{\alpha}$. Over ocean, kernel SW flux biases are ΔF_{AOD}. Color bar shows the relative number density.

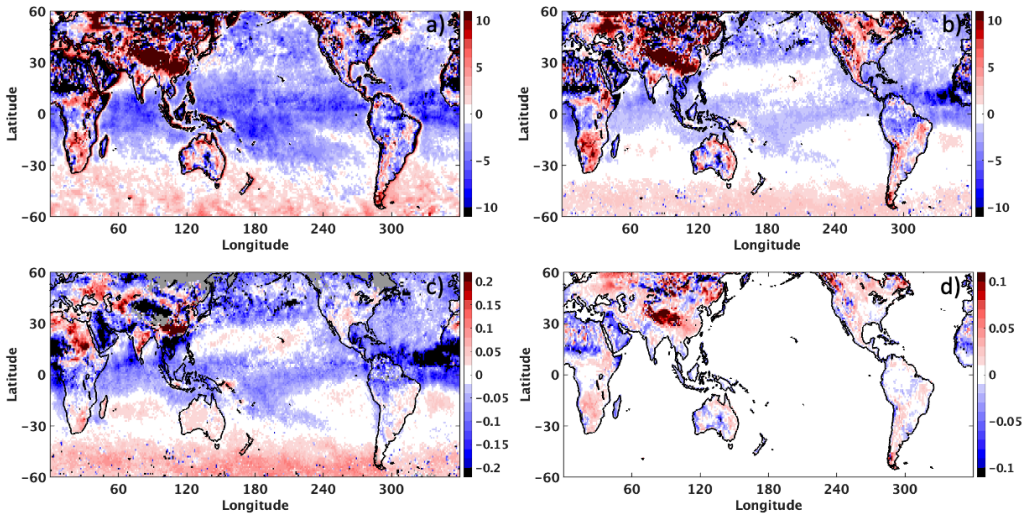

Figure 7. TOA clear-sky reflected SW flux biases of multi-model mean (MMM) (a), and TOA clear-sky reflected SW flux biases due to AOD biases (shown in c) and surface albedo biases (shown in d) calculated from radiative kernels (b), aerosol optical depth biases of MMM relative to MODIS (c), and land surface albedo biases of MMM relative to MODIS (d) for April 2010.

Over ocean, the spatial distribution of ΔF corresponds well with ΔF_{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. 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 ΔF and ΔF_{AOD}+ΔF_{α} 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 ΔF 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. Only by comparing the clear-sky TOA SW fluxes from CERES with those from the models confirms that the models indeed underestimate the dust outflow there which result in low biases of SW flux. One may argue that overestimating dust

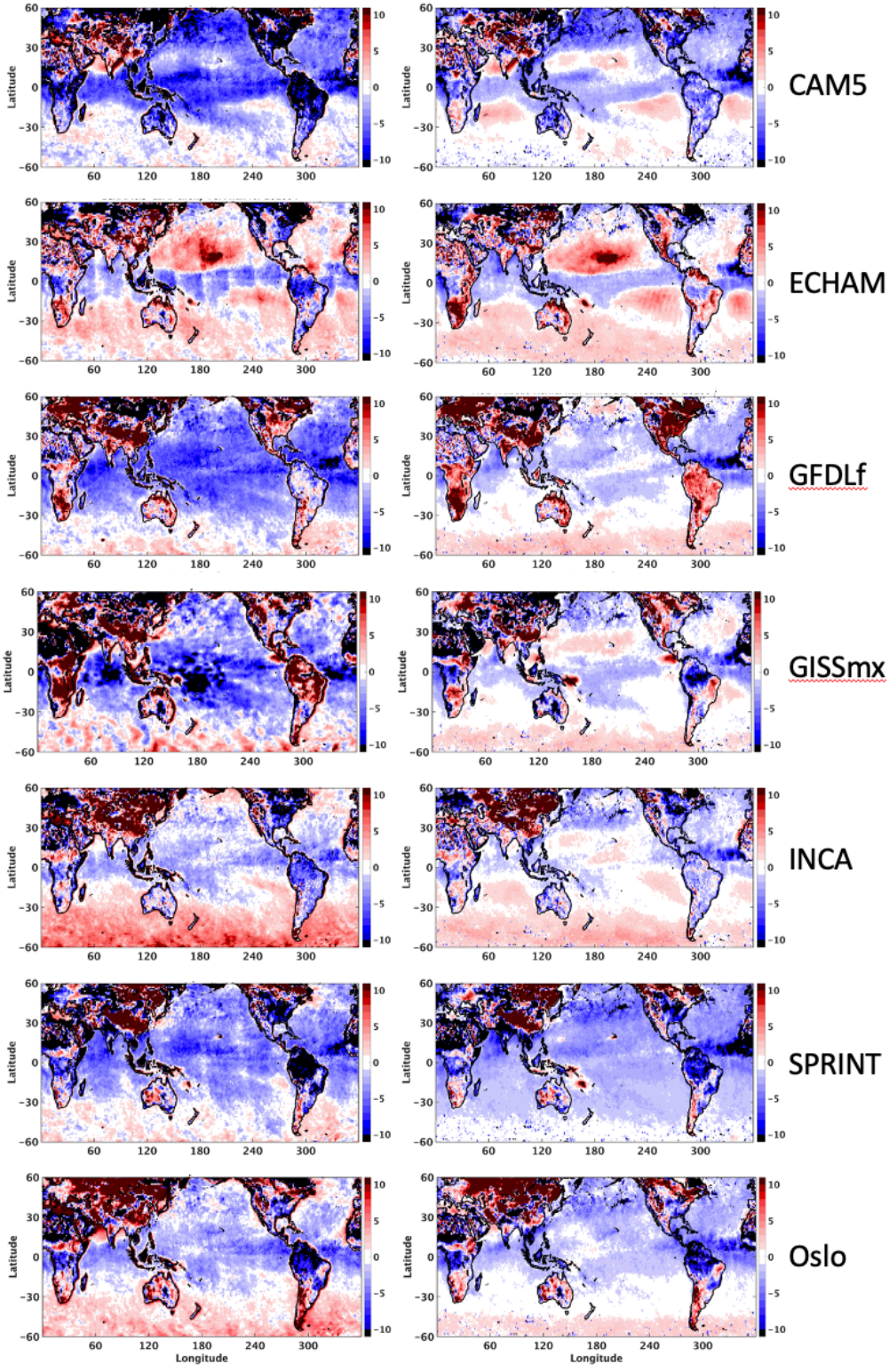

Figure 8. TOA clear-sky reflected SW flux biases fo AeroCom models relative to CERES EBAF for April 2010 (left panels). Right panels use radiative kernel calculations to determine the flux biases associated with AOD and surface albedo biases.

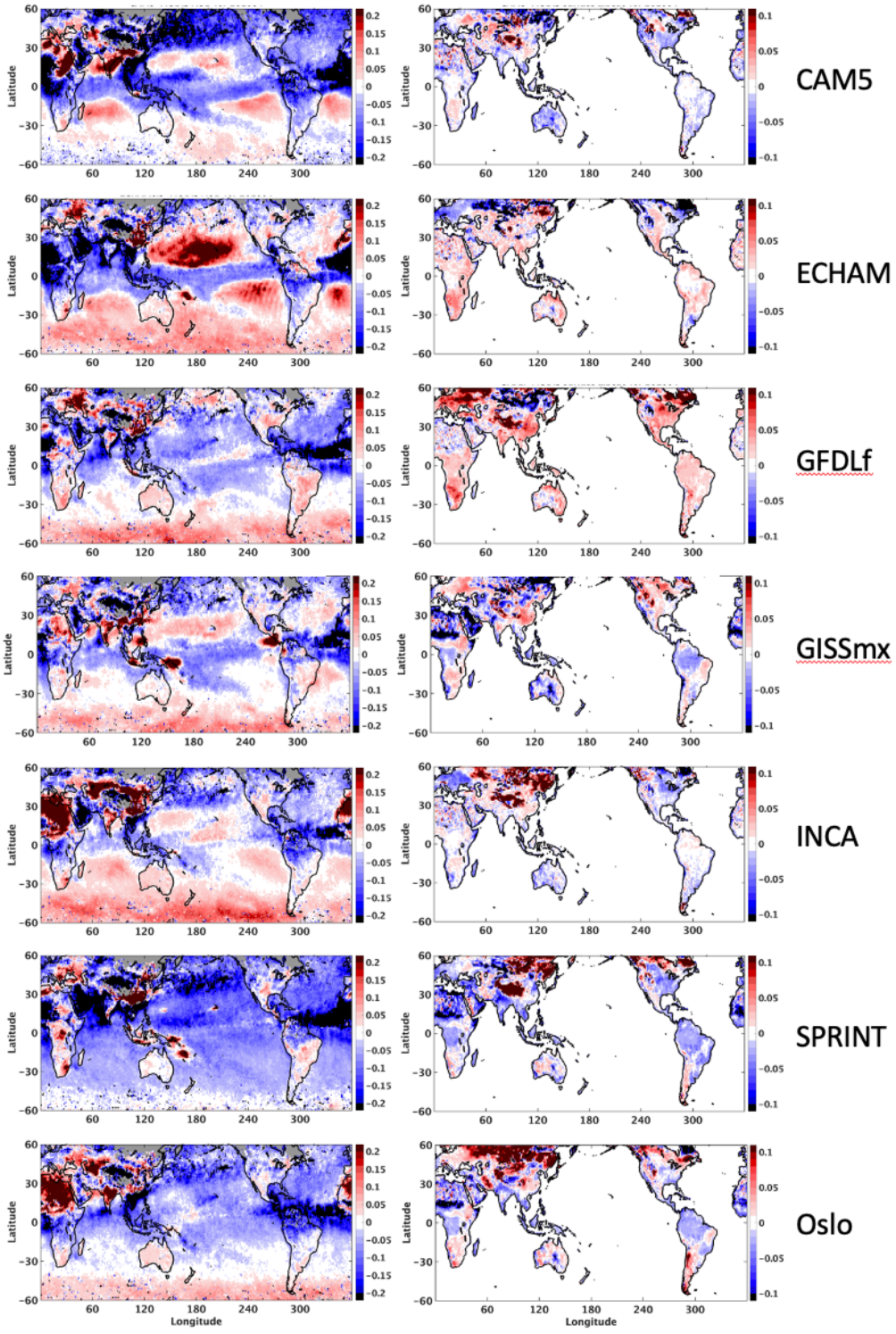

Figure 9. AOD biases (left panels) and land surface albedo biases (right panels) of AeroCom models relative to MODIS retrievals for April 2010.

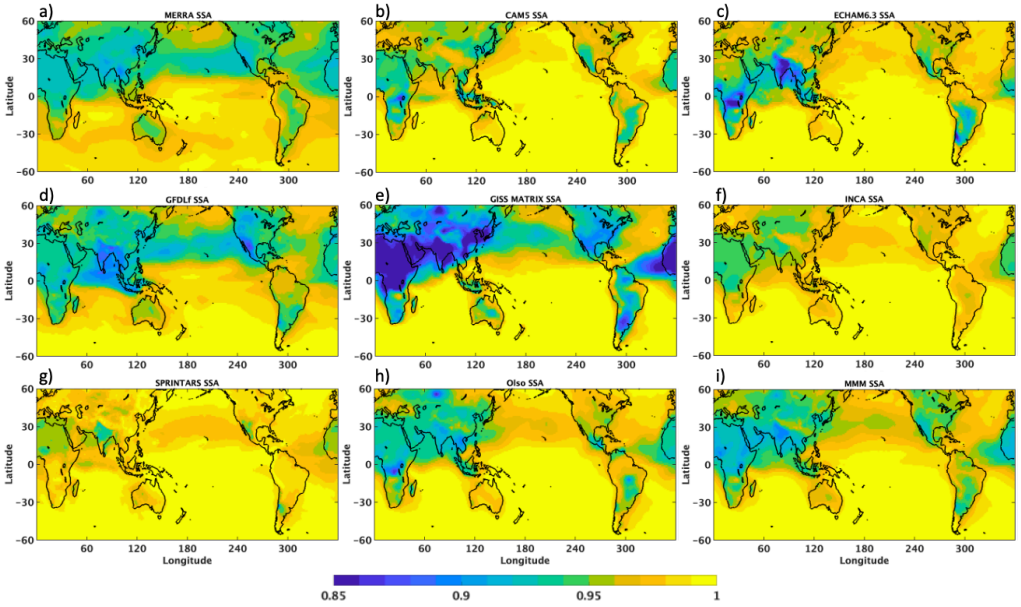

Figure 10. Single-scattering albedo for (a) MERRA-2, (b) CAM5, (c) ECHAM, (d) GFDLf, (e) GISS MATRIX, (f) INCA, (g) SPRINTARS, (h) Oslo, and (i) multi-model mean for April. MERRA-2 result is for 2007, and all others are for 2010.

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 ΔF_χ. 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 (Figure 10). Thus the AOD kernels over ocean are less sensitive to changes in AOD than if an AeroCom model (less absorbing) is used as the base state. Over land, single scattering albedo (SSA) from many AeroCom models agree reasonable 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 ΔF_χ. 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 its aerosols are very absorbing) 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 load-

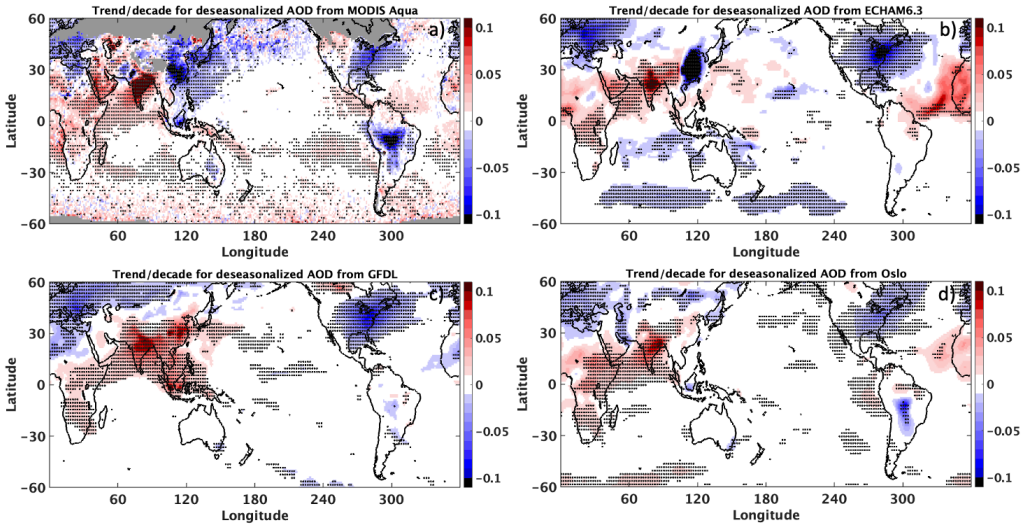

Figure 11. Aerosol optical depth trends (per decade) calculated using satellite retrievals from Aqua MODIS (a) and using ECHAM-HAM (b), GFDL (c), and Oslo (d) model simulations. Data from July 2002 to December 2017 are used here. All models use emissions from CEDS until 2014, but use different emissions between 2015 and 2017.

ing and land surface albedo in the models then adjusting the aerosol composition, size distribution, effective refractive indices, and aerosol hygroscopic 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

ECHAM-HAM, GFDL (AMIP run), and Oslo models also provide historical simulations which can be 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. However, each model handles the emissions between 2015 and 2017 differently. ECHAM-HAM used the CMIP6 Shared Socioeconomic Pathways (SSP) 3-7.0 emission scenario (Gidden et al., 2019). Oslo used the SSP2-4.5 middle of the road emission scenario (Fricko et al., 2017). GFDL used CEDS anthropogenic and biomass burning emissions of 2014 for 2015-2017, and the dust and sea-salt emissions are calculated using the actual wind speed produced by the model. 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 Peninsular, 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 2015-2017 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 direct aerosol radiative effect (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):

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

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

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

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:

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

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 from CERES EBAF and the surface albedo datasets, and from model simulations. Both observation-based and model simulations show increasing trends over eastern United States and the Atlantic Ocean (DARE becomes less negative), and decreasing trends over central Africa, India, and Indian Ocean (DARE becomes more negative). These DARE trends are consistent with the AOD trends over these regions. The increasing trend off the coast of eastern China in EBAF data is largely absent in the models, as is in the AOD trend. ECHAM-HAM model shows a strong increasing trend over eastern China, whereas trends from EBAF, GFDL, and Oslo are very muted. The observational-based DARE trend and AOD trends show good agreement over many regions, which adds confidence in both products.

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. 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 aerosol optical depth. 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

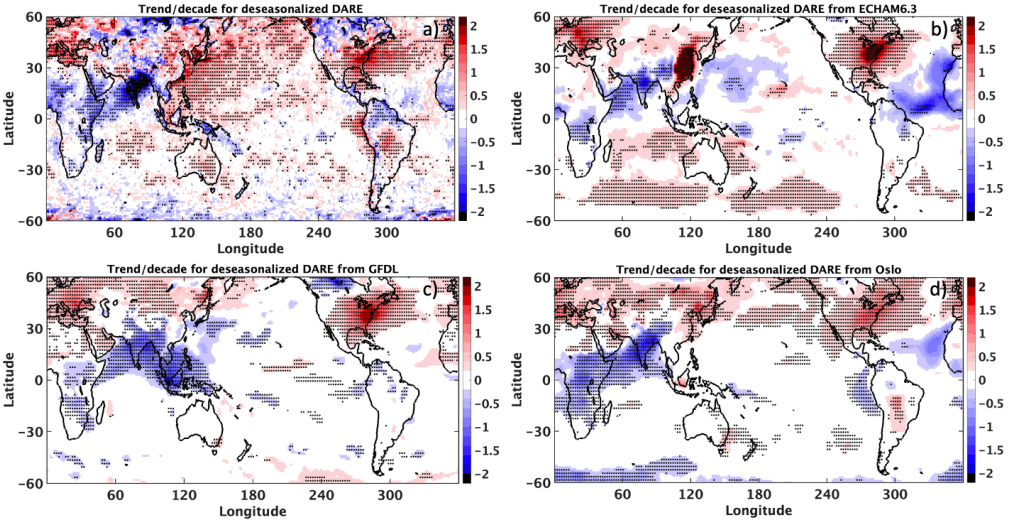

Figure 12. TOA SW direct aerosol radiative effect trends (Wm^{-2} per decade) calculated using CERES EBAF (a) and using ECHAM-HAM (b), GFDL (c), and Oslo (d) model simulations. Data from July 2002 to December 2017 are used here.

Table 4. Aerosol direct radiative effect trends (Wm^{-2} per decade) over China, India, and USA. Only land areas in the indicated latitude and longitude ranges are included in the regional trend calculation.

	China (25-40°N, 110-120°)	India (7-22°N, 70-90°)	USA (30-42°N, 265-283°)
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
USA	0.2±0.2	-1.1±0.3	1.0±0.2

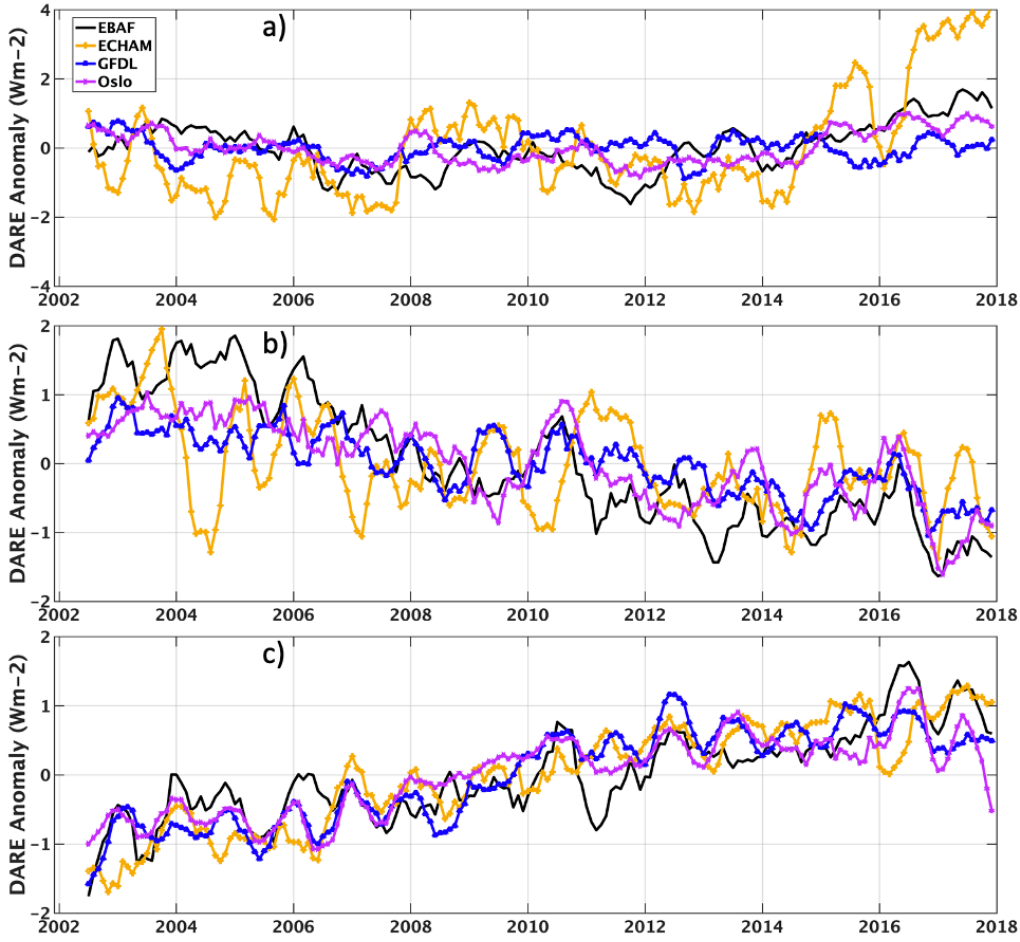

Figure 13. Deseasonalized anomalies (Wm^{-2}) of TOA SW direct aerosol radiative effect over eastern China (a), India (b), and eastern USA (c) calculated using CERES EBAF (black lines), ECHAM-HAM (orange line), GFDL (blue lines), and Oslo (purple) model simulations. Data from July 2002 to December 2017 are used here.

developed from MERRA-2 reanalysis to attribute their contributions to SW flux biases. Over ocean, the 60S-60N mean TOA SW flux bias in MMM is reduced by 25% after accounting for the contribution from AOD bias. Over land, the 60S-60N 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, Indian Ocean, Arabian Peninsular, 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 dataset.

Using independently derived satellite datasets (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. Implementing these changes will improve the global and regional SW flux agreement between models and satellite observations, and reduce the diversity among the models.

Appendix A

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

A1 CAM5

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 10000 nm in dry diameter) and 8 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 and longwave 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 et al., 2018; Matsui & Moteki, 2020).

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 shortwave and 14 longwave 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 (sulphate, 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 SO_2 (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 sulphate. 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-hour relaxation time, towards the NCEP-NCAR re-analysis (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 hour for shortwave and 3 hour for longwave. The shortwave 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_2O, CO_2, and O_2 formulations, and shortwave absorption by CH_4 and N_2O. In addition, the effects of the shortwave 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 (S. 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

INCA (INteraction with Chemistry and Aerosols) is a chemistry-aerosol model coupled to a land surface and a dynamical model. INCA simulates dust, sea salt, black carbon (BC), NO3, SO4, SO2, 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 sulphate (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 ($\sim$0.5625°$\times$ 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; Søvde et al., 2012). The model is run in a 2.25° $\times$ 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 ref-

erences 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_{α} 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.

Acknowledgments

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 GES DISC. The V6 MODIS Bidirectional Reflectance Distribution Func- tion (BRDF)/albedo products (MCD43C1) were obtained from the Land Processes Distributed Active Archive Center (LP DACC) through https://lpdaac.usgs.gov/products/mcd43c1v006/.

This research has been supported by NASA CERES project. We would like to thank Dr. John P. Dunne for providing constructive comments on on earlier version of the paper. Ragnhild B.Skeie and Gunnar Myhre were funded through the Norwegian Research Council (grant no. 250573) and the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement 820829 (CONSTRAIN). Hitoshi Matsui was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Japan Society for the Promotion of Science (MEXT/JSPS) KAKENHI Grant Numbers JP17H04709, JP16H01770, JP19H04253, JP19H05699, JP19KK0265, JP20H00196, and JP20H00638, MEXT Arctic Challenge for Sustainability (ArCS, JPMXD1300000000) and ArCS-II (JPMXD1420318865) projects, and the Environment Research and Technology Development Fund 2-2003 (JPMEERF20202003) of the Environmental Restoration and Conservation Agency. David Neubauer acknowledges funding from the European Union's Horizon 2020 Research and Innovation programme, project FORCeS, under grant agreement no. 821205. The ECHAM-HAMMOZ model is developed by a consortium composed of ETH Zurich, Max-Planck-Institut für Meteorologie, Forschungszentrum Jülich, University of Oxford, the Finnish Meteorological Institute, and the Leibniz Institute for Tropospheric Research and managed by the Center for Climate Systems Modeling (C2SM) at ETH Zurich. Toshihiko Takemura is supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant no. JP19H05669), the Environment Research and Technology Development Fund (S-20) of the Environmental Restoration and Conservation Agency, Japan, and the NEC SX supercomputer system of the National Institute for Environmental Studies, Japan. Susanne E. Bauer and Kostas Tsigaridis acknowledge funding from the NASAs Modeling, Analysis and Prediction Program. Resources supporting this work were provided by the NASA High?End

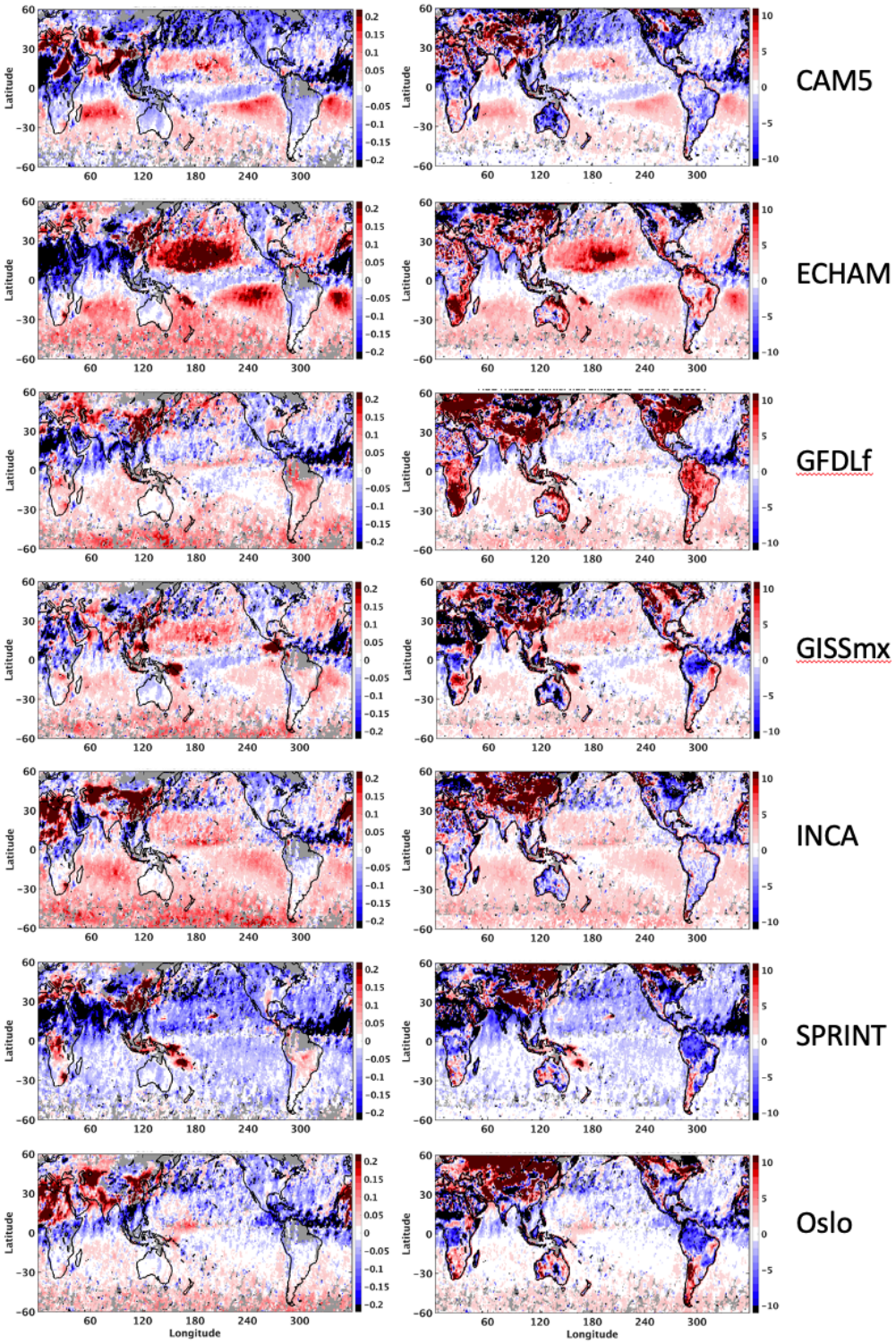

Figure B1. The monthly mean aerosol optical depth biases of AeroCom models relative to MISR retrieval (left panels), TOA SW flux biases due to AOD biases (models – MISR) and land surface albedo biases (models – MODIS) calculated from their respective radiative kernels (right panels) for April 2010.

Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center.

References

Bauer, S., Wright, D., Koch, D., Lewis, E. R., McGraw, R., Chang, L.-S., ... Ruedy, R. (2008). MATRIX (Multiconfiguration Aerosol TRacker of mIXing state): An aerosol microphysical module for global atmospheric models. *Atmos. Chem. Phys.*, *8*, 6603-6035. doi: 10.5194/acp-8-6003-2008

Bauer, S. E., Tsigaridis, K., Faluvegi, G., Kelley, M., Lo, K. K., Miller, R. L., ... Wu, J. (2020). Historical (1850-2014) Aerosol Evolution and Role on Climate Forcing Using the GISS ModelE2.1 Contribution to CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*(8). doi: 10.1029/2019MS001978

Bellouin, N., Boucher, O., Haywood, J., & Reddy, M. S. (2005). Global estimate of aerosol direct radiative forcing from satellite measurements. *Nature*, *438*(22), 1138-1141.

Bellouin, N., Jones, A., Haywood, J., & Christopher, S. A. (2008). Updated estimate of aerosol direct radiative forcing from satellite observations and comparison against the Hadley centre climate model. *J. Geophys. Res.*, *113*(D10205). doi: 10.1029/2007JD009385

Bohren, C., & Huffman, D. (1998). *Absorption and scattering of light by small particles*. Hoboken, N. J.: John Wiley.

Carrer, D., Roujean, J. L., & Meurey, C. (2010). Comparing operational MSG/SEVIRI Land Surface albedo products from Land SAF with ground measurements and MODIS. *IEEE Trans. Geosci. Remote Sens.*. doi: 10.1109/TGRS.2009.2034530

Cescatti, A., Marcolla, B., Santhana Vannan, S. K., Pan, J. Y., Roman, M. O., Yang, X., ... Schaaf, C. B. (2012). Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network. *Remote Sens. Environ.*, *121*, 323-334. doi: 10.1016/j.rse.2012.02.019

d'Almeida, G. A., Koepke, P., & Shettle, E. P. (1991). *Atmospheric aerosols - global climatology and radiative characteristics*. Hampton, VA: A. Deepak Publishing.

Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Wielicki, B. A., ... Sun, M. (2013). Geostationary enhanced temporal interpolation for CERES flux products. *J. Atmos. Oceanic Technol.*, *30*, 1072-1090. doi: 10.1175/JTECH-D-12-00136.1

Durack, P. J., & Taylor, K. E. (2018). *PCMDI AMIP SST and sea-ice boundary conditions version 1.1.4.* Retrieved 01/10/2021, from `EarthSystemGridFederation` doi: 10.22033/ESGF/input4MIPs.2204

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. Retrieved from `https://www.geosci-model-dev.net/9/1937/2016/` doi: 10.5194/gmd-9-1937-2016

Freidenreich, S. M., & Ramaswamy, V. (2005). Refinement of the Geophysical Fluid Dynamics Laboratory solar benchmark computations and an improved parameterization for climate models. *J. Geophys. Res.*, *110*(D17). doi: 10.1029/2004JD005471

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., ... Riahi, K. (2017). The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century. *Global Environmental Change*, *42*, 251-267. doi: 10.1016/j.gloenvcha.2016.06.004

Garay, M. J., Witek, W. L., Kahn, R. A., Seidel, F. C., Limbacher, J. A., Bull, M. A., ... Yu, Y. (2020). Introducing the 4.4 km spatial resolution Multi-

Angle Imaging SpectroRadiometer (MISR) aerosol product. *Atmos. Meas. Tech.*, *13*, 593-628. doi: 10.5194/amt-13-593-2020

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., ... Takahashi, K. (2019). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geoscientific Model Development*, *12*(4), 1443-1475. doi: 10.5194/gmd-12-1443-2019

Ginoux, P., Chin, M., Tegen, I., Prospero, J., Holben, B., Dubovik, O., & Lin, S.-J. (2001). Sources and global distributions of dust aerosols simulated with the GOCART model. *J. Geophys. Res.*, *106*, 20255–20273.

Gliß, J., Mortier, A., Schulz, M., Andrews, E., Balkanski, Y., Bauer, S. E., ... Tsyro, S. G. (2021). AeroCom phase III multi-model evaluation of the aerosol life cycle and optical properties using ground- and space-based remote sensing as well as surface in situ observations. *Atmos. Chem. Phys.*, *21*, 87-128. doi: 10.5194/acp-21-87-2021

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., ... Zhang, Q. (2018). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, *11*(1), 369-408. doi: 10.5194/gmd-11-369-2018

Holben, B., Tanre, D., Smirnov, A., Eck, T., Slutsker, I., Abuhassan, N., ... Zibordi, G. (2001). An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET. *J. Geophys. Res.*, *106*(D11), 12067-12097.

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

Jin, Z., Charlock, T. P., Jr., W. L. S., & Rutledge, K. (2004). A parameterization of ocean surface albedo. *Geophys. Res. Lett.*, *31*(22). doi: 10.1029/2004GL021180

Johnson, G. C., Lyman, J. M., & Loeb, N. G. (2016). Improving estimates of Earth's energy imbalance. *Nat. Climate Change*, *6*, 639-640. doi: 10.1038/nclimate3043

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... others (1996). The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, *77*, 437-472.

Kinne, S., Schulz, M., Textor, C., Guibert, S., & coauthors. (2006). An AeroCom initial assessment - optical properties in aerosol component modules of global models. *Atmos. Chem. Phys.*, *6*, 1815-1834.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., & Hsu, N. C. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.*, *6*, 2989-3034. doi: 10.5194/amt-6-2989-2013

Li, X., & Strahler, A. H. (1992). Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. *IEEE Trans. Geosci. Remote Sensing*, *30*, 276-292.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J., ... Kato, S. (2018). Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *J. Climate*, *31*, 895-918. doi: 10.1175/JCLI-D-17-0208.1

Loeb, N. G., & Manalo-Smith, N. (2005). Top-of-atmosphere direct radiative effect of aerosols over global oceans from merged CERES and MODIS observations. *J. Climate*, *18*, 3506-3526.

Loeb, N. G., Manalo-Smith, N., Su, W., Shankar, M., & Thomas, S. (2016). CERES top-of-atmosphere Earth radiation budget climate data record: Accounting for in-orbit changes in instrument calibration. *Remote Sens.*, *8*(182). doi:

10.3390/rs8030182

Loeb, N. G., Su, W., Bellouin, N., & Ming, Y. (2021). Changes in clear-sky shortwave aerosol direct radiative effects since 2002. *J. Geophys. Res.*, *126*(e2020JD034090). doi: 10.1029/2020JD034090

Loeb, N. G., Wang, H., Allan, R. P., Andrews, T., Armour, K., Cole, J. N. S., ... Wyser, K. (2020). New Generation of Climate Models Track Recent Unprecedented Changes in Earth's Radiation Budget Observed by CERES. *Geophys. Res. Lett.*, *47*(5). doi: 10.1029/2019GL086705

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., ... Wong, T. (2009). Towards optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Climate*, *22*, 748-766. doi: 10.1175/2008JCLI2637.1

Lucht, W., Schaaf, C. B., & Strahler, A. H. (2000). An anogorithm for the retrieval of albedo from space using semiempirical BRDF models. *IEEE Trans. Geosci. Remote Sensing*, *38*, 977-998.

Lund, M. T., Myhre, G., Haslerud, A. S., Skeie, R. B., Griesfeller, J., Platt, S. M., ... Schulz, M. (2018). Concentrations and radiative forcing of anthropogenic aerosols from 1750 to 2014 simulated with the Oslo CTM3 and CEDS emission inventory. *Geoscientific Model Development*, *11*(12), 4909-4931. doi: 10.5194/gmd-11-4909-201

Martensson, E. M., Nilsson, E. D., de Leeuw, G., Cohen, L. H., & Hansson, H.-C. (2003). Laboratory simulations and parameterization of the primary marine aerosol production. *J. Geophys. Res.*, *108*. doi: 10.1029/2002JD002263

Matsui, H. (2017). Development of a global aerosol model using a two-dimensional sectional method: 1. model design. *J. Adv. Model. Earth Syst.*, *9*, 1921-1947. doi: 10.1002/2017MS000936

Matsui, H., Hamilton, D. S., & Mahowald, N. M. (2018). Black carbon radiative effects highly sensitive to emitted particle size when resolving mixing-state diversity. *Nat. Commun.*, *9*(3446). doi: 10.1038/s41467-018-05635-1

Matsui, H., & Mahowald, N. (2017). Development of a global aerosol model using a two-dimensional sectional method: 2. evaluation and sensitivity simulations. *J. Adv. Model. Earth Syst.*, *9*, 1887-1920. doi: 10.1002/2017MS000937

Matsui, H., & Moteki, N. (2020). High sensitivity of arctic black carbon radiative effects to subgrid vertical velocity in aerosol activation. *Geophys. Res. Lett.*, *47*(e2020GL088978). doi: 10.1029/2020GL088978

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... et al. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. *J. Adv. Model. Earth Syst.*, *11*(4), 998-1038. doi: 10.1029/2018MS001400

Miller, R. L., Cakmur, R. V., Perlwitz, J., Geogdzhayev, I. V., Ginoux, P., Koch, D., ... Tegen, I. (2006). Mineral dust aerosols in the NASA Goddard Institute for Space Sciences ModelE atmospheric general circulation mode. *J. Geophys. Res.*, *111*(D6), D06208. doi: 10.1029/2005JD005796

Monahan, E. C., Spiel, D. E., & Davidson, K. L. (1986). *A model of marine aerosol generation via whitecaps and wave disruption bt – oceanic whitecaps: And their role in air-sea exchange processes,*. Springer. Dordrecht, Netherlands. doi: 10.1007/978-34594-009-4668-2_16

Muench, S., & Lohmann, U. (2020). Developing a Cloud Scheme With Prognostic Cloud Fraction and Two Moment Microphysics for ECHAM-HAM. *J. Adv. Model. Earth Syst.*, *12*(8), e2019MS001824. doi: 10.1029/2019MS001824

Myhre, G., Bellouin, N., Berglen, T. F., Berntsen, T. K., Boucher, O., Grini, A., ... Tanre, D. (2007). Comparison of the radiative properties and direct radiative effect of aerosols from a global aerosol model and remote sensing data over ocean. *Tellus B*, *59*(1), 115-129. doi: 10.1111/j.1600-0889.2006.00226.x

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K.,

... Zhou, C. (2013). Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations. *Atmos. Chem. Phys.*, *4*(1853-1877). doi: 10.5194/acp-13-1853-2013

O'Donnell, D., Tsigaridis, K., & Feichter, J. (2011). Estimating the direct and indirect effects of secondary organic aerosols using ECHAM5-HAM. *Atmos. Chem. Phys.*, *11*, 8635-8659. doi: 10.5194/acp-11-8635-2011

Paulot, F., Paynter, D., Ginoux, P., Naik, V., & Horowitz, L. W. (2018). Changes in the aerosol direct radiative forcing from 2001 to 2015: observational constraints and regional mechanisms. *Atmos. Chem. Phys.*, *18*, 13265-13281. doi: 10.5194/acp-18-13265-2018

Pincus, R., Batstone, C. P., Hofmann, R. J. P., Taylor, K. E., & Glecker, P. J. (2008). Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *J. Geophys. Res.*, *113*(D14209), D14209. doi: 10.1029/2007JD009334

Pincus, R., & Stevens, B. (2013). Paths to accuracy for radiation parameterizations in atmospheric models. *J. Adv. Model. Earth Syst.*, *5*, 225-233. doi: 10.1002/jame.20027

Randles, C. A., Kinne, S., Myhre, G., Schulz, M., Stier, P., Fischer, J., ... Lu, P. (2013). Intercomparison of shortwave radiative transfer schemes in global aerosol modeling: results from the AeroCom Radiative Transfer Experiment. *Atmos. Chem. Phys.*, *13*(5), 2347-2379. doi: 10.5194/acp-13-2347-2013

Remer, L. A., & Kaufman, Y. J. (2006). Aerosol direct radiative effect at the top of the atmophere over cloud free ocean derived from four years of MODIS data. *Atmos. Chem. Phys.*, *6*, 237-253.

Roman, M. O., Gatebe, C. K., Shuai, Y., Wang, Z., Gao, F., Masek, J. G., ... Schaaf, C. B. (2013). Use of in situ and airborne multiangle data to assess MODIS- and Landsat-based estimates of directional reflectance and albedo. *IEEE Trans. Geosci. Remote Sensing*, *51*, 1393-1404. doi: 10.1109/TGRS.2013.2243457

Rutan, D., Rose, F., Roman, M., Manalo-Smith, N., Schaaf, C. B., & Charlock, T. (2009). Development and assessment of broadband surface albedo from Clous and the Earth's Radiant Energy System Clouds and Radiaiton Swath data product. *J. Geophys. Res.*, *114*(D08125). doi: 10.1029/2008JD010669

Sayer, A. M., Munchak, L. A., Hsu, N. C., Levy, R. C., Bettenhausen, C., & Jeong, M. (2014). MODIS Collection 6 aerosol products: Comeparison between Aqua's e-Deep Blue, Dark Target, and "merged" data sets, and usage recommendations. *J. Geophys. Res.*, *119*, 13965-13989. doi: 10.1002/2014JD022453

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., ... Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1), 135-148. doi: 10.1016/S0034-4257(02)00091-3

Schulz, M., Cozic, A., & Szopa, S. (2009). Lmdzt-inca dust fore-cast model developments and associated validation efforts. *IOP Conference Series: Earth and Environmental Science*, *7*(12014). doi: 10.1088/1755-1307/7/1/012014

Schulz, M., Textor, C., Kinne, S., Balkanski, Y., & coauthors. (2006). Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. *Atmos. Chem. Phys.*, *6*, 5225-5246.

Schutgens, N., Sayer, A. M., Heckel, A., Hsu, C., Jethva, H., de Leeuw, G., ... Stier, P. (2020). An AeroCom–AeroSat study: intercomparison of satellite AOD datasets for aerosol model evaluation. *Atmos. Chem. Phys.*, *20*(21), 12431-12457. doi: 10.5194/acp-20-12431-2020

Sekiguchi, M., & Nakajima, T. (2008). A k-distribution-based radiation code and its computational optimization for an atmospheric general circulation model. *J. Quant. Spect. Rad. Trans.*, *109*, 2779-2793. doi: 10.1016/j.jqsrt.2008.07.013

Søvde, O. A., Prather, M. J., Isaksen, I. S. A., Berntsen, T. K., Stordal, F., Zhu, X.,

... Hsu, J. (2012). The chemical transport model Oslo CTM3. *Geoscientific Model Development*, *5*(6), 1441-1469. doi: 10.5194/gmd-5-1441-2012

Stier, P., Schutgens, N. A. J., Bellouin, N., Bian, H., Boucher, O., Chin, M., ... Zhou, C. (2013). Host model uncertainties in aerosol radiative forcing estimates: results from the AeroCom Prescribed intercomparison study. *Atmos. Chem. Phys.*, *13*(6), 3245-3270. doi: 10.5194/acp-13-3245-2013

Su, W., Bodas-Salcedo, A., Xu, K.-M., & Charlock, T. P. (2010). Comparison of the tropical radiative flux and cloud radiative effect profiles in a climate model with Clouds and the Earth's Radiant Energy System (CERES) data. *J. Geophys. Res.*, *115*(D01105), D01105. doi: 10.1029/2009JD012490

Su, W., Loeb, N. G., Schuster, G. L., Chin, M., & Rose, F. G. (2013). Global all-sky shortwave direct radiative forcing of anthropogenic aerosols from combined satellite observations and GOCART simulations. *J. Geophys. Res.*, *118*, 1-15. doi: 10.1029/2012JD018294

Taberner, M., Pinty, B., Govaerts, Y., Liang, S., Verstraete, M. M., Gobron, N., & Widlowski, J. (2010). Comparison of MISR and MODIS land surface albedos: Methodology. *J. Geophys. Res.*, *115*(D05101). doi: 10.1029/2009JD012665

Takemura, T., Egashira, M., Matsuzawa, K., Ichijo, H., O'ishi, R., & Abe-Ouchi, A. (2009). A simulation of the global distribution and radiative forcing of soil dust aerosols at the last glacial maximum. *Atmos. Chem. Phys.*, *9*(9), 3061-3073. doi: 10.5194/acp-9-3061-2009

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., & Nakajima, T. (2005). Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model. *J. Geophys. Res.*, *110*(D2). doi: 10.1029/2004JD005029

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., ... Kimoto, M. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, *12*(7), 2727-2765. doi: 10.5194/gmd-12-2727-2019

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., ... Lohmann, U. (2019). The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 1: Aerosol evaluation. *Geosci. Model Dev.*, *12*, 1643-1677. doi: 10.5194/gmd-12-1643-2019

Thorsen, T. J., Ferrare, R. A., Kato, S., & Winker, D. M. (2020). Aerosol direct radiative effect sensitivity analysis. *J. Climate*, *33*, 6119-6139. doi: 10.1175/JCLI-D-19-0669.1

Tsigaridis, K., Koch, D., & Menon, S. (2013). Uncertainties and importance of sea spray composition on aerosol direct and indirect effects. *J. Geophys. Res.*, *118*(1), 220-235. doi: 10.1029/2012JD018165

Wang, H., & Su, W. (2013). Evaluating and understanding top of the atmosphere cloud radiative effects in Intergovernmental Panel on Climate Change (IPCC) fifth assessment report (AR5) cloupled model intercomparison project phase 5 (CMIP5) models using satellite observations. *J. Geophys. Res.*, *118*, 1-17. doi: 10.1029/2012JD018619

Wang, R., Balkanski, O., Y.and Boucher, Ciais, P., Schuster, G. L., Chevallier, F., Samset, B. H., ... Valari, M. (2016). Estimation of global black carbon direct radiative forcing and its uncertainty constrained by observations. *J. Geophys. Res.*, *121*, 5948-5971. doi: 10.1002/2015jd024326

Wang, Z., Lin, L., Xu, Y., Che, H., Zhang, X., Zhang, H., ... Xie, B. (2021). Incorrect asian aerosols affecting the attribution and projection of regional climate change in cmip6 models. *npj Climate and Atmospheric Science*, *4*(2), 1. doi: 10.1038/s41612-020-00159-2

Wang, Z., Schaaf, C. B., Chopping, M. J., Strahler, A. H., Wang, J., Roman, M. O., ... Shuai, Y. (2012). Evaluation of Moderate-resolution Imaging Spectroradiometer (MODIS) snow albedo product (MCD43A) over tundra. *Remote*

Sens. Environ., *117*, 264-280. doi: 10.1016/j.rse.2011.10.002

Wang, Z., Schaaf, C. B., Strahler, A. H., Chopping, M. J., Roman, M. O., Shuai, Y., ... Fitzjarrald, D. R. (2014). Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow- covered period. *Remote Sens. Environ.*, *140*, 60-77. doi: 10.1016/j.rse.2013.08.025

Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Smith, G. L., & Cooper, J. E. (1996). Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System experiment. *Bull. Amer. Meteor. Soc.*, *77*, 853-868.

Wild, M., Folini, D., Schar, C., Loeb, N. G., Dutton, E. G., & Konig-Langlo, G. (2013). The global energy balance from a surface perspective. *Clim. Dyn.*, *40*, 3107-3134. doi: 10.1007/s00382-012-1569-8

Yu, H., Kaufman, Y. J., Chin, M., Feingold, G., Remer, L. A., Anderson, T. L., & coauthors. (2006). A review of measurement-based assessments of the aerosol direct radiative effect and forcing. *Atmos. Chem. Phys.*, *6*, 613-666.

Zanatta, M., Gysel, M., Bukowiecki, T., N.and Müller, Weingartner, E., Areskoug, H., Fiebig, M., ... Laj, P. (2016). A European aerosol phenomenology-5: Climatology of black carbon optical properties at 9 regional background sites across Europe. *Atmos. Environ.*, *145*, 1352-2310. doi: 10.1016/j.atmosenv.2016.09.035

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., ... Xiang, B. (2018). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 1. Simulation Characteristics With Prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*(3), 691-734. doi: 10.1002/2017MS001208