The GFDL Global Atmospheric Chemistry-Climate Model AM4.1: Model Description and Simulation Characteristics

Larry W. Horowitz[1], Vaishali Naik[1], Fabien Paulot[1], Paul A. Ginoux[1], John P. Dunne[1], Jingqiu Mao[2], Jordan Schnell[4], Xi Chen[3], Jian He[3], Jasmin G. John[1], Meiyun Lin[3], Pu Lin[5], Sergey Malyshev[1], David Paynter[1], Elena Shevliakova[1], Ming Zhao[1]

[1]NOAA Geophysical Fluid Dynamics Laboratory, Princeton NJ USA

[2]Department of Chemistry and Biochemistry & Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK 99775, USA

[3]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton NJ USA

[4]Department of Earth and Planetary Sciences and Institute for Sustainability and Energy, Northwestern University, Evanston, IL, USA

[5]Cooperative Institute for Modeling the Earth System, Princeton University, Princeton NJ 08540

Corresponding author: Larry W. Horowitz (Larry.Horowitz@noaa.gov)

Key Points:

- A new atmospheric chemistry-climate model (AM4.1) has been developed for the Geophysical Fluid Dynamics Laboratory (GFDL) 4th-generation suite.
- AM4.1 includes an advanced dynamical core and physical parameterizations, with enhanced vertical resolution and revised aerosol and chemistry interactions.
- The AM4.1 model exhibits substantially enhanced fidelity compared to previous-generation GFDL atmospheric models.

Abstract

We describe the baseline model configuration and simulation characteristics of GFDL's Atmosphere Model version 4.1 (AM4.1), which builds on developments at GFDL over 2013–2018 for coupled carbon-chemistry-climate simulation as part of the sixth phase of the Coupled Model Intercomparison Project. In contrast with GFDL's AM4.0 development effort, which focused on physical and aerosol interactions and which is used as the atmospheric component of CM4.0, AM4.1 focuses on comprehensiveness of Earth system interactions. Key features of this model include doubled horizontal resolution of the atmosphere (~200 km to ~100 km) with revised dynamics and physics from GFDL's previous-generation AM3 atmospheric chemistry-climate model. AM4.1 features improved representation of atmospheric chemical composition, including aerosol and aerosol precursor emissions, key land-atmosphere interactions, comprehensive land-atmosphere-ocean cycling of dust and iron, and interactive ocean-atmosphere cycling of reactive nitrogen. AM4.1 provides vast improvements in fidelity over AM3, captures most of AM4.0's baseline simulations characteristics and notably improves on AM4.0 in the representation of aerosols over the Southern Ocean, India, and China—even with its interactive chemistry representation—and in its manifestation of sudden stratospheric warmings in the coldest months. Distributions of reactive nitrogen and sulfur species, carbon monoxide, and ozone are all substantially improved over AM3. Fidelity concerns include degradation of upper atmosphere equatorial winds and of aerosols in some regions.

Plain Language Summary

GFDL has developed a coupled chemistry-climate Atmospheric Model (AM4.1) as part of its 4th-generation coupled model development activities. AM4.1 includes comprehensive chemistry for representing ozone and aerosols, and has been developed for use in chemistry and air quality applications, including advanced land-atmosphere-ocean coupling. With fidelity near to that of AM4.0, AM4.1 features vastly improved representation of climate mean patterns and variability from previous GFDL atmospheric chemistry-climate models.

1. Introduction

Atmospheric chemistry and composition are intrinsically coupled to the Earth's climate system. The pivotal role of chemistry-climate interactions in regulating climate has been recognized for over 30 years (Ramanathan et al., 1987). Advances in fundamental knowledge of atmospheric chemistry and availability of long-term measurements, combined with advances in computing, have provided an opportunity to enhance the comprehensiveness of our representation of

atmospheric chemistry-climate interactions. For the fifth phase of the Coupled Model Intercomparison Project (CMIP5), GFDL contributed its first coupled chemistry-climate model, CM3 (Donner et al., 2011), which allowed us to represent interactive aerosols and ozone, rather than prescribing concentrations from offline models (e.g., Horowitz et al., 2003). Much has been learned in the preceding years as to the strengths and weaknesses of CM3 chemistry, including the critical role of prognostic aerosol interactions (e.g., Naik et al., 2013a; Levy et al., 2013). As such, interactive aerosols were included in all of GFDL's 4th-generation atmospheric model development efforts, targeted for the sixth phase of the Coupled Model Intercomparison Project (CMIP6). The high computational cost of interactive atmospheric chemistry, however, was avoided in GFDL's recent CM4.0 development (Held et al., 2019) by prescribing ozone and other oxidants. A full interactive chemistry representation, along with a comprehensive carbon cycle, was reserved in this generation of GFDL models for Earth System Model development in ESM4.1 (Dunne et al., submitted). Thus, CM4 included a focus on ocean resolution, while ESM4.1 focused on a comprehensive representation of the Earth system.

The overall goal of AM4.1 development was to merge a suite of mostly parallel sets of updates and innovations into GFDL's 4th-generation atmospheric model. These updates include a revised chemical mechanism from AM3 to AM4.1 to take advantage of new laboratory kinetic data (e.g., Sander et al., 2011; Mao et al., 2013b; as implemented by Li et al., 2016), in particular for photooxidation of biogenic volatile organic compounds (BVOCs). Analysis of deficiencies in AM3 chemistry has pointed out improper treatment of nitrate aerosols and gas-aerosol interactions and biases in wet and dry deposition. We also wanted to leverage successful follow-on development efforts from AM3 targeted to implement reduced nitrogen cycling (Paulot et al., 2016, 2017b) and improved representation of the seasonal cycle in sulfate (Paulot et al., 2017a).

Finally, we wanted to provide the ability for the atmospheric model to handle a more diverse suite of land-atmosphere and ocean-atmosphere linkages for comprehensive Earth system representation of not only heat and hydrology but also CO_2, dust, reactive nitrogen, and organic carbon.

The focus of the present study is to document the atmosphere physics and chemistry developed for AM4.1, as distinguished from the physical climate–focused AM4.0 (Zhao et al., 2018ab), for standalone atmospheric applications. A more comprehensive discussion of coupled atmosphere-ocean-land Earth system interactions in ESM4.1 is provided by Dunne et al. (submitted). We focus our analysis on evaluating the AMIP configuration of AM4.1 used for CMIP6, and document differences in results between AM4.1 and other GFDL CMIP models, including AM4.0 (CMIP6) and AM3 (CMIP5). In the case of comparisons with AM3, the differences in results reflect updates to both the model configuration and the emissions (as discussed in Section 2.4).

2. Model description

A general schematic of AM4.1 forcing, dynamics, physics, aerosol, and chemistry interactions is provided in Figure 1. The following sections discuss the formulations for these components in reference to their AM4.0 (Zhao et al., 2018ab) counterparts.

2.1 Physical model formulation

The physical formulation of AM4.1 is similar to that of AM4.0, but the model top has been raised from 100 Pa (~45 km) to 1 Pa (~80 km), and the number of vertical levels has been increased from 33 to 49, similar to the 48-level structure of AM3. This enhanced vertical extent and resolution allows AM4.1 to represent stratospheric dynamics and chemistry, and

stratosphere-troposphere coupling. The time step used in the dynamical core for gravity wave and the Lagrangian dynamics is reduced from 150 s in AM4.0 to 130 s in AM4.1 for increased numerical stability.

Like AM4.0, AM4.1 includes 5 tracers for water (specific humidity, liquid water, ice water, cloud amount, and liquid droplet number concentration), and uses the same large-scale and convective cloud parameterizations as in AM4.0. Cloud parameterizations in AM4.1 were re-tuned slightly compared to AM4.0 in order to improve agreement with observed top-of-atmosphere shortwave and longwave radiative fluxes, in response to initially excessive reflection from convective clouds over sub-Saharan Africa, North Indian Ocean, and the western tropical Pacific. In particular, the scale factor applied to the fall speed of ice clouds (*c1* in Zhao et al., 2018b) was reduced from 0.90 in AM4.0 to 0.85 in AM4.1 to increase ice water path and decrease outgoing longwave radiation. The cloud erosion timescale (τ_{eros}) in convectively active regions is decreased slightly from 6.9 hours to 5.6 hours to increase the absorbed shortwave radiation. The cloud erosion timescale under other conditions is unchanged from AM4.0.

As described by Zhao et al. (2018b), non-orographic gravity wave drag in AM4.0/AM4.1 is parameterized following Alexander and Dunkerton (1999), but the parameters used in AM4.1 are modified from those in AM4.0 to improve stratospheric circulation. In AM4.1, the magnitudes of the non-orographic gravity wave flux above 350 hPa for the tropics, northern extratropics, and southern extratropics (*St*, *Sn*, and *Ss*) are set to 0.004, 0.005, and 0.0035 $m^2\ s^{-2}$, respectively.

Land hydrology and ecosystem dynamics are represented in AM4.1 by the GFDL Land Model version 4.1 (LM4.1; Shevliakova et al., submitted), replacing the LM4.0 model used in AM4.0

(Zhao et al., 2018b). LM4.1 includes advanced vegetation and canopy competition, fire, land use representation, and dynamic atmospheric dust coupling.

2.2 Atmospheric chemistry and aerosol scheme

AM4.1 includes interactive tropospheric and stratospheric gas-phase and aerosol chemistry. The bulk aerosol scheme, including 18 transported aerosol tracers (see Supplementary material, Table S1), is similar to that in AM4.0 (Zhao et al., 2018b), with the following updates: (1) ammonium and nitrate aerosols are treated explicitly, with the sulfate–nitrate–ammonia thermodynamic equilibrium simulated using ISORROPIA (Fountoukis and Nenes, 2007), as described by Paulot et al. (2016); (2) oxidation of sulfur dioxide and dimethyl sulfide to produce sulfate aerosol is driven by the gas-phase oxidant concentrations (OH, H_2O_2, and O_3) and cloud pH simulated by the online chemistry scheme (Paulot et al., 2016), and (3) the rate of aging of black and organic carbon aerosols from hydrophobic to hydrophilic forms varies with calculated concentrations of hydroxyl radical (OH), as described by Liu et al. (2011).

Unlike AM4.0, the AM4.1 model has an online representation of gas-phase tropospheric and stratospheric chemistry. The combined tropospheric and stratospheric chemistry scheme includes 18 prognostic (transported) aerosol tracers, 59 prognostic gas phase, 6 prognostic ideal tracers, and 41 diagnostic (non-transported) chemical tracers (Table S2), with 43 photolysis reactions, 190 gas-phase kinetic reactions, and 15 heterogeneous reactions (Table S3). The tropospheric chemistry includes reactions of the NO_x-HO_x-O_x-CO-CH_4 system and oxidation schemes for other non-methane volatile organic compounds. The stratospheric chemistry accounts for the major ozone loss cycles (O_x, HO_x, NO_x, ClO_x, and BrO_x) and heterogeneous reactions on liquid and solid stratospheric aerosols as in Austin et al. (2013). The base chemical mechanism is

updated from that in AM3 (Naik et al., 2013a), using gas-phase and heterogeneous chemistry updates from Mao et al. (2013ab), similar to the configuration described by Schnell et al. (2018). We include heterogeneous reactions of HO_2, NO_2, N_2O_5, and NO_3 on the surfaces of all simulated aerosol types, with specified gamma values (given in Table S3). Note in particular that $\gamma(HO_2)$ is reduced from the value of 1 recommended by Mao et al. (2013a) to 0.2. We also include the heterogeneous oxidation of SO_2 on aerosols following Zheng et al. (2015). The heterogeneous chemistry included in AM4.1 has a much stronger effect on oxidant levels than that in AM3, which used $\gamma(N_2O_5)=0.1$, $\gamma(NO_3)=0.1$, $\gamma(NH_3)=0.05$, $\gamma(HO_2)=0$, and $\gamma(NO_2)=0$, applied only to sulfate aerosols. The chemical system is solved using an implicit Euler backward method with Newton-Raphson iteration, as in Horowitz et al. (2003). Photolysis rates are calculated interactively using the FAST-JX version 7.1 code, as described by Li et al. (2016), accounting for the radiative effects of simulated aerosols and clouds.

Dry deposition velocities for all aerosols are calculated interactively using a wind-driven resistance method (Gallagher et al., 2002), accounting for the effect of whitecaps over the ocean (Williams, 1982). The treatment of wet deposition accounts for slower removal by frozen precipitation due to the Bergeron process in mixed-phase clouds (Liu et al., 2011). Dry and wet deposition for gases are as described by Paulot et al. (2016).

Optical properties of aerosols are precalculated, as in AM4.0, using Mie theory assuming sphericity. The extinction efficiency, single scattering albedo, and asymmetry parameter are tabulated as a function of wavelength, aerosol type, aerosol size (for dust and sea salt), and relative humidity. Sulfate is assumed to be internally mixed with black carbon for the calculation of optical properties. Unlike AM4.0, radiative effects of nitrate aerosol are included in AM4.1 (as in Paulot et al., 2017b).

2.3 AMIP (1980–2014) simulation configuration

We conduct AMIP simulations with AM4.1 over the period 1979–2014 using observed gridded sea surface temperature (SST) and sea-ice concentration boundary conditions from the reconstructions of Taylor et al. (2000). Historical reconstructions of monthly solar spectral irradiances are from Matthes et al. (2017). For radiation calculations, global monthly mean concentrations of greenhouse gases (GHGs), including nitrous oxide (N_2O), and ozone depleting substances (ODSs, including CFC-11, CFC-12, CFC-113, and HCFC-22) are specified from Meinshausen et al. (2017). Global mean mixing ratios of methane (CH_4) and N_2O are specified at the surface as lower boundary conditions for chemistry. Carbon dioxide (CO_2) mixing ratio is restored to observed global-mean values with a one-year timescale. The simulated global-mean CO_2 and CH_4 concentrations are used for radiation calculations.

2.4 Emissions

Annually varying time series of monthly anthropogenic and biomass burning emissions of ozone precursors and aerosols (and their precursors) are from the Community Emissions Data System (CEDS; Hoesly et al., 2018) and the data set of van Marle et al. (2017), respectively, developed in support of CMIP6. Wildfire emissions are distributed vertically between the surface and 6 km, with location- and biome-dependent vertical profiles, as recommended by Dentener et al. (2006), similar to the treatment in AM3 (Donner et al., 2011; Naik et al., 2013a). Natural emissions of NO_x, CO, non-methane volatile organic compounds (NMVOC), hydrogen (H_2), and ammonia (NH_3) are generally the same as those considered by Naik et al. (2013a), namely from the Precursors of Ozone and their Effects in the Troposphere (POET) inventory for present day (corresponding to year 2000) (Granier et al., 2005). Emissions of NH_3 from sea bird colonies, not

accounted for in AM3, are included in AM4.1 following Riddick et al. (2012). The treatment of marine ammonia emissions is also revised as described below.

Biogenic emissions of isoprene and monoterpenes are calculated online using the Parameterized Canopy Environment Emission Activity (PCEEA algorithm; Guenther et al., 2006) in the Model of Emissions of Gases and Aerosols from Nature (MEGAN v2.1; Guenther et al., 2012) as a function of simulated air temperature and shortwave radiative fluxes, implemented as described by Rasmussen et al (2012). Leaf area indices for 17 plant functional types are based on AVHRR and MODIS data and are mapped to five vegetation types (Emmons et al. 2010). These vegetation types and leaf area indices are independent of those simulated by the LM4.1 dynamic vegetation model, due to a lack of coupling between the dynamic vegetation properties simulated by LM4.1 and the atmospheric emissions module. We do not apply the soil moisture or CO_2 responses from Guenther et al. (2012). Future model development plans include coupling biogenic emissions to LM4.1. Sea salt emissions are based on the parameterization of Monahan (1986) as in CM3 (Donner et al., 2011), but are modulated by sea surface temperature following Jaeglé et al. (2011). Ocean ammonia emissions are calculated following Paulot et al. (2015), using the simulated seawater concentration of NH_4^+ in ESM4.1. Other marine emissions, including primary organic aerosols (POA) and dimethyl sulfide (DMS), are calculated similarly to in CM3. DMS emissions are calculated using an empirical formula as a function of a prescribed monthly climatology of DMS concentration in sea water (Lana et al., 2011) and calculated wind speed at 10 m, as described by Chin et al. (2002). Thus, oceanic emissions of POA, DMS, ammonia, and sea salt are dependent on the simulated meteorology in the model.

Emission totals for year 2014 are shown in Table 1. Time series of annual global emissions in AM4.1 (using CMIP6 inventories) are shown for select species in Figure 2, and compared with corresponding totals in AM3 (using CMIP5 inventories).

Sources of secondary organic aerosols (SOA) include an anthropogenic source from oxidation of the simulated C_4H_{10} hydrocarbon tracer by hydroxyl radical (with a 10% per-carbon yield) and a biogenic pseudo-emission assuming a 10% per-carbon yield from emissions of BVOCs, including isoprene and monoterpenes, from vegetation. This yield is in the range of values suggested by recent studies using more detailed schemes for SOA production (e.g., Bates and Jacob, 2019; Pai et al., 2020). In year 2014, the sources of SOA are 83.84 Tg a^{-1} from BVOCs and 3.49 Tg a^{-1} from anthropogenic hydrocarbon oxidation.

Lightning NO_x emissions are calculated interactively as a function of subgrid convection in AM4.1, as diagnosed by the double plume convection scheme described by Zhao et al. (2018b). The lightning NO_x source is calculated as a function of convective cloud-top height, following the parameterization of Price et al. (1997), and is injected with the vertical distribution of Pickering et al. (1998), as in AM3 (Naik et al., 2013a). The global total production of NO_x by lightning is 3.59 Tg N for year 2014.

Dust emissions are calculated dynamically online in the land component, LM4.1, as a function of wind speed, topography, vegetation cover, snow cover, soil moisture, and land type, as described by Evans et al. (2016) and Shevliakova et al. (submitted).

As in AM3, direct injection of SO_2 from volcanic eruptions and emissions of carbonyl sulfide (COS) are not considered in AM4.1. Instead, we specify time series of stratospheric aerosol optical properties, accounting for not only the volcanic contribution to stratospheric aerosol

abundance but also other natural and anthropogenic contributions. Tropospheric emissions of SO_2 from continuously degassing and explosive volcanoes are treated in the same way as in AM3 (Donner et al., 2011), with a climatological total of 3.59 Tg S a^{-1}.

3. Results: Physical climate simulation (AMIP, 1980–2014)

Surface air temperature

Comparison of surface air temperature over land with observations from CRU TS (Figure 3) illustrates the substantial decrease in overall root mean square error (RMSE) achieved in AM4.1 (RMSE = 1.92°C) from the previous generation full-chemistry AM3 (RMSE = 2.18°C) and similar, if slightly degraded, pattern to AM4.0 (RMSE = 1.85°C). The most notable difference from AM3 to AM4.0 and AM4.1 is an improvement in boreal warm biases and South American cold biases.

Precipitation

Comparison of precipitation with observations from GPCP v2.3 (Figure 4) also illustrates the substantial decrease in overall RMSE achieved in AM4.1 (RMSE = 0.83 mm d^{-1}) from the previous generation full-chemistry AM3 (RMSE = 1.02 mm d^{-1}) and a similar pattern to AM4.0 (RMSE = 0.85 mm d^{-1}). The most notable difference from AM3 to AM4.0 and AM4.1 is an improvement in Amazon dry biases, and in wet biases over Australia and the Indian Ocean.

Circulation

Comparison of zonal mean zonal winds with the ERA40 reanalysis (Figure 5) illustrates a substantial decrease in overall RMSE in AM4.1 (RMSE = 1.32 m s^{-1}) from the previous generation high-top full-chemistry AM3 (RMSE = 1.75 m s^{-1}). The AM4.1 RMSE is greater than

254 that in the low-top AM4.0 (RMSE = 1.00 m s^{-1}), owing to a westerly wind bias in the equatorial
255 stratosphere, and a weak, equatorward-shifted Arctic stratospheric jet in AM4.1. The
256 representation of the stratospheric wintertime westerly polar jet associated with the Antarctic
257 vortex is significantly improved in AM4.1 (not shown) compared with AM3 (Donner et al.,
258 2011), in which the westerlies were excessively strong (leading to a too-cold Antarctic vortex).
259 We plan to work towards further improving the stratospheric circulation in future versions of
260 AM4.1 through improvement in our representation of parameterized gravity wave drag.
261 Tropospheric circulation patterns in AM4.1 are very similar to those in AM4.0.

262 *Stratospheric variability*

263 Comparison of statistics for sudden stratospheric warmings with the ERA40 reanalysis (Figure 6)
264 illustrates an improvement in AM4.1 with respect to capturing events in the coldest months
265 (December–January), which were largely missed in AM4.0 (Zhao et al., 2018a), even though
266 AM4.0 already performs quite well among low-top atmospheric models (Charlton-Perez et al.,
267 2013). In the surrounding months (November, February), AM4.1 overestimates warming events,
268 whereas AM4.0 matches the reanalysis data fairly well.

269 *Radiation fluxes*

270 Comparison of top-of-atmosphere (TOA) net radiation with CERES EBAF observations (Figure
271 7) further illustrates the substantial decrease in overall root mean square error (RMSE) achieved
272 in AM4.1 (RMSE = 7.2 W m^{-2}) from the previous generation full-chemistry AM3 (RMSE = 8.6
273 W m^{-2}) and similar, if slightly degraded, pattern to AM4.0 (RMSE = 6.8 W m^{-2}). The most
274 notable difference from AM3 to AM4.0 and AM4.1 is associated with an improvement in areas
275 of tropical convection along the intercontinental convergence zone (ITCZ) that had previously

been too absorbing and increased absorption in northern boreal regions that had been previously too reflective, as discussed by *Zhao et al*. (2018a). The most notable differences between AM4.0 and AM4.1 are associated with a decrease in the global TOA from a nearly unbiased AM3 (0.02 W m^{-2}) to a slight negative bias in AM4.0 (-0.14 W m^{-2}) and substantial low bias in AM4.1 (-0.80 W m^{-2}). This increase in bias is due in part to increased albedo of northern boreal regions associated with snow masking depth in LM4.1, as discussed by Shevliakova et al. (submitted), and also in part to differences over Antarctic associated with the prescribed albedo of snow on glaciers that was modified late in the development cycle of ESM4.1 to address Southern Ocean dynamics, as discussed by Dunne et al. (submitted).

Lightning flash frequency

Figure S1 shows the lightning flash frequency retrieved from the spaceborne Optical Transient Detector (OTD) and Lightning Imaging Sensor (LIS) (Cecil et al., 2014), compared with simulated values from AM3 and AM4.1. In both AM3 and AM4.1, lightning flash frequency is parameterized as a function of convective cloud top height, following Price et al. (1997), but the two models use different parameterizations of cumulus convection (Zhao et al., 2018b). While the overall correlation between model and observations is lower in AM4.1 than AM3, there are some notable areas of improvement in the representation of flash frequency, including a reduction of the high biases present in AM3 over the Amazon and the maritime continent, improving agreement with observations.

4. Results: Simulation of atmospheric composition

4.1 Ozone

In this section, we evaluate model simulations of ozone, including surface ozone concentrations relevant for air quality and column ozone abundances relevant for climate.

Surface ozone

We focus on the seasonal mean of the maximum daily 8-hour average (MDA8) surface ozone over the period 2005–2014, when observations are available from densely clustered monitoring sites across northern mid-latitude populated regions (Figure 8 for MAM, Figure 9 for JJA**).** Observations were obtained from the Tropospheric Ozone Assessment Report (TOAR) Database for 2005–2014 (Schultz et al., 2017) and a monitoring network operated since 2013 by China's Ministry of Environmental Protection (CNMEP, http://106.37.208.233:20035/). Observations are averaged onto the same 1°x1° grid as AM4.1. We compare simulated ozone from the AM4.1 AMIP simulation with that from the AM3 AMIP simulation.

Surface MDA8 ozone in AM3 is biased high by 12 ppb on average during MAM (Figure 8b) and by up to 20 ppb over the eastern U.S. during summer (Figure 9b), as documented in previous studies (Lin et al., 2012a; Lin et al., 2012b; Fiore et al., 2014; Rieder et al., 2015; Lin et al., 2017). AM4.1 shows substantially reduced biases in mean ozone for both spring and summer over the eastern U.S. and Europe (Figures 8c and 9c). This dramatic improvement in the simulation of surface ozone concentrations results from a combination of updates to the chemical mechanism from AM3 to AM4.1, including updates to the isoprene oxidation scheme and the representation of heterogeneous reactions (Mao et al., 2013b), and the change from CMIP5 emissions in AM3 to CMIP6 emissions in AM4.1 (Section 2.3). The shallow surface layer of the model (30 m thick) may also have an impact on the comparison with surface sites. Zhao et al. (2018a) found a

318 significant improvement in diagnosed 2-m temperatures associated with this shallower surface
319 layer.

320 To further explore the causes of the differences in surface ozone abundances between AM3 and
321 AM4.1, we conduct two additional simulations—an AM4.1 simulation with nudged meteorology
322 and an additional AM4.1 nudged simulation with AM3-like chemistry (AM4.1_AM3Chem; Lin
323 et al., 2019). The two experiments use the same CMIP6 emissions and have nearly identical
324 meteorology (as a result of the nudging), allowing us to isolate the influence of changes in
325 chemistry alone. Seasonal-mean MDA8 ozone from these simulations are plotted in
326 Supplementary Figures S2–S5.

327 Similar to the results from AM3 (Figures 8b and 9b), surface MDA8 ozone in AM4.1_AM3Chem
328 is biased high by 11 ppb on average during spring (Figure S2b) and by up to 20 ppb over the
329 eastern U.S. during summer (Figure S3b). Switching the chemistry scheme from AM3 to AM4
330 leads to substantial reductions in mean ozone biases for both spring and summer over the eastern
331 U.S. and Europe (Figures S2c and S3c), but the model underestimates springtime MDA8 ozone
332 over central eastern China by 20 ppb (Figure S2c versus CNMEP observations in Figure S2a).

333 Observations show more severe springtime ozone pollution over central eastern China and Mexico
334 than in the U.S. and Europe. This regional contrast is not simulated in either of our experiments.
335 Particularly, the enhanced heterogeneous chemistry in AM4 (Section 2.2; Mao et al., 2013ab)
336 likely leads to excessive heterogeneous loss of HO_x and NO_x radicals over eastern China and
337 Mexico, where aerosol loadings are high during the spring season. For summer over the
338 southeastern U.S., where high mean-state ozone biases are found in many current-generation
339 CTMs and CCMs (Fiore et al., 2009; Young et al., 2018), the AM4.1 experiment shows remarkable

agreement with observations. However, on the basis of analysis conducted for an intensive field campaign, Travis et al. (2016) suggested that the common model biases in simulating summertime ozone over the southeastern U.S. may reflect a combination of excessive NO_x emissions (too high by 50%) and the deep model surface layer that cannot resolve near-surface ozone gradients. A balanced view is needed to interpret the reduced ozone biases in the AM4.1 experiment.

Our results suggest the complexity of various sources, sinks, transport, and chemistry in influencing the simulation of surface ozone. In the future, process-based assessments, not only for means but also for variability and extreme events, are needed to fully evaluate how the choices of different emission datasets, chemistry, and deposition schemes affect simulations of surface ozone and related tracers.

Tropospheric ozone column

We compare climatological annual mean tropospheric ozone columns simulated by AM3 (mean over 2000–2008) and AM4.1 (2005–2014) with those derived from the OMI-MLS (Ziemke et al., 2019) (Figure 10). In the analysis shown here, AM3's native ozone output on model levels is used to calculate tropospheric ozone column using the WMO tropopause definition, while for AM4.1, the tropospheric ozone column (tropoz) is diagnosed at every time step, by applying the WMO tropopause definition using model simulated temperature. The global mean tropospheric ozone column simulated by AM3 and AM4.1 are 35 DU and 31 DU, respectively, compared to the OMI/MLS value of 30 DU. While AM3 showed consistent high biases globally except over the Antarctic, AM4.1 shows an interhemispheric pattern in the biases with high values in the Northern Hemisphere mid-latitudes and over continents and low values in the Southern Hemisphere extra-tropics. This pattern is consistent with global chemistry-climate models

362 evaluated against the OMI/MLS climatology by Young et al. (2013) for a slightly different time
363 period. An interesting feature in AM4.1 is the strong positive bias over Oceania, possibly related
364 to the different biomass burning emissions applied in the two models. AM3 exhibited an average
365 high bias of 21.7%, which has been reduced to 7.3% in AM4.1; accordingly the RMSE has been
366 reduced considerably, from 7.1 DU in AM3 to 4.6 DU in AM4.1.

367 *Total Ozone Column*

368 Figure 11 shows the evaluation of modeled time series of total column ozone against two
369 datasets for 1980–2015, namely Multi-Satellite Merged Total Column NASA and NOAA
370 product from Frith et al. (2013; SBUV; open triangles) and version 3.4 of the National Institute
371 of Water and Atmospheric Research - Bodeker Scientific (NIWA-BS; closed circles) total
372 column ozone database. AM3 results are plotted for 1980–2008 period, while AM4.1 results are
373 for 1980–2014. The comparison is shown for the annual average globally, in the tropics, and in
374 southern and northern midlatitudes, and for March in the Arctic and October in the Antarctic.
375 Globally (Figure 11a), absolute values of total column ozone for AM3 were biased high
376 compared to both datasets, whereas AM4.1, on the other hand, is biased low. Both models
377 generally capture the trend in total column ozone, although the evaluation of AM3 is truncated at
378 2008. As suggested by the greater correlation coefficients for AM4.1 compared with AM3,
379 AM4.1 is able to better capture the observed interannual variability and trends of global mean
380 total column ozone. In the tropics (Figure 11b), total ozone column values remain lower than
381 observed in AM4.1, as opposed to higher in AM3. Consistent with observations, both the models
382 simulate negligible trends in total column ozone in the tropics, however AM4.1 exhibits greater
383 skill in capturing the observed evolution of total column ozone. In the northern mid-latitudes
384 (Figure 11c), AM4.1 differs more from observations than AM3 does, although with fairly similar

skill in simulating the observed time evolution of total column ozone. The comparison is opposite for the southern mid-latitudes (Figure 11d), where AM4.1 is much closer to observed values than AM3 with similar correlations. In the Arctic in March (Figure 11e), AM4.1 reproduces the observed total ozone column values slightly better than AM3, however both have fairly low skill in reproducing the observed evolution. In the Antarctic in October (Figure 11f), AM4.1 exhibits greater skill in simulating ozone depletion compared to AM3 both in terms of trends and absolute values. This improvement likely results from the improved dynamical representation of the Antarctic polar vortex in AM4.1 (Section 3). Overall, AM4.1 compares slightly better against observations of total column ozone than AM3.

4.2 Carbon monoxide

The simulated tropospheric CO columns are evaluated against CO retrievals from the MOPITT (Measurements of Pollution in The Troposphere) instrument in Figure 12. We use the MOPITT V8 Joint (NIR+TIR) retrievals (Deeter et al., 2019) during 2001–2014, which are available from the NASA Earthdata archive (https://earthdata.nasa.gov). The model is interpolated to the gridded monthly MOPITT observations and the averaging kernel for each grid is applied to the simulated monthly mean CO profiles.

The tropospheric CO columns are in general higher in AM4.1 than AM3, in better agreement with MOPITT retrievals in terms of magnitudes (RMSE reduced from (2.6–2.7)x10^{17} cm^{-2} to (1.6–1.8)x10^{17} cm^{-2}) and spatial distribution (r^2 increased from 0.7–0.9 to 0.8–0.9). Compared to AM3, AM4.1 reduces the underestimations in column CO in the Northern Hemisphere, but overestimates column CO in the Southern Hemisphere, especially during summer. This is in part due to lower OH levels in AM4.1 than AM3.

To evaluate surface CO, we use measurements from a globally distributed network of air sampling sites maintained by the Global Monitoring Laboratory (GML) of the National Oceanic and Atmospheric Administration (NOAA) (Pétron et al., 2019; data available at ftp://aftp.cmdl.noaa.gov/data/trace_gases/co/flask/). Surface CO observations during 1988–2014 are used to evaluate model performance (Figure 13).

AM4.1 simulates higher surface CO concentrations than AM3 over the Southern Hemisphere, and slightly overestimates surface CO concentrations by < 5 ppb when compared to surface observations. Over the Northern Hemisphere, AM4.1 largely reduces the negative biases that occurred in AM3, with a mean bias of ±20 ppb over most GMD sites. This is consistent with the comparisons to the MOPITT retrievals shown above. In addition, compared to AM3, AM4.1 better captures the seasonal cycles (with correlation coefficient $R > 0.5$) at most sites and better captures the latitudinal gradient as well ($R = 1.0$ versus $R = 0.9$).

Comparisons of surface CO concentrations over pristine sites show significant improvement in AM4.1 over AM3 across latitudes from South to North. In the Southern Hemisphere, such as at South Pole (SPO), Ushuaia (USH), and Easter Island (EIC) sites, the underestimation of surface CO concentrations by AM3 are reduced in AM4.1. In the Northern Hemisphere, such as at Mauna Loa (MLO), Barrow (BRW), and Alert (ALT) sites, both surface concentrations and monthly variations are improved significantly in AM4.1 compared to AM3. These improvements are mainly associated with improved chemistry in AM4.1.

4.3 Aerosols

We first evaluate concentrations of aerosols in surface air. Figure 14 (top panels) compare simulated concentrations of sulfate and nitrate aerosols from AM4.1 with observations over the

United States from the IMPROVE network. The model successfully captures the wide range of observed sulfate aerosol concentrations. While nitrate concentrations are well correlated with observations (R=0.74), simulated concentrations are generally too high (normalized mean bias (NMB) = +80%). This bias is larger than in Paulot et al. (2016), where nitrate aerosols are assumed to deposit rapidly like nitric acid. Simulated concentrations of sulfate and nitrate in precipitation are compared with observations from the NADP network in the lower panels of Figure 14. The rainwater abundances of sulfate and nitrate are well correlated with observations, but with a low bias for sulfate (NMB = -19%) and a high bias for nitrate (NMB = +35%).

Figure 15 compares simulated concentrations of sulfate, dust, and sea salt aerosols from AM3, AM4.0, and AM4.1 with observations from the University of Miami network (Savoie and Prospero, 1977). The model successfully captures the wide range of observed sulfate aerosol concentrations. For sulfate, the RMS error versus observations is reduced in AM4.1 (0.20 μg m^{-3}) from AM3 and AM4.0 (both 0.22 μg m^{-3}), and the correlation is improved (r = 0.93 in AM4.1, r = 0.89 in AM3 and AM4.0). The agreement between simulated and observed dust improves from AM3 to AM4.0, but then degrades in AM4.1, reflecting the shift from prescribed to interactive source regions for dust in LM4.1 (Shevliakova et al., submitted). The RMS error for simulated sea salt is reduced significantly in AM4.1 (0.35 μg m^{-3}) compared with AM3 (0.47 μg m^{-3}) and AM4.0 (0.49 μg m^{-3}), as a result of updates to the emissions and deposition parameterizations in AM4.1.

We next evaluate the simulated AOD against measurements from the AERONET sunphotometer network (Holben et al., 1998) in Figure 16. Here we use the quality assured and cloud screened level 2 version 2 AOD data (Smirnov et al., 2000). For comparison, we also show the results from AM4.0 (middle) and AM3 (bottom). Both AM4.0 and AM4.1 exhibit higher correlation

(0.89 and 0.9) and lower RMS (0.07 and 0.08) with AERONET observations than AM3 (0.81 and 0.09, respectively). In particular, the large positive biases in the tropics and equatorial regions are reduced, which reflects the more efficient removal of aerosol by convective precipitation (Paulot et al., 2016). AM4.1 exhibits a greater positive bias than AM4.0 over the Midwest United States, associated with higher dust loading and nitrate aerosol (not included in AM3 and AM4.0).

Figures 17 and S6 and compare the simulated regional monthly mean AOD simulated by AM3, AM4.0, and AM4.1 with observations from the MODIS (Levy et al., 2007) and MISR (Kahn et al., 2009) instruments. AM4.0 and AM4.1 have reduced the seasonal contrast between winter and summer months, in better agreement with observational constraints. The spring maximum over East Asia and the North Pacific is also better captured with AM4.0 and AM4.1. The AM3 high bias over the Caribbean Sea and maritime continent are reduced consistent with the comparison against AERONET.

These improvements primarily reflect changes in the treatment of aerosol removal, including reduced removal by frozen precipitation formed by the Bergeron process and more efficient scavenging by convective precipitation (Paulot et al., 2016). AM4.1 exhibits greater bias over Asia than AM4.0, which primarily reflects higher optical depth from dust and ammonium nitrate. Uncertainties in Asian SO_2 and NH_3 emissions (Zhang et al., 2009) and aerosol hygroscopic growth may also contribute to the AM4.1 high bias over this region (Paulot et al., 2018).

4.4 Hydroxyl Radical (OH) and Methane Lifetime

Here, we evaluate the climatological mean hydroxyl (OH) radical simulated by AM4.1, as OH is the primary atmospheric oxidant determining the abundance and lifetime of several short-lived

474 climate forcers, including methane. The simulation of OH depends on the chemical mechanism,
475 particularly the representation of isoprene photooxidation (Archibald et al., 2010; Bates and
476 Jacob, 2019). Differences in emissions, meteorology, and photochemical mechanisms across
477 models also lead to differences in OH (Murray et al., submitted 2019; Wild et al., 2020).
478 Climatological mean (1980–2014) global airmass-weighted tropospheric OH simulated by
479 AM4.1 is 10.4×10^5 molecules cm^{-3}, about 18% lower than that simulated by AM3, but is within
480 the range of values reported for ACCMIP models for the 2000s (Naik et al., 2013b).
481 Consequently, the mean whole-atmosphere chemical lifetime of methane (calculated as the
482 global methane burden divided by global total loss) in AM4.1 is 8.5 years; lifetime against loss
483 by reaction with tropospheric OH is 9.7 years, which is 13% greater than the AM3 value of 8.6
484 years (1981–2000), but still lower than the observationally derived estimate of 11.2±1.3 years
485 (Prather et al., 2012). Figure 18 shows the comparison of tropospheric OH distribution for 12
486 regions simulated by AM4.1 with estimates from AM3, ACCMIP ensemble mean, and the
487 climatology of Spivakovsky et al. (2000). AM4.1 simulates reduced OH levels compared to
488 AM3 throughout the troposphere, possibly because of differences in emissions and chemical
489 mechanisms between the two model versions. In particular, the lower lightning NO_x in AM4.1
490 versus AM3 acts to lower OH because of the strong sensitivity of OH to lightning NO_x emissions
491 (Murray et al., 2013). Relative to the Spivakovsky et al. climatology, AM4.1 exhibits a reduced
492 high bias compared with AM3, but has too low OH, particularly in the tropical upper
493 troposphere.

494 **5. Sensitivities to Greenhouse Gases, Aerosols, and SST Perturbations**

495 Table 2 shows the net radiative flux perturbations that result from historical changes in
496 anthropogenic forcing agents, and from idealized changes in CO_2 and SST. Comparison of these

497 radiative metrics between AM3, AM4.0 and AM4.1 indicates that effective radiative forcings
498 (ERF) from pre-industrial to present-day changes in greenhouse gases and aerosols are nearly
499 identical between AM4.0 and AM4.1. However, the ERF from quadrupling CO_2 is significantly
500 lower in AM4.1, mostly because of the inclusion of interactive ozone (colder stratospheric
501 temperatures reduce the rate of ozone chemical loss), but also partially resulting from increased
502 dust emissions from LM4.1 (related to increased fires under elevated-CO_2 conditions). The Cess
503 feedback, the change in net radiative flux resulting from an increase of SSTs by 2K, is
504 significantly more negative in ESM4.1 (corresponding to a weaker Cess sensitivity), likely
505 resulting from increased emissions of salt, dust, and BVOCs with increasing temperatures in
506 ESM4.1. While comparison with previous-generation models is complicated by changes in the
507 AMIP configuration since the AM3 model simulations were conducted (in particular, updating
508 the "present-day" conditions from representing 1990 conditions to 2014 conditions), some
509 assessment of these differences can be made using AM4.0 simulations conducted for 1990
510 conditions (as in Zhao et al., 2018a). The most important differences between AM3 and AM4.0
511 are a decrease in the magnitude of the negative aerosol ERF from AM3 to AM4.0, an increase in
512 the 4xCO2 ERF consistent with an update to the treatment of CO_2 radiative bands (Zhao et al.,
513 2018b), and a strengthening of the negative Cess feedback. The decrease in the magnitude of the
514 aerosol ERF from AM3 to AM4.0 has been attributed by Zhao et al. (2018b) to a decrease in the
515 strength of the aerosol indirect effect, resulting from the increase in horizontal resolution and
516 improvements to the representations of aerosol convective wet deposition (Paulot et al., 2016)
517 and aerosol activation.

6. Summary

AM4.1 includes considerable advances in resolution and physics as in AM4.0 (Zhao et al, 2018ab) as well as a comprehensively revised suite of chemistry parameterizations to improve consistency in treatment across species and with advances in the underlying science over the last decade. AM4.1 is able to maintain the fidelity of AM4.0 while substantially increasing in comprehensiveness and associated climate-chemistry interactions and feedbacks. This development effort has also led to considerable improvement in model fidelity compared to GFDL's previous-generation coupled chemistry-climate model (AM3) with respect to observed atmospheric composition for aerosol, CO, ozone, as well as climate phenomena such as sudden stratospheric warmings.

Acknowledgements

We thank GFDL's Model Development Team, Modeling Systems Division, and Operations group for keeping ESM4.1 development and implementation going through many technical challenges. We thank the GFDL Data Portal team for their tireless work on developing and maintaining the vast infrastructure needed to enable publication of data output from our CMIP6 models to the Earth System Grid. We are also grateful to the many groups that freely share the observations used here to evaluate various aspects of chemistry and composition simulated by the models and to the input4MIPs activity (https://esgf-node.llnl.gov/projects/input4mips/) for making available the boundary condition and forcing datasets needed for performing model simulations. This manuscript benefited from internal GFDL reviews by Leo Donner and Songmiao Fan. Discussions with Songmiao Fan were helpful in improving the formulation of dry and wet deposition of aerosols. Data is provided at https://doi.org/10.22033/ESGF/CMIP6.1407. Model code is provided at https://data1.gfdl.noaa.gov/nomads/forms/esm4/. The input data is provided at ftp://data1.gfdl.noaa.gov/users/ESM4/ESM4Documentation/GFDL-ESM4/inputData/ESM4_rundir.tar.gz

References

Alexander, M. J., and Dunkerton, T. J. (1999). A spectral parameterization of mean-flow forcing due to breaking gravity waves. *J. Atmos. Sci.*, **56**, 4167–4182.

546 Austin, J., and Wilson, R. J. (2006). Ensemble simulations of the decline and recovery of
547 stratospheric ozone. *J. Geophys. Res.*, **111**, D16314, doi:10.1029/2005JD006907.

548 Austin, J., Horowitz, L. W., Schwarzkopf, M. D., Wilson, R. J., and Levy, H. II (2013).
549 Stratospheric ozone and temperature simulated from the preindustrial era to the present day.
550 *J. Climate*, **26**, 3528-3543, DOI:10.1175/JCLI-D-12-00162.1.

551 Bates, K. H., and Jacob, D. J. (2019). A new model mechanism for atmospheric oxidation of
552 isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary
553 organic aerosol. *Atmos. Chem. Phys.*, **19**, 9613–9640, https://doi.org/10.5194/acp-19-9613-
554 2019.

555 Cecil, D. J., Buechler, D. E., and Blakeslee, R. J. (2014). Gridded lightning climatology from
556 TRMM-LIS and OTD: Dataset description. *Atmospheric Research*, **135–136**, 404-414,
557 https://doi.org/10.1016/j.atmosres.2012.06.028.

558 Charlton, A., and Polvani, L. M. (2007). A new look at stratospheric sudden warmings. Part I:
559 Climatology and modeling benchmarks. *Journal of Climate*, **20**, 449–469.

560 Charlton, A., Polvani, L. M., Perlwitz, J., Sassi, F., Manzini, E., Shibata, K., et al. (2007). A new
561 look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations.
562 *Journal of Climate*, **20**, 470–488.

563 Charlton-Perez, A. J., Baldwin, M. P., Birner, T., Black, R. X., Butler, A. H., Calvo, N., et al.
564 (2013). On the lack of stratospheric dynamical variability in low-top versions of the CMIP5
565 models. *Journal of Geophysical Research: Atmospheres*, **118**, 2494–2505.
566 https://doi.org/10.1002/jgrd.50125

567 Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B.N., Duncan, B.N., et al. (2002).
568 Tropospheric Aerosol Optical Thickness from the GOCART Model and Comparisons with
569 Satellite and Sun Photometer Measurements. *J. Atmos. Sci.*, **59**, 461–483,
570 https://doi.org/10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2.

571 Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S.,
572 Worden, H. M., Ziskin, D., and Andreae, M. O. (2019). Radiance-based retrieval bias
573 mitigation for the MOPITT instrument: the version 8 product. *Atmos. Meas. Tech.*, **12**, 4561–
574 4580, https://doi.org/10.5194/amt-12-4561-2019.Dentener, F., Kinne, S., Bond, T., Boucher,
575 O., Cofala, J., Generoso, S., et al. (2006). Emissions of primary aerosol and precursor gases
576 in the years 2000 and 1750 prescribed data-sets for AeroCom. *Atmos. Chem. Phys., 6*(12),
577 4321-4344. doi:10.5194/acp-6-4321-2006

578 Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., et al. (2011).
579 The dynamical core, physical parameterizations, and basic simulation characteristics of the

atmospheric component AM3 of the GFDL global coupled model CM3. *Journal of Climate*, **24**(13), 3484-3519.

Dunne, J. P, Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., et al. (submitted). The GFDL Earth System Model version 4.1 (GFDL-ESM4.1): Model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*.

Evans, S., Ginoux, P., Malyshev, S., and Shevliakova, E. (2016). Climate-vegetation interaction and amplification of Australian dust variability. *Geophys. Res. Lett.*, **43**, 11,823– 11,830, doi:10.1002/2016GL071016.

Frith, S. M., et al. (2013). Multi-Satellite Merged Ozone (O3) Profile and Total Column 1 Month Zonal Mean L3 Global 5.0 degree Latitude Zones V1, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), 10.5067/MEASURES/OZONE/DATA310.

Fiore, A. M., Oberman, J. T.,Lin, M. Y., Zhang, L., Clifton, O. E., Jacob, D. J., et al. (2014). Estimating North American background ozone in U.S. surface air with two independent global models: Variability, uncertainties, and recommendations. *Atmos. Environ.*, **96**, 284-300, doi: 10.1016/j.atmosenv.2014.07.045.

Fiore, A. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P., et al. (2009). Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. *J. Geophys. Res.*, **114**, doi: 10.1029/2008jd010816.

Fountoukis, C. and Nenes, A. (2007). ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+-Ca^{2+}-Mg^{2+}-NH_4^+-SO_4^{2-}-NO_3^--Cl^--H_2O aerosols. *Atmos. Chem. Phys.*, **7**, 4639–4659, doi:10.5194/acp-7-4639-2007.

Gallagher, M. W., E. Nemitz, J. R. Dorsey, D. Fowler, M. A. Sutton, M. Flynn, and J. Duyzer, (2002). Measurements and parameterizations of small aerosol deposition velocities to grassland, arable crops, and forest: Influence of surface roughness length on deposition. *J. Geophys. Res.*, **107**, 4154, doi:10.1029/2001JD000817.

Granier, C., J. F. Lamarque, A. Mieville, J. F. Muller, J. Olivier, J. Orlando, J. Peters, G. Pétron, G. Tyndall, and S. Wallens (2005). POET, a database of surface emissions of ozone precursors, available at: http://www.aero.jussieu.fr/projet/ACCENT/POET.php.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmospheric Chemistry and Physics*, **6**(11), 3181-3210.

612 Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., &
613 Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1
614 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions.
615 *Geoscientific Model Development*, **5**(6), 1471-1492.

616 Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., et al. (2019).
617 Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in*
618 *Modeling Earth Systems*, **11**. https://doi.org/10.1029/2019MS001829.

619 Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., et al.
620 (2018). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from
621 the Community Emissions Data System (CEDS). *Geoscientific Model Development (Online)*,
622 **11**(PNNL-SA-123932).

623 Holben, B. N., and Coauthors (1998). AERONET—A federated instrument network and data
624 archive for aerosol characterization. *Remote Sens. Environ.*, **66**, 1–16.

625 Horowitz, L. W., Walters, S., Mauzerall, D. L., Emmons, L. K., Rasch, P. J., Granier, C., ... &
626 Tyndall, G. S. (2003). A global simulation of tropospheric ozone and related tracers:
627 Description and evaluation of MOZART, version 2 (DOI 10.1029/2002JD002853). *J.*
628 *Geophys. Res.*, **108**(24), ACH-16.

629 Jaeglé, L., Quinn, P. K., Bates, T. S., Alexander, B., & Lin, J. T. (2011). Global distribution of
630 sea salt aerosols: new constraints from in situ and remote sensing observations. *Atmospheric*
631 *Chemistry and Physics*, **11**(7), 3137-3157.

632 Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., et al. (2009).
633 MISR aerosol product attributes and statistical comparisons with MODIS. *IEEE*
634 *Transactions on Geoscience and Remote Sensing*, **47**, 4095– 4114.
635 https://doi.org/10.1109/TGRS.2009.2023115

636 Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., et al. (2011). An
637 updated climatology of surface dimethlysulfide concentrations and emission fluxes in the
638 global ocean. *Global Biogeochem. Cycles*, **25**, GB1004, doi:10.1029/2010GB003850.

639 Levy, H. II, Horowitz, L. W., Schwarzkopf, M. D., Ming, Y., Golaz, J.-C., Naik, V., and
640 Ramaswamy, V. (2013). The roles of aerosol direct and indirect effects in past and future
641 climate change. *J. Geophys. Res.*, **118**, 4521-4532, doi:10.1002/jgrd.50192.

642 Levy, R. C., Remer, L., Mattoo, S., Vermote, E. F., & Kaufman, Y. J. (2007). Second-generation
643 operational algorithm: Retrieval of aerosol properties over land from inversion of moderate
644 resolution imaging spectroradiometer spectral reflectance. *Journal of Geophysical Research*,
645 **112**, D13211.

Li, J., Mao, J., Min, K.-E., Washenfelder, R. A., Brown, S. S., Kaiser, J., Keutsch, F.N., et al. (2016). Observational constraints on glyoxal production from isoprene oxidation and its contribution to organic aerosol over the Southeast United States. *J. Geophys. Res.*, **121**, 9849-9861.

Lin, M., Horowitz, L. W., Payton, R., Fiore, A. M., and Tonnesen, G. (2017). US surface ozone trends and extremes from 1980 to 2014: quantifying the roles of rising Asian emissions, domestic controls, wildfires, and climate. *Atmos. Chem. Phys.*, **17**, 2943-2970, doi: 0.5194/acp-17-2943-2017.

Lin, M., S. Malyshev, E. Shevliakova, F. Paulot, L. W. Horowitz, S. Fares, et al. (2019). Sensitivity of ozone dry deposition to ecosystem-atmosphere interactions: A critical appraisal of observations and simulations. *Global Biogeochem. Cycles*, **33**, 1264–1288, doi:10.1029/2018GB006157.

Lin, M., A. M. Fiore, O. R. Cooper, L. W. Horowitz, A. O. Langford, H. Levy, et al. (2012a). Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions. *J. Geophys. Res.*, **117**, doi: 10.1029/2012jd018151.

Lin, M., A. M. Fiore, L. W. Horowitz, O. R. Cooper, V. Naik, J. Holloway, et al. (2012b). Transport of Asian ozone pollution into surface air over the western United States in spring. *J. Geophys. Res.*, **117**, doi: 10.1029/2011jd016961.

Lin, M., Horowitz, L. W., Xie, Y., Paulot, F., Malyshev, S., Shevliakova, E., et al. (2020). Ecosystem-atmosphere interactions exacerbate climate penalty on air pollution extremes in Europe. *Nat. Clim. Chang.*, **10,** 444–451, https://doi.org/10.1038/s41558-020-0743-y.

Liu, J., Fan, S., Horowitz, L. W., and Levy, H. II (2011). Evaluation of factors controlling long-range transport of black carbon to the Arctic. *J. Geophys. Res.*, **116**, D04307, doi:10.1029/2010JD015145.

Mao, J., Horowitz, L. W., Naik, V., Fan, S., Liu, J., and Fiore, A. M. (2013a). Sensitivity of tropospheric oxidants to biomass burning emissions: implications for radiative forcing. *Geophys. Res. Lett.*, **40**, 1241– 1246, doi:10.1002/grl.50210.

Mao, J., Fan, S., Jacob, D. J., and Travis, K. R. (2013b). Radical loss in the atmosphere from Cu-Fe redox coupling in aerosols. *Atmos. Chem. Phys.*, **13**, 509– 519, doi:10.5194/acp-13-509-2013.

Matthes, K., Funke, B., Anderson, M. E., Barnard, L., Beer, J., Charbonneau, P., et al. (2017). Solar Forcing for CMIP6 (v3.2). *Geosci. Model Dev.*, **10**, doi:10.5194/gmd-10-2247-2017.

678 Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., et al.
679 (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geosci.*
680 *Model Dev.*, **10**, 2057–2116, https://doi.org/10.5194/gmd-10-2057-2017.

681 Monahan, E. C., Spiel, D. E., and Davidsona, K. L. (1986). A model of marine aerosol
682 generation via whitecaps and wave disruption. *Oceanic Whitecaps,* E. C. Monahan and G.
683 Mac Niocaill, Eds., D. Reidel, 167–174.

684 Murray, L. T., Fiore, A. M., Shindell, D. T., Naik, V., and Horowitz, L.W. (submitted). Large
685 uncertainties in global hydroxyl projections tied to fate of reactive nitrogen and carbon. Proc.
686 Natl. Acad. Sci.

687 Murray, L. T., Logan, J. A., and Jacob, D. J. (2013). Interannual variability in tropical
688 tropospheric ozone and OH: the role of lightning, *J. Geophys. Res.*, **118**, 1-13,
689 doi:10.1002/jgrd.50857.

690 Naik, V., Horowitz, L. W., Fiore, A. M., Ginoux, P., Mao, J., Aghedo, A. M., and Levy, H.
691 (2013a). Impact of preindustrial to present-day changes in short-lived pollutant emissions on
692 atmospheric composition and climate forcing. *J. Geophys. Res. Atmos.*, **118**, 8086– 8110,
693 doi:10.1002/jgrd.50608.

694 Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J.-F., Lin, M., et al
695 (2013b). Preindustrial to present-day changes in tropospheric hydroxyl radical and methane
696 lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project
697 (ACCMIP), *Atmos. Chem. Phys.,* **13**, 5277–5298, https://doi.org/10.5194/acp-13-5277-2013.

698 Pai, S. J., Heald, C. L., Pierce, J. R., Farina, S. C., Marais, E. A., Jimenez, J. L., et al. (2020). An
699 evaluation of global organic aerosol schemes using airborne observations. *Atmos. Chem.*
700 *Phys.*, **20**, 2637–2665, https://doi.org/10.5194/acp-20-2637-2020.

701 Paulot, F., Jacob, D. J., Johnson, M. T., Bell, T. G., Baker, A. R., Keene, W. C., Lima, I. D.,
702 Doney, S. C., and Stock, C. A. (2015), Global oceanic emission of ammonia: Constraints
703 from seawater and atmospheric observations, *Global Biogeochem. Cycles*, **29**, 1165– 1178,
704 doi:10.1002/2015GB005106.

705 Paulot, F., Ginoux, P., Cooke, W. F., Donner, L. J., Fan, S., Lin, M. Y., et al. (2016). Sensitivity
706 of nitrate aerosols to ammonia emissions and to nitrate chemistry: implications for present
707 and future nitrate optical depth. *Atmospheric Chemistry and Physics*, **16**(3), 1459-1477.

708 Paulot, F., Fan, S., and Horowitz, L. W. (2017a). Contrasting seasonal responses of sulfate
709 aerosols to declining SO_2 emissions in the Eastern US: Implications for the efficacy of SO_2
710 emission controls. *Geophysical Research Letters*, **44**(1), 455-464.

Paulot, F., Paynter, D., Ginoux, P., Naik, V., Whitburn, S., Van Damme, M., et al. (2017b). Gas-aerosol partitioning of ammonia in biomass burning plumes: Implications for the interpretation of spaceborne observations of ammonia and the radiative forcing of ammonium nitrate. *Geophysical Research Letters*, **44**(15), 8084-8093.

Paulot, F., Malyshev, S., Nguyen, T., Crounse, J. D., Shevliakova, E., and Horowitz, L. W. (2018). Representing sub-grid scale variations in nitrogen deposition associated with land use in a global Earth System Model: implications for present and future nitrogen deposition fluxes over North America. *Atmos. Chem. Phys.*, **18**, 17,963–17,978, doi: 10.5194/acp-18-17963-2018.

Pétron, G., Crotwell, A. M., Dlugokencky, E., and Mund, J.W. (2019). Atmospheric Carbon Monoxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1988-2018, Version: 2019-08, https://doi.org/10.15138/33bv-s284.

Pickering, K. E., Wang, Y., Tao, W.-K., Price, C., and Müller, J.-F. (1998). Vertical distributions of lightning NO_x for use in regional and global chemical transport models. *J. Geophys. Res.*, **103**(D23), 31203–31216, doi:10.1029/98JD02651.

Pincus, R., Forster, P.M., and Stevens, B. (2016). The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6. *Geosci. Model Dev.*, **9**, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016.

Prather, M., Holmes, C., and Hsu, J. (2012). Reactive greenhouse gas scenarios: systematic exploration of uncertainties and the role of atmospheric chemistry. *Geophys. Res. Lett.*, **39**, L09803, doi:10.1029/2012GL051440.

Price, C., Penner, J., and Prather, M. (1997). NO_x from lightning: 1. Global distribution based on lightning physics. *J. Geophys. Res.*, **102**, 5929–5941.

Ramanathan, V., Callis, L., Cess, R., Hansen, J., Isaksen, I., Kuhn, W., et al. (1987). Climate-chemical interactions and effects of changing atmospheric trace gases. *Rev. Geophys.*, **25**(7), 1441–1482, doi:10.1029/RG025i007p01441.

Riddick, S., Dragosits, U., Blackall, T., Daunt, F., Wanless, S., and Sutton, M. (2012). The global distribution of ammonia emissions from seabird colonies. *Atmos. Environ.*, **55**, 319–327.

Rieder, H. E., Fiore, A. M., Horowitz, L. W., and Naik, V. (2015). Projecting policy-relevant metrics for high summertime ozone pollution events over the eastern United States due to climate and emission changes during the 21st century. *J. Geophys. Res.*, **120**(2), 784-800, doi: 10.1002/2014jd022303.

744 Sander, S. P., Abbatt, J., Barker, J. R., Burkholder, J. B., Friedl, R. R., Golden, D. M., et al.
745 (2011), Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies
746 Evaluation Number 17, *JPL Publication 10–6, Jet Propulsion Laboratory, Pasadena*.

747 Savoie, D. L., and Prospero, J. M. (1977). Aerosol concentration statistics for the Northern
748 Tropical Atlantic. *J. Geophys. Res.*, **82**(37), 5954–5964, doi:10.1029/JC082i037p05954.

749 Schultz, M. G., Schroder, S., Lyapina, O., Cooper, O. R., Galbally, I. Petropavlovskikh, et al.
750 (2017). Tropospheric Ozone Assessment Report: Database and metrics data of global surface
751 ozone observations. *Elementa-Sci Anthrop*, **5**, doi: 10.1525/elementa.244.

752 Schnell, J. L., Naik, V., Horowitz, L. W., Paulot, F., Mao, J., Ginoux, P., Zhao, M., and Ram, K.
753 (2018). Exploring the relationship between surface $PM_{2.5}$ and meteorology in Northern India.
754 *Atmos. Chem. Phys.*, **18**, 10157-10175, https://doi.org/10.5194/acp-18-10157-2018, 2018.

755 Shevliakova, E., Malyshev, S., Martinez Cano, I., Milly, P. C. D., Pacala, S., Ginoux, P., et al.
756 (submitted). The land component LM4.1 of the GFDL Earth system model ESM4.1:
757 biophysical and biogeochemical processes and interactions with climate. *Journal of*
758 *Advances in Modeling Earth Systems*.

759 Smirnov, A., Holben, B., Eck, T., Dubovik, O., & Slutsker, I. (2000). Cloud-screening and
760 quality control algorithms for the AERONET database. *Remote Sensing of Environment*, **73**,
761 337–349.

762 Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones,
763 D. B. A. et al. (2000). Three-dimensional climatological distribution of tropospheric OH:
764 Update and evaluation, *J. Geophys. Res.*, **105**, 8,931– 8,980.

765 Taylor, K.E., Williamson, D., and Zwiers, F. (2000). The sea surface temperature and sea ice
766 concentration boundary conditions for AMIP II simulations. *PCMDI Report 60*, Program for
767 Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory,
768 25 pp. Available online: https://pcmdi.llnl.gov/report/pdf/60.pdf.

769

770 Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., et al. (2016). Why do
771 models overestimate surface ozone in the Southeast United States? *Atmos. Chem. Phys.*, **16**,
772 13,561-13,577, doi: 10.5194/acp-16-13561-2016.

773 van Marle, M. J. E., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A.-L., Field, R. D., et al.
774 (2017). Historic global biomass burning emissions based on merging satellite observations
775 with proxies and fire models (1750 - 2015). *Geoscientific Model Development,* **10**, 3329-
776 3357. doi:10.5194/gmd-2017-32.

Wild, O., Voulgarakis, A., O'Connor, F., Lamarque, J.-F., Ryan, E. M., and Lee, L. (2020). Global sensitivity analysis of chemistry-climate model budgets of tropospheric ozone and OH: Exploring model diversity, *Atmos. Chem. Phys,* **20***,* 4047-4058, https://doi.org/10.5194/acp-20-4047-2020.

Williams, R. M. (1982). A model for the dry deposition of particles to natural water surfaces. *Atmos. Environ.*, **16**, 1933–1938, 1982.

Young, P. J., Naik, V., Fiore, A. M., Gaudel, A., Guo, J., Lin, M. Y., et al. (2018). Tropospheric Ozone Assessment Report: Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends. *Elementa-Sci Anthrop*, **6**, doi: 10.1525/elementa.265.

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., et al. (2009). Asian emissions in 2006 for the NASA INTEX-B mission. *Atmos. Chem. Phys.*, **9**, 5131–5153, https://doi.org/10.5194/acp-9-5131-2009.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018a). 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.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018b). The GFDL global atmosphere and land model AM4.0/LM4.0: 2. Model description, sensitivity studies, and tuning strategies. *J. Adv. Model. Earth Syst.*, **10**, https://doi.org/10.1002/2017MS001209.

Zheng, B., Zhang, Q., Zhang, Y., He, K. B., Wang, K., Zheng, G. J., Duan, F. K., Ma, Y. L., and Kimoto, T. (2015). Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic aerosol formation during the January 2013 haze episode in North China. *Atmos. Chem. Phys.*, **15**, 2031.

Figure captions

Figure 1. Schematic description of forcing, dynamics, physics, aerosol, and chemistry interactions in AM4.1. Terms depicted in gray (left) are prescribed as inputs to the model, while chemical processes included in the orange box are calculated interactively within the atmospheric model. The light blue box (top) includes physical processes calculated in AM4.1. The green box (bottom left) represents the land component (LM4.1), which is coupled to AM4.1. The dark blue box (bottom right) includes specified ocean surface boundary conditions.

Figure 2. Global annual totals (in Tg a^{-1}, using mass as indicated on y-axis label) for anthropogenic (fossil fuel + biomass burning + ship + aircraft) emissions of NO, CO, SO_2, NH_3, BC, and primary OM in AM3 (CMIP5 emissions) and AM4.1 (CMIP6 emissions) AMIP simulations.

Figure 3. Annual mean surface air temperatures (°C) in AM4.1 AMIP simulation (1980-2014) and CRU-TS-3.22 observations (1979-2013). Differences between simulated and observed surface air temperatures in AM4.1, AM4.0, and AM3 AMIP simulations.

Figure 4. Annual mean precipitation (mm day^{-1}) for 1980-2014 in AM4.1 AMIP simulation and GPCP v2.3 observations. Differences between simulated and observed precipitation in AM4.1, AM4.0, and AM3 AMIP simulations.

Figure 5. Annual mean zonal mean zonal wind (m s^{-1}) in AM4.1 AMIP simulation (1980-2014) and ERA40 reanalysis (1981-2000). Differences between simulated and observed zonal winds in AM4.1, AM4.0, and AM3 AMIP simulations.

Figure 6. Monthly and annual (ANN) stratospheric sudden warming (SSW) frequency for 1870–2014 from AM4.0 and AM4.1, and 1957–2002 from ERA40. SSW is defined as in Charlton and Polvani (2007). Error bars indicate the 95% confidence interval (the statistical test of the SSW frequency is calculated as in Charlton et al., 2007).

Figure 7. Annual mean net radiation flux at top of atmosphere (W m^{-2}) in AM4.1 AMIP simulation (1980-2014) and CERES EBAF v2.8 observations (2000-2015). Differences between simulated and observed net radiation flux in AM4.1, AM4.0, and AM3 AMIP simulations.

Figure 8. MAM mean surface MDA8 ozone mixing ratios (ppbv) for 2000-2008 from: (a) TOAR observations regridded to the same 1° x 1° grid as AM4.1, (b) AM3 AMIP simulation, (c) AM4.1 AMIP simulation. Here, mn is the mean and rmsd is the root-mean-square deviation between observations and simulations.

Figure 9. Same as Figure 8, but for JJA.

Figure 10. Climatological mean tropospheric ozone column in AM3 (upper left; Dobson Units, DU), AM4.1 (lower left; DU), and the % bias compared to the OMI/MLS satellite estimate of the Tropospheric Ozone Column (Ziemke et al., 2019) for AM3 (upper right; %) and AM4.1 (lower right; %). RMSE is provided in DU.

Figure 11. Comparison of time series of total ozone column (DU) for the annual mean (a) global mean (90°S-90°N), (b) tropics (25°S-25°N), (c) northern mid-latitudes (35°N-60°N), (d) southern mid-latitudes (35°S-35°N), and for the (e) March mean in the Arctic (60°N-90°N) and (f) October mean in the Antarctic (60°S-90°S) from AM3 (red) and AM4.1 (blue) against NASA and NOAA observations from the multi-satellite merged ozone total column [Frith et al., 2013] (SBUV; open triangles) and version 3.4 of the NIWA-BS total column ozone database [Bodeker et al., 2005] (NIWA; closed circles). The numbers in each panel indicate linear correlation coefficient (R) for model against each of the measurement datasets (top for NIWA and bottom for SBUV).

Figure 12. Absolute difference in tropospheric CO column between AM3 and MOPITT (left panel) and AM4.1 and MOPITT (right panel) for winter (December-January-February, DJF, top) and summer (June-July-August, JJA, bottom).

Figure 13. Comparison of surface CO mixing ratios (ppbv) from AM4.1 (red) and AM3 (blue) against NOAA Global Monitoring Division (GMD) flask observations (Pétron et al., 2019, for 1988-2014). Left panels show model bias (top) and correlation coefficient (bottom) versus observations, plotted by station latitude. Right panels show monthly time series comparisons at selection stations. The root mean square error (RMSE) and correlation coefficient (R) are indicated on plots.

Figure 14. Comparison of AM4.1 (2000-2014) against IMPROVE (a,b) and NADP (c,d) observations of concentrations in surface air (top) and in precipitation (bottom) of sulfate (left) and nitrate (right).

Figure 15. Comparison of simulated (AM3, 1979-2008; AM4.0, 1980-2014; and AM4.1, 1980-2014) and observed (University of Miami) annual mean surface concentrations ($\mu g\ m^{-3}$) of (first row) sulfate, (second row) dust, and (third row) sea salt sodium at 28 locations and (bottom) their ratios (simulated/observed) at each location (for AM4.1 only). Shaded contours indicate simulated surface concentrations (top colorbar) and symbols indicate the ratio of simulated/observed concentrations (bottom colorbar, symbol points upwards if ratio greater than one, downwards if less than one).

Figure 16. Comparison of simulated aerosol optical depths (550 nm) with AERONET observations over the 2000-2014 period for (top) AM4.1, (middle) AM4.0, and (bottom) AM3 AMIP simulation. Dashed lines in left panels denote slopes of 0.5 and 2. Color in right panels

868 shows the percentage difference between model and AERONET (i.e., 100%× [model-
869 AERONET]/AERONET).

870 **Figure 17.** Monthly climatology (2003–2014) of aerosol optical depth simulated by AM3
871 (purple line), AM4.0 (green line) and AM4.1 (orange line) and measured by MODIS (TERRA:
872 star, AQUA: cross) and MISR (filled circles) satellite instruments. Each panel represents a
873 spatial average over the corresponding region on the background map. The numbers in each box
874 show the correlation coefficients (left) and normalized root mean square error (right) compared
875 to MODIS-TERRA (purple: AM3, green: AM4.0, orange: AM4.1).

876 **Figure 18.** Climatological (1980-2014) annual mean airmass-weighted tropospheric OH
877 concentration averaged globally (top-most row) and regionally for individual atmospheric
878 subdomains from AM4.1 (black) compared with those from AM3 (1980-2008, red), ACCMIP
879 ensemble mean (orange), and climatological mean values from Spivakovsky et al. (2000)
880 (purple). Values for AM4.1 and AM3 also show +/- standard deviation about the mean.

881 **Tables**

882 **Table 1.** Annual total emissions for year 2014 in AM4.1.

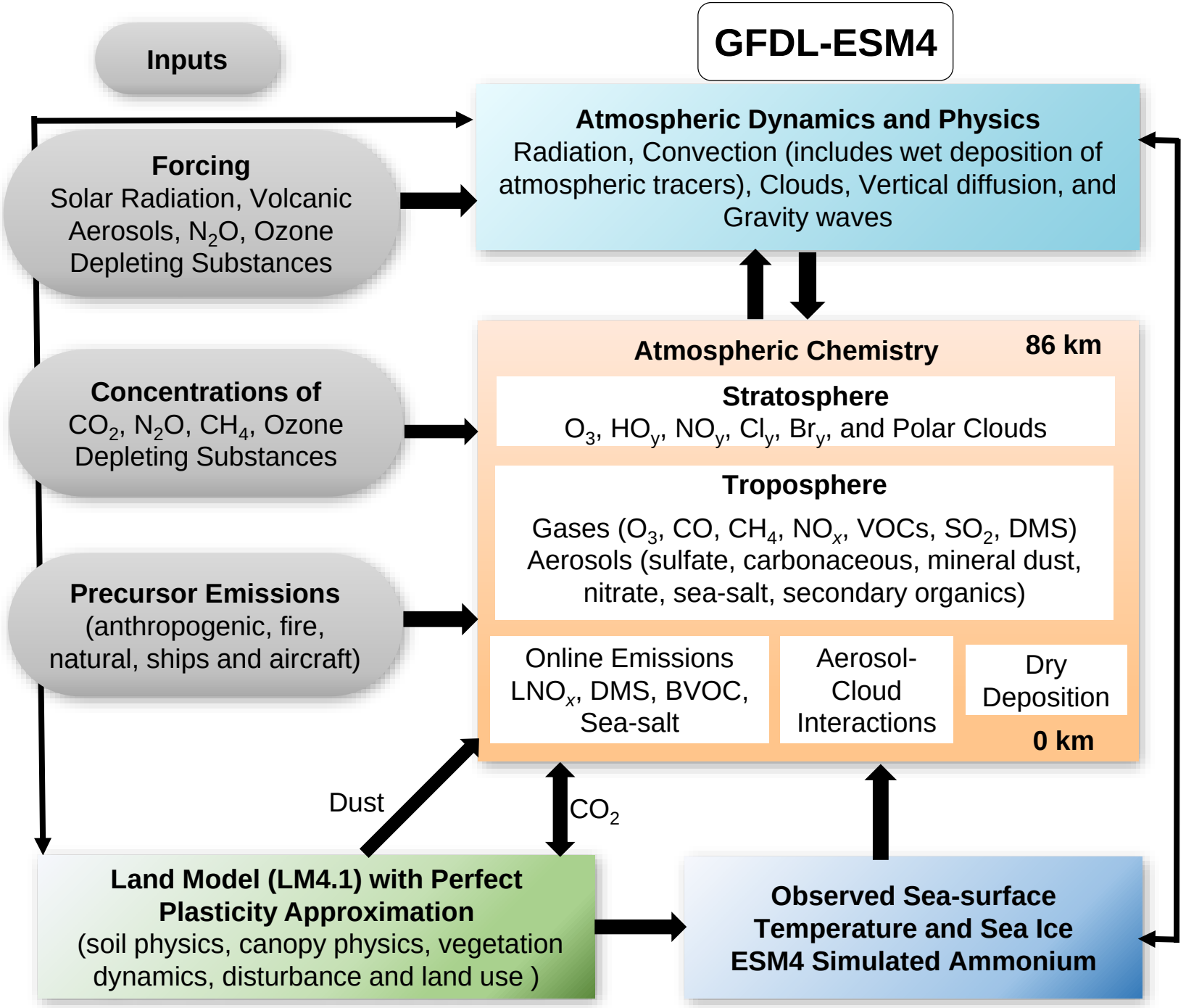
Species	Units	anthro	biomass burning	biogenic / natural	ocean	animals	soil	ship	aircraft	Total
ACET	TgC/yr	1.47	0.98	15.09	0	0	0	0	0	17.53
BC	TgC/yr	7.83	1.77	0	0	0	0	0.17	0	9.76
C2H4	TgC/yr	4.88	3.82	0	0	0	0	0.14	0	8.83
C2H5OH	TgC/yr	2.40	0.07	4.82	0	0	0	0	0	7.29
C2H6	TgC/yr	5.22	2.71	0.80	0.78	0	0	0.17	0	9.67
C3H6	TgC/yr	9.50	5.85	0.85	1.29	0	0	0.16	0	17.66
C3H8	TgC/yr	5.05	0.53	1.63	1.05	0	0	0.49	0	8.76
C4H10	TgC/yr	52.93	2.34	0	0	0	0	1.10	0	56.38
C10H16	TgC/yr	0	1.24	57.37	0	0	0	0	0	58.61
CH2O	TgC/yr	1.00	1.94	0	0	0	0	0	0	2.94
CH3OH	TgC/yr	0.30	3.24	85.61	0	0	0	0	0	89.14
CO	Tg/yr	612.40	356.68	159.24	19.80	0	0	0.69	0.57	1149.37
DMS	Tg/yr	0	0	0	42.72	0	0	0	0	42.72
DUST	Tg/yr	0	0	0	0	0	2507.67	0	0	2507.67
H2	Tg/yr	24.50	9.01	0	2.98	0	2.98	0.03	0	39.48
ISOP	Tg/yr	0.00	0.57	499.78	0	0	0	0.00	0	500.36
NH3	Tg/yr	60.82	4.30	0	3.89	0.15	2.95	0.02	0	72.13
NO	TgN/yr	35.52	6.23	3.29	0	0	3.59	6.89	0.93	56.45
OM	Tg/yr	31.26	26.81	0	16.21	0	0	0.20	0	74.48
SSALT	Tg/yr	0	0	0	6254.24	0	0	0	0	6254.24
SO2	TgS/yr	51.26	1.14	3.59	0	0	0	4.44	0.14	60.56

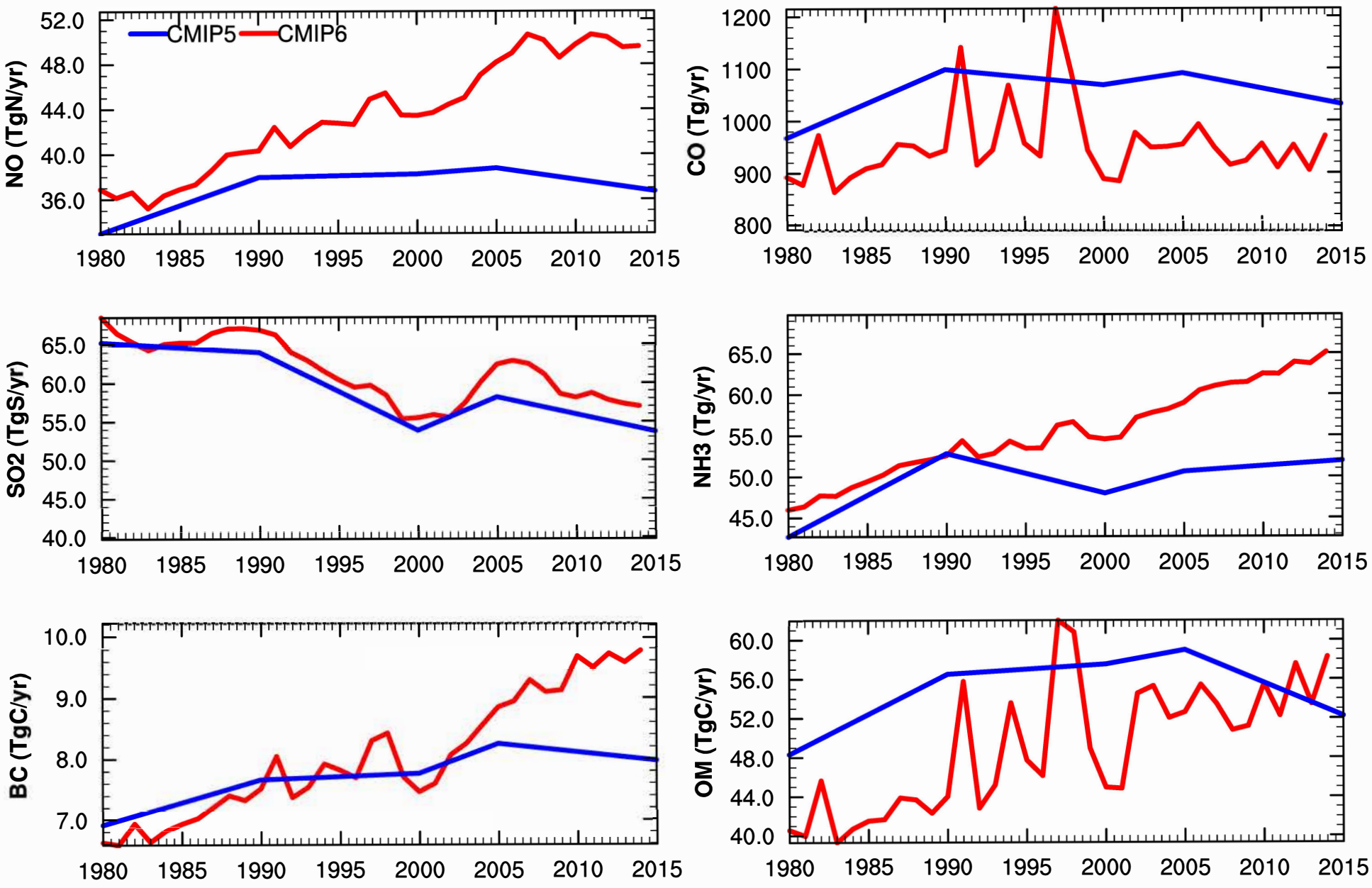
883

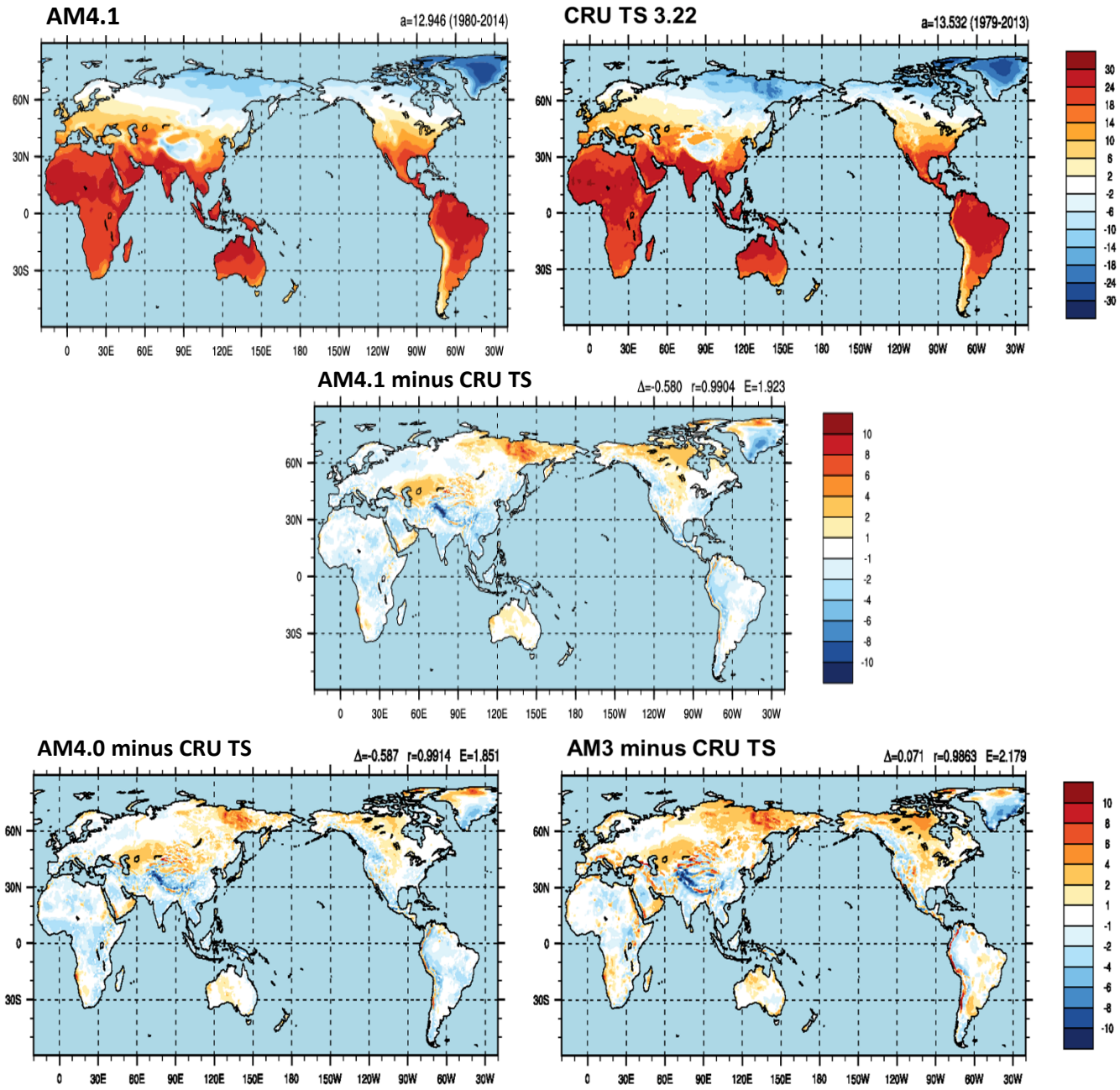
884 **Table 2.** Effective radiative forcings and feedbacks (in W m^{-2}) to greenhouse gases, aerosols,
885 anthropogenic forcings, land use, quadrupling CO_2, and SST perturbations. The AM3 (1990) and
886 AM4.0 (1990) experiments are identical to those described by Zhao et al. (2018a). The AM4.0
887 (2014) and AM4.1 (2014) experiments use 30-year climatological simulations, i.e., the piClim
888 experiments from the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al.,
889 2016).

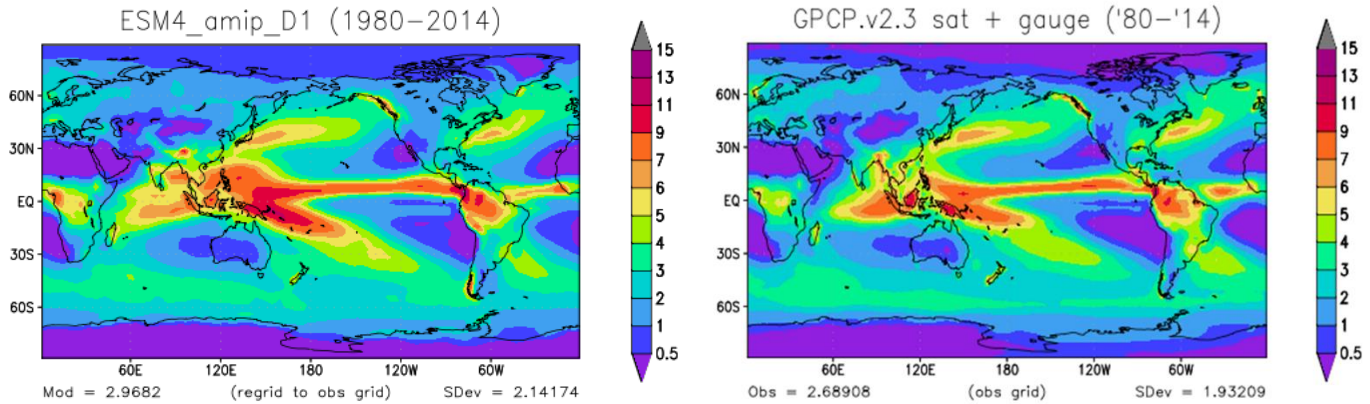
	AM3 (1990)	AM4.0 (1990)	AM4.0 (2014)	AM4.1 (2014)
GHG ERF	2.63	2.61	3.14	3.22
Aerosol ERF	-1.69	-0.96	-0.73	-0.70
Anthro ERF	N/A	N/A	2.33	2.37
LU ERF	N/A	N/A	-0.33	-0.28
4xCO2 ERF	7.19	N/A	8.23	7.72
Cess feedback (SST +2K)	-2.86	-3.59	-3.64	-4.14

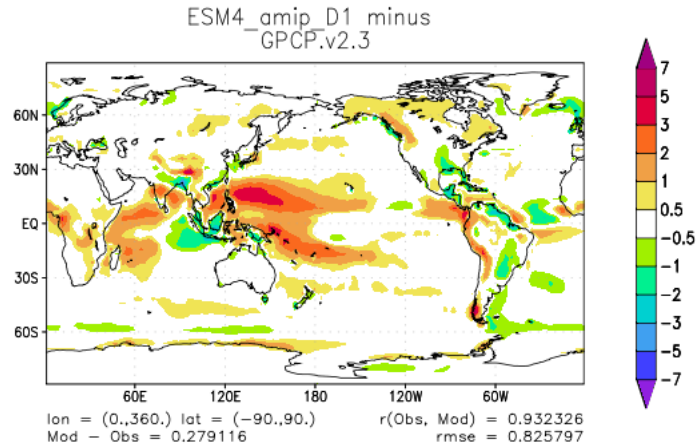
890

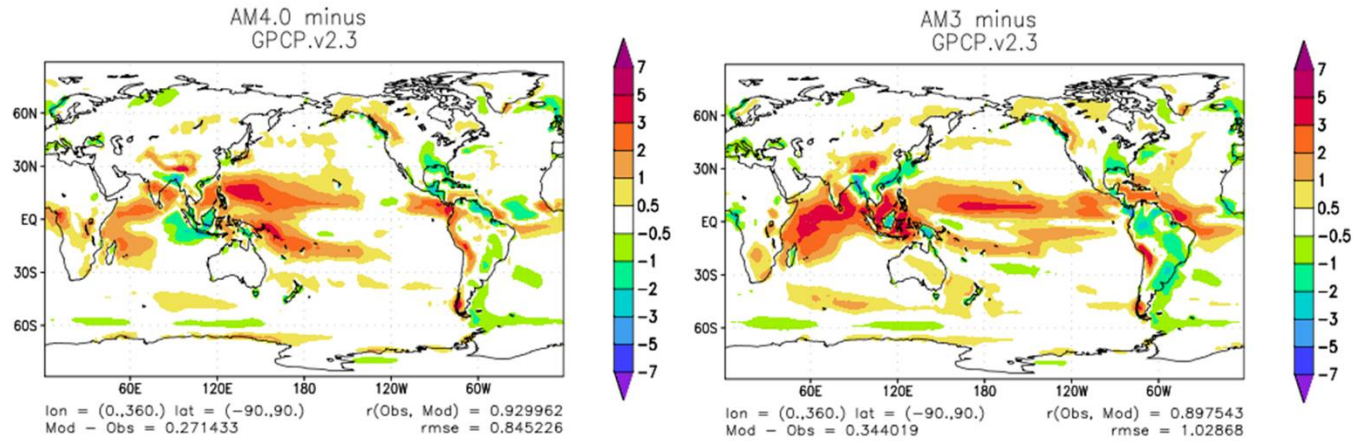

GFDL-ESM4
Inputs
Forcing
Solar Radiation, Volcanic Aerosols, N_2O, Ozone Depleting Substances
Concentrations of
CO_2, N_2O, CH_4, Ozone Depleting Substances
Precursor Emissions
(anthropogenic, fire, natural, ships and aircraft)
Atmospheric Dynamics and Physics
Radiation, Convection (includes wet deposition of atmospheric tracers), Clouds, Vertical diffusion, and Gravity waves
Atmospheric Chemistry
86 km
Stratosphere
O_3, HO_y, NO_y, Cl_y, Br_y, and Polar Clouds
Troposphere
Gases (O_3, CO, CH_4, NO_x, VOCs, SO_2, DMS)
Aerosols (sulfate, carbonaceous, mineral dust, nitrate, sea-salt, secondary organics)
Online Emissions
LNO_x, DMS, BVOC, Sea-salt
Aerosol-Cloud Interactions
Dry Deposition
0 km
Dust
CO_2
Land Model (LM4.1) with Perfect Plasticity Approximation
(soil physics, canopy physics, vegetation dynamics, disturbance and land use)
Observed Sea-surface Temperature and Sea Ice
ESM4 Simulated Ammonium

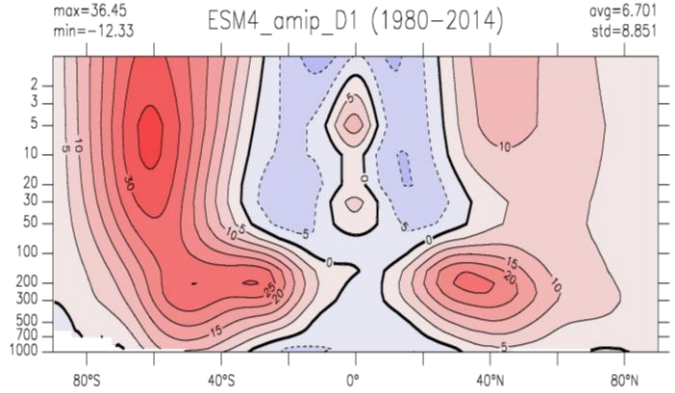

CMIP5
CMIP6
NO (TgN/yr)
52.0
48.0
44.0
40.0
36.0
CO (Tg/yr)
1200
1100
1000
900
800
SO2 (TgS/yr)
65.0
60.0
55.0
50.0
45.0
40.0
NH3 (Tg/yr)
65.0
60.0
55.0
50.0
45.0
BC (TgC/yr)
10.0
9.0
8.0
7.0
OM (TgC/yr)
60.0
56.0
52.0
48.0
44.0
40.0
1980
1985
1990
1995
2000
2005
2010
2015

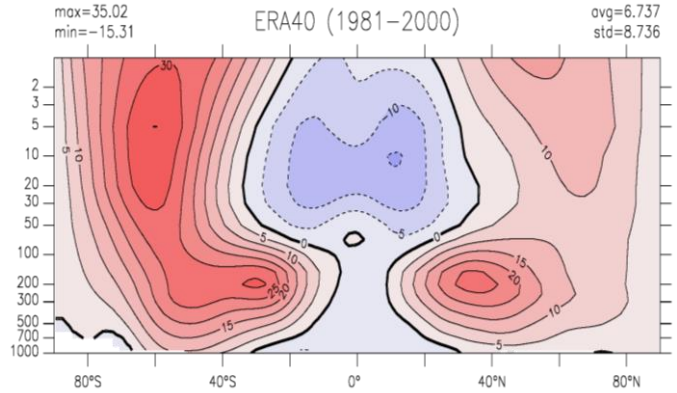

AM4.1
a=12.946 (1980-2014)
CRU TS 3.22
a=13.532 (1979-2013)
30
24
18
14
10
6
2
-2
-6
-10
-14
-18
-24
-30
AM4.1 minus CRU TS
Δ=-0.580 r=0.9904 E=1.923
10
8
6
4
2
1
-1
-2
-4
-6
-8
-10
AM4.0 minus CRU TS
Δ=-0.587 r=0.9914 E=1.851
AM3 minus CRU TS
Δ=0.071 r=0.9863 E=2.179
60N
30N
0
30S
0 30E 60E 90E 120E 150E 180 150W 120W 90W 60W 30W

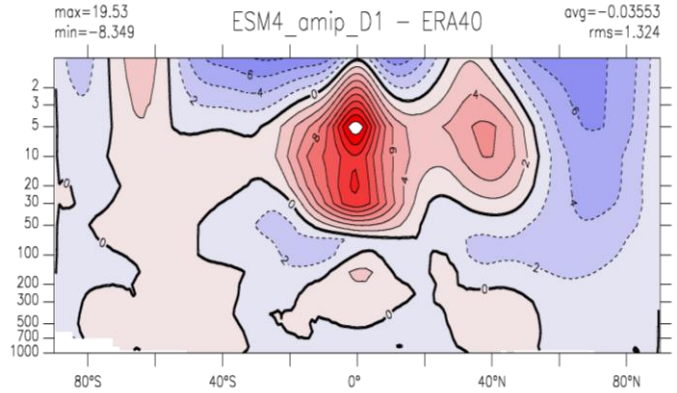

ESM4_amip_D1 (1980–2014)
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
15
13
11
9
7
6
5
4
3
2
1
0.5
Mod = 2.9682
(regrid to obs grid)
SDev = 2.14174
GPCP.v2.3 sat + gauge ('80–'14)
Obs = 2.68908
(obs grid)
SDev = 1.93209

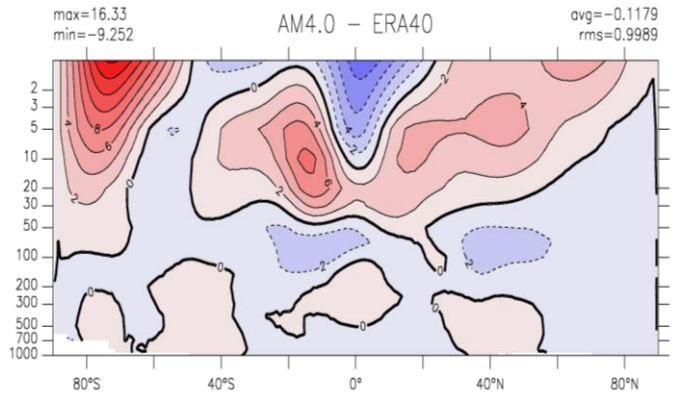

ESM4_amip_D1 minus
GPCP.v2.3
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
7
5
3
2
1
0.5
−0.5
−1
−2
−3
−5
−7
lon = (0.,360.) lat = (−90.,90.)
Mod − Obs = 0.279116
r(Obs, Mod) = 0.932326
rmse = 0.825797

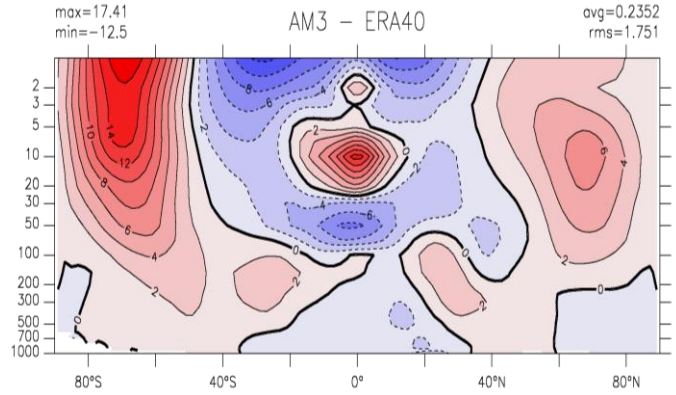

AM4.0 minus
GPCP.v2.3
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
7
5
3
2
1
0.5
−0.5
−1
−2
−3
−5
−7
lon = (0.,360.) lat = (−90.,90.)
Mod − Obs = 0.271433
r(Obs, Mod) = 0.929962
rmse = 0.845226
AM3 minus
GPCP.v2.3
lon = (0.,360.) lat = (−90.,90.)
Mod − Obs = 0.344019
r(Obs, Mod) = 0.897543
rmse = 1.02868

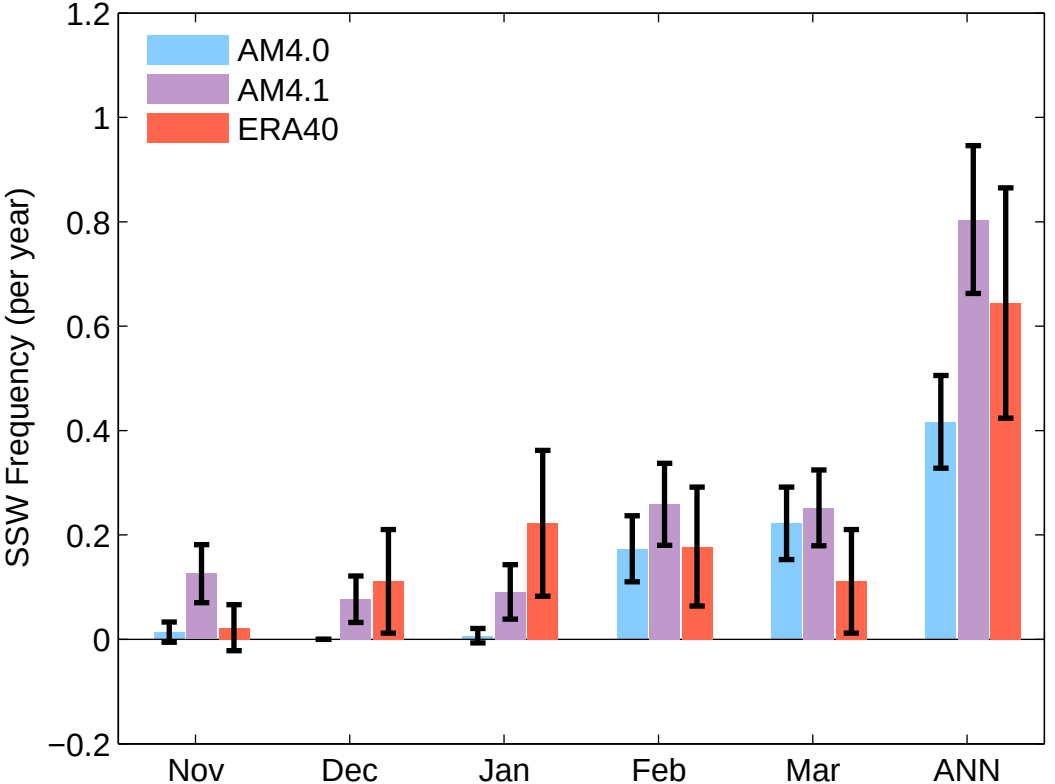

max=36.45
min=−12.33
ESM4_amip_D1 (1980−2014)
avg=6.701
std=8.851
2
3
5
10
20
30
50
100
200
300
500
700
1000
80°S
40°S
0°
40°N
80°N

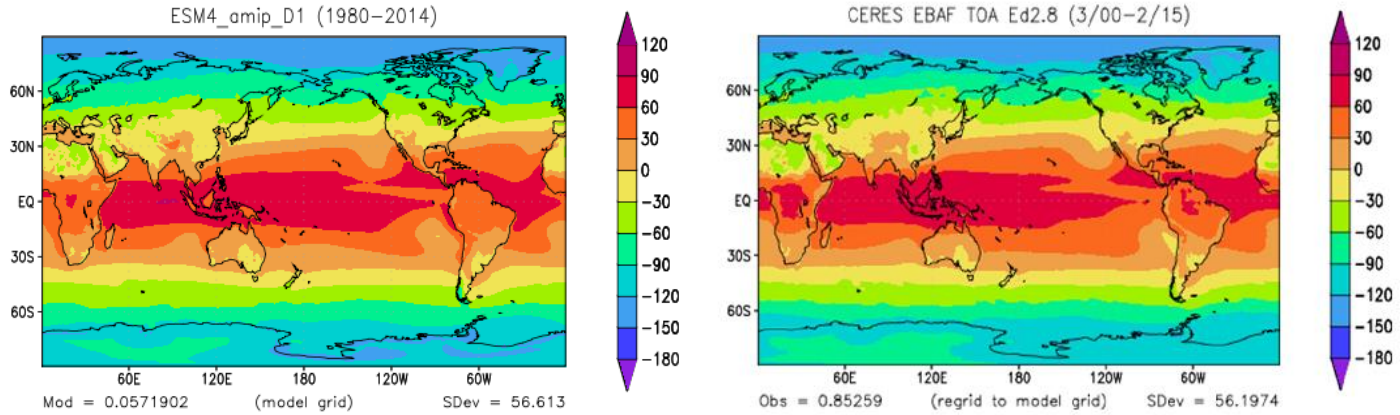

max=35.02
min=−15.31
ERA40 (1981−2000)
avg=6.737
std=8.736
2
3
5
10
20
30
50
100
200
300
500
700
1000
80°S
40°S
0°
40°N
80°N

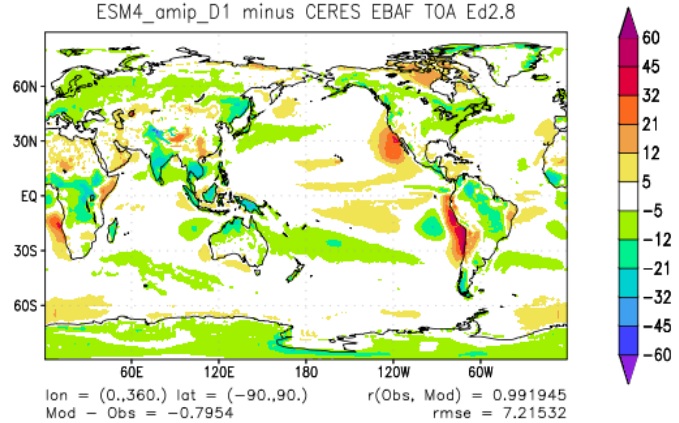

max=19.53
min=−8.349
ESM4_amip_D1 − ERA40
avg=−0.03553
rms=1.324
2
3
5
10
20
30
50
100
200
300
500
700
1000
80°S
40°S
0°
40°N
80°N

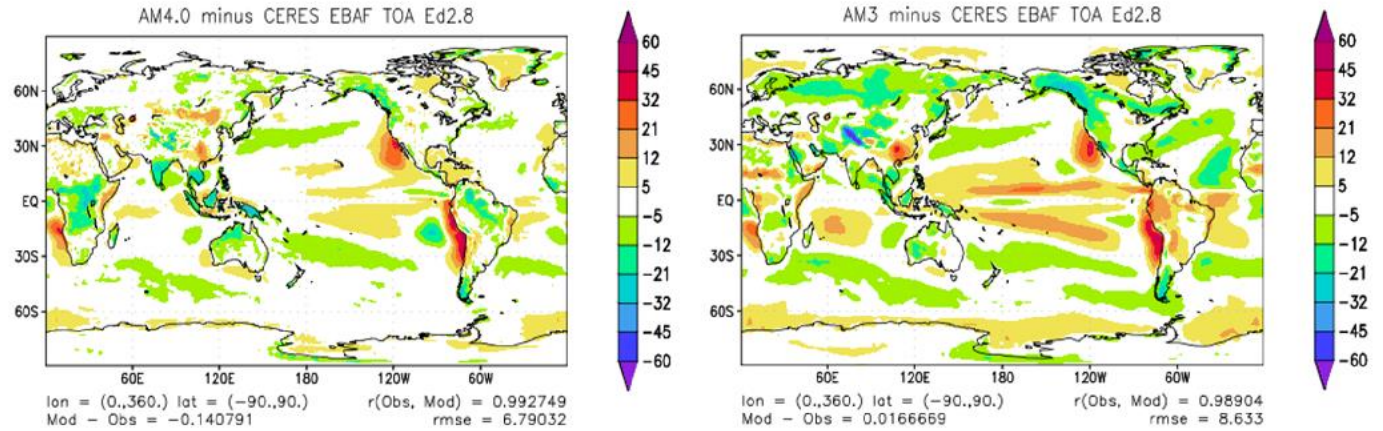

max=16.33
min=−9.252
AM4.0 − ERA40
avg=−0.1179
rms=0.9989
2
3
5
10
20
30
50
100
200
300
500
700
1000
80°S
40°S
0°
40°N
80°N

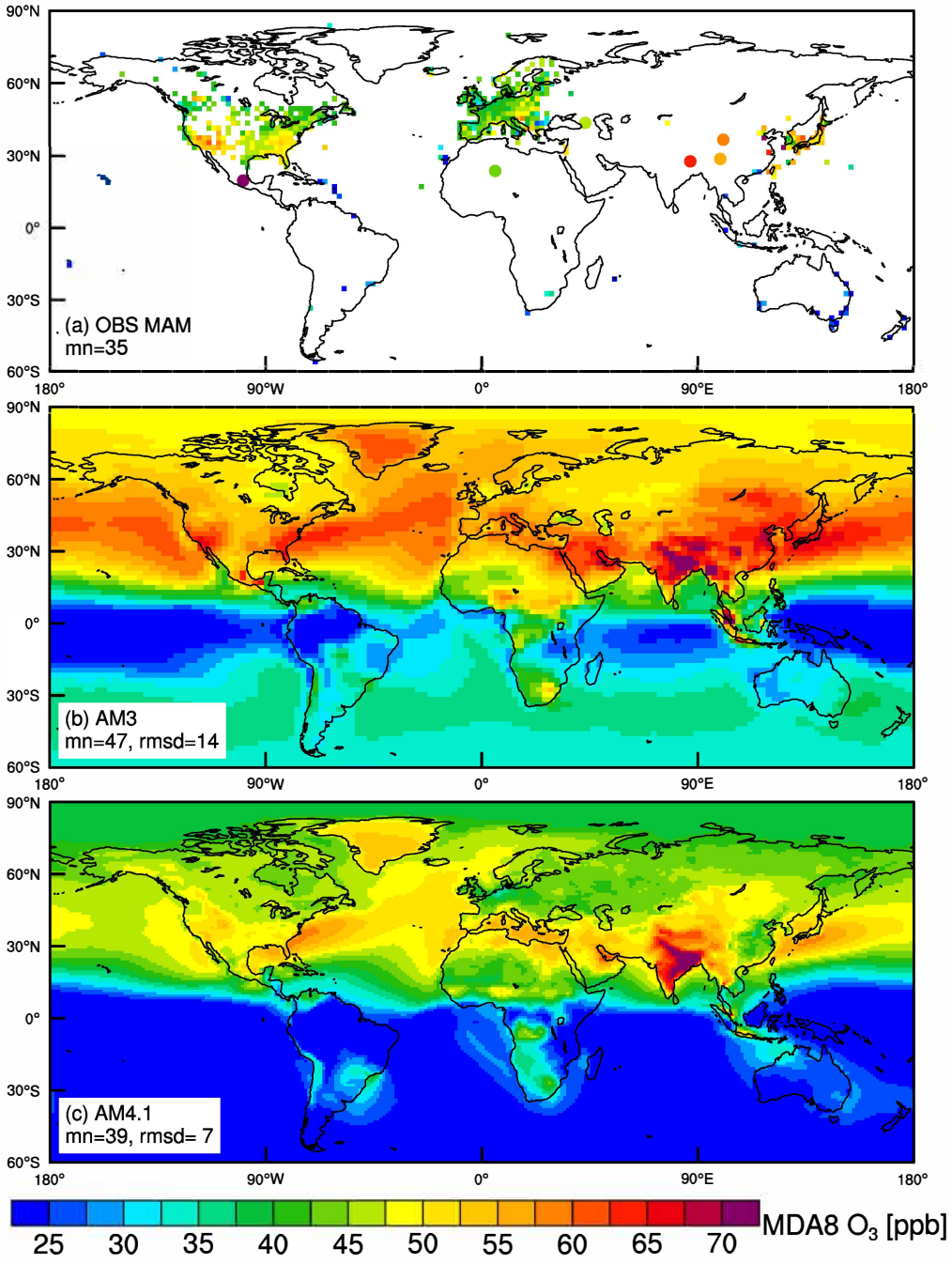

max=17.41
min=−12.5
AM3 − ERA40
avg=0.2352
rms=1.751
2
3
5
10
20
30
50
100
200
300
500
700
1000
80°S
40°S
0°
40°N
80°N

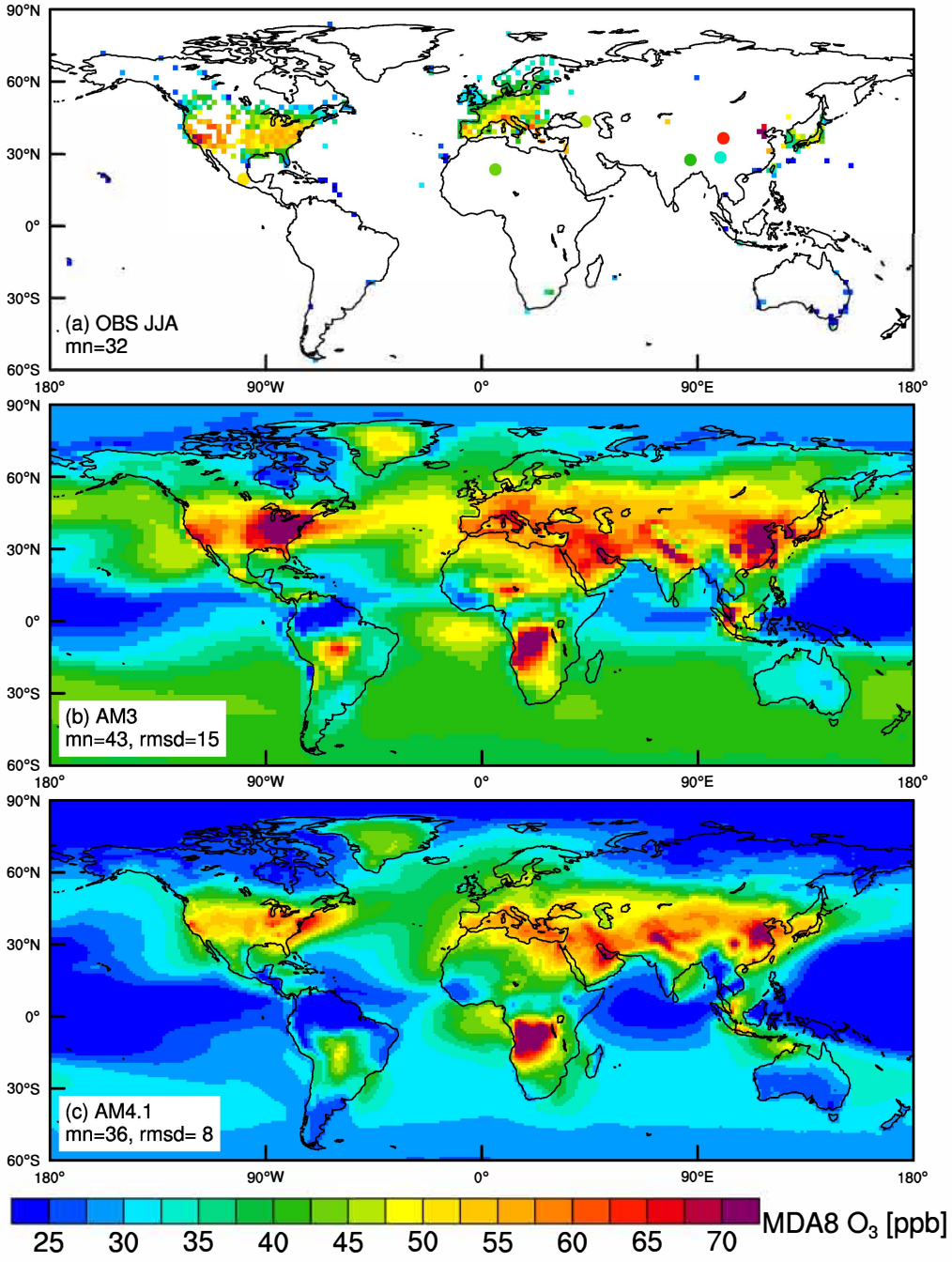

AM4.0
AM4.1
ERA40
SSW Frequency (per year)
1.2
1
0.8
0.6
0.4
0.2
0
−0.2
Nov
Dec
Jan
Feb
Mar
ANN


ESM4_amip_D1 (1980–2014)
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
120
90
60
30
0
−30
−60
−90
−120
−150
−180
Mod = 0.0571902 (model grid) SDev = 56.613
CERES EBAF TOA Ed2.8 (3/00–2/15)
Obs = 0.85259 (regrid to model grid) SDev = 56.1974

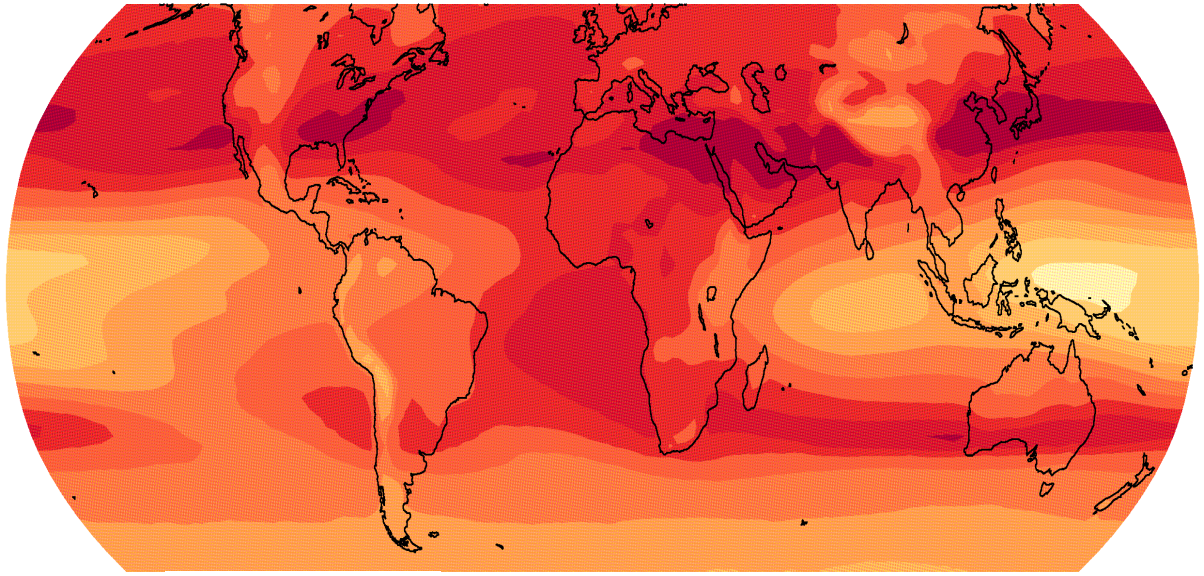

ESM4_amip_D1 minus CERES EBAF TOA Ed2.8
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
60
45
32
21
12
5
−5
−12
−21
−32
−45
−60
lon = (0.,360.) lat = (−90.,90.) r(Obs, Mod) = 0.991945
Mod − Obs = −0.7954 rmse = 7.21532


AM4.0 minus CERES EBAF TOA Ed2.8
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
60
45
32
21
12
5
−5
−12
−21
−32
−45
−60
lon = (0.,360.) lat = (−90.,90.) r(Obs, Mod) = 0.992749
Mod − Obs = −0.140791 rmse = 6.79032
AM3 minus CERES EBAF TOA Ed2.8
lon = (0.,360.) lat = (−90.,90.) r(Obs, Mod) = 0.98904
Mod − Obs = 0.0166669 rmse = 8.633

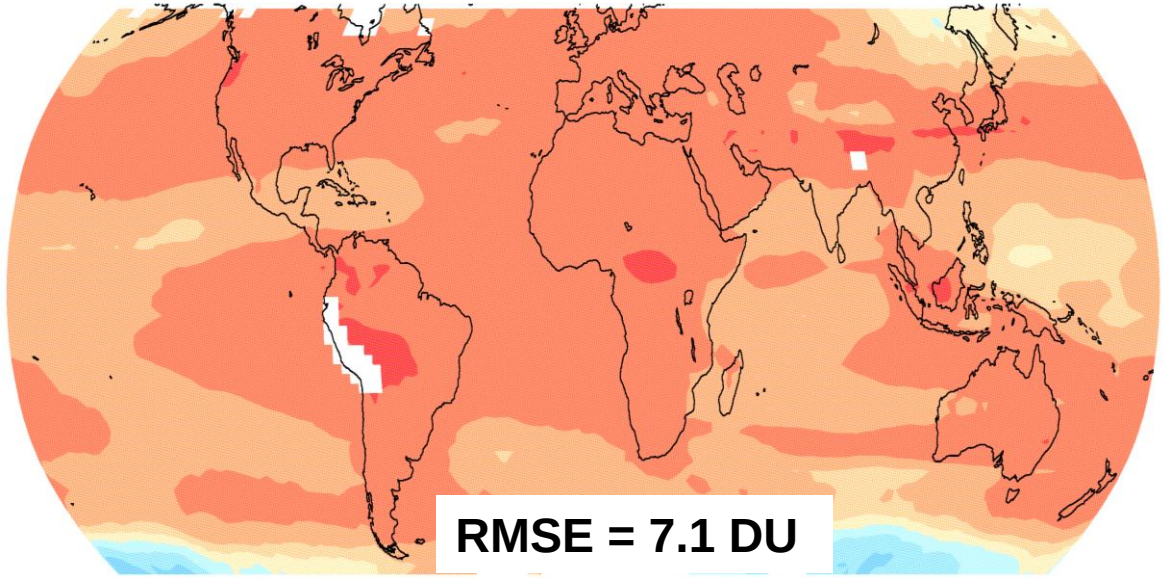

90°N
60°N
30°N
0°
30°S
60°S
180°
90°W
0°
90°E
180°
(a) OBS MAM
mn=35
(b) AM3
mn=47, rmsd=14
(c) AM4.1
mn=39, rmsd= 7
25
30
35
40
45
50
55
60
65
70
MDA8 O3 [ppb]


90°N
60°N
30°N
0°
30°S
60°S
180°
90°W
0°
90°E
180°
(a) OBS JJA
mn=32
(b) AM3
mn=43, rmsd=15
(c) AM4.1
mn=36, rmsd= 8
25
30
35
40
45
50
55
60
65
70
MDA8 O3 [ppb]

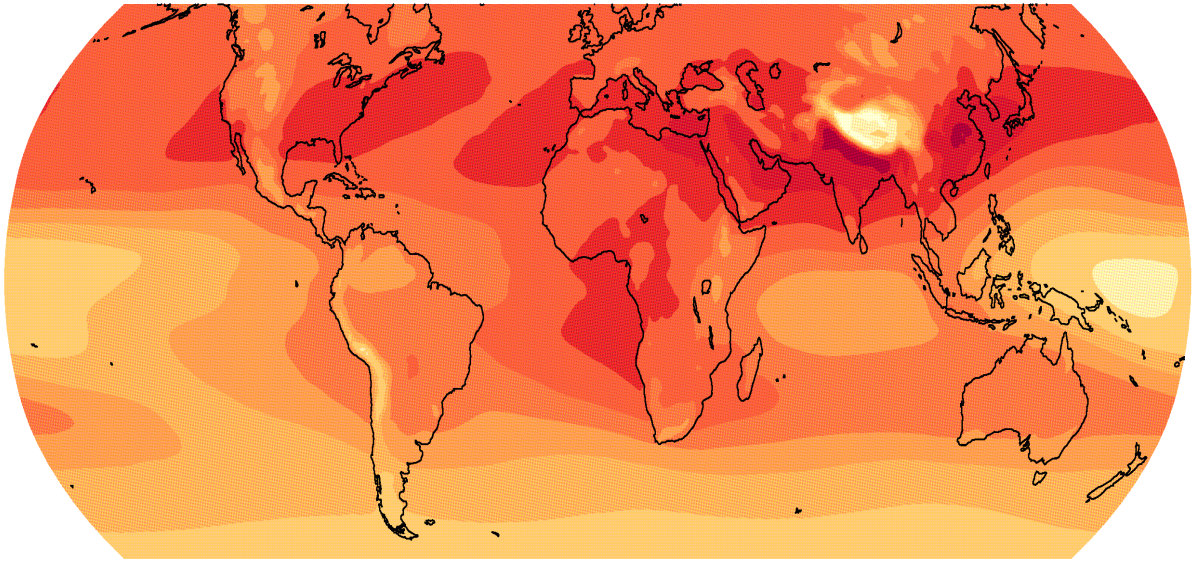
2000-2008 AM3 DU

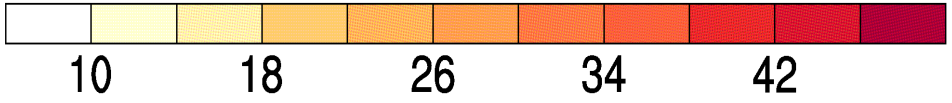


% Bias = AM3 – OMI/MLS

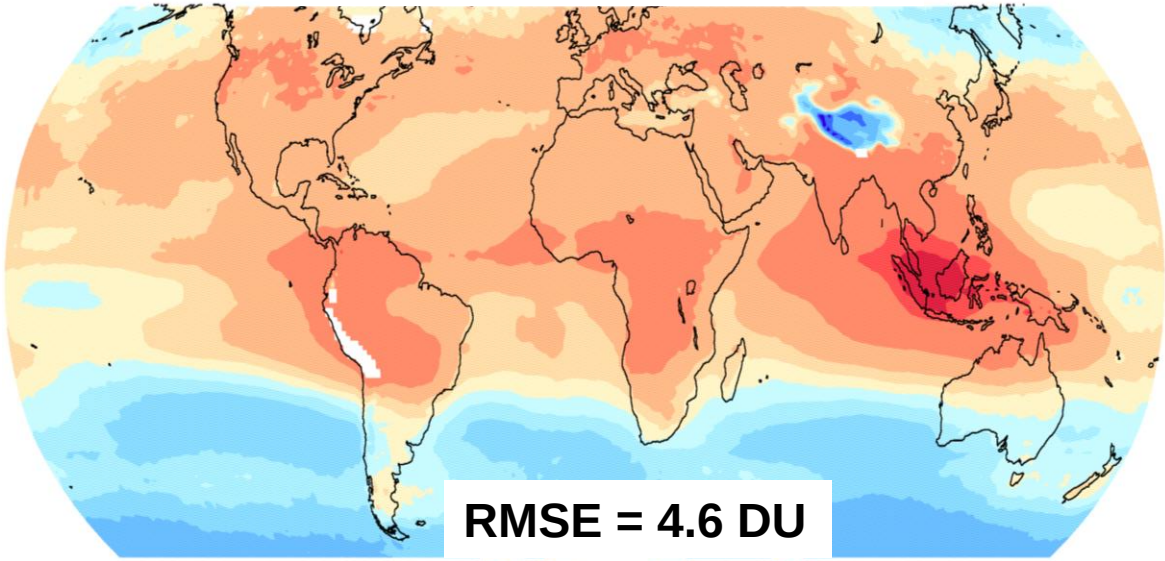


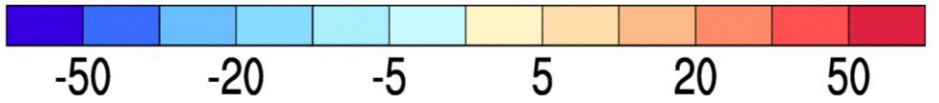
2005-2014 AM4.1 DU

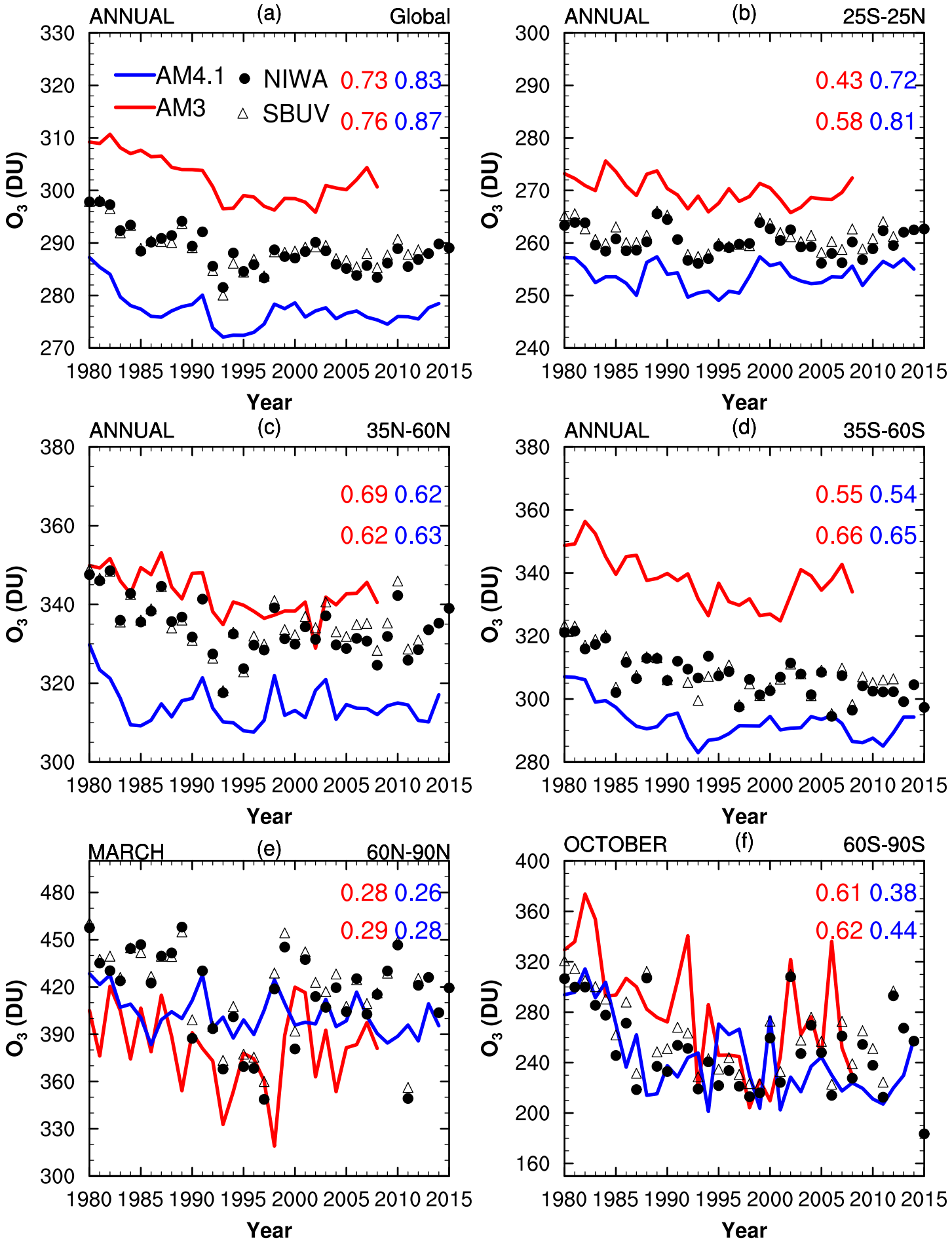




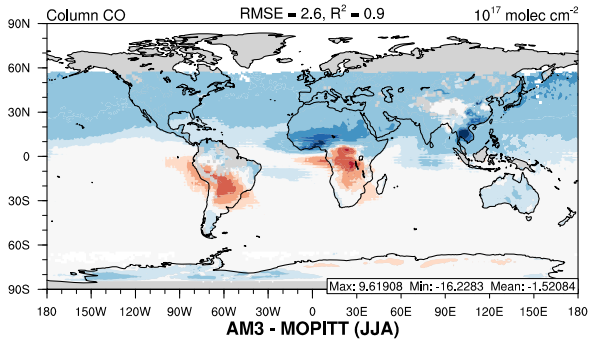
% Bias = AM4.1 – OMI/MLS



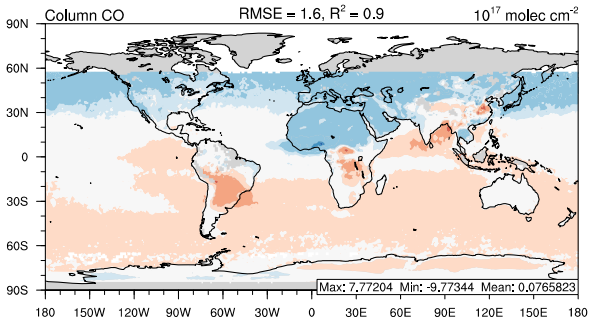



ANNUAL
(a)
Global
AM4.1
AM3
NIWA
SBUV
0.73 0.83
0.76 0.87
O3 (DU)
Year
ANNUAL
(b)
25S-25N
0.43 0.72
0.58 0.81
ANNUAL
(c)
35N-60N
0.69 0.62
0.62 0.63
ANNUAL
(d)
35S-60S
0.55 0.54
0.66 0.65
MARCH
(e)
60N-90N
0.28 0.26
0.29 0.28
OCTOBER
(f)
60S-90S
0.61 0.38
0.62 0.44

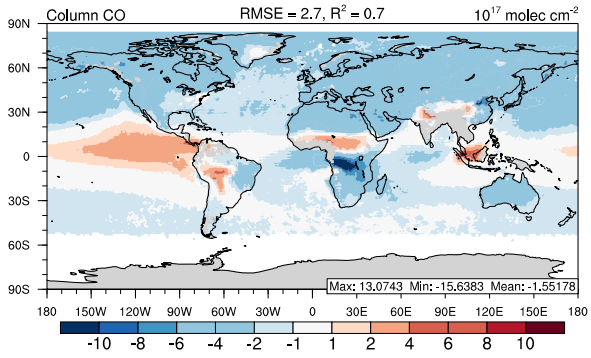
AM3 - MOPITT (DJF)



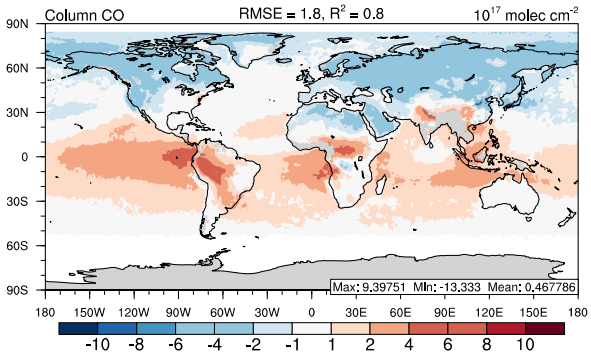
AM4.1 - MOPITT (DJF)

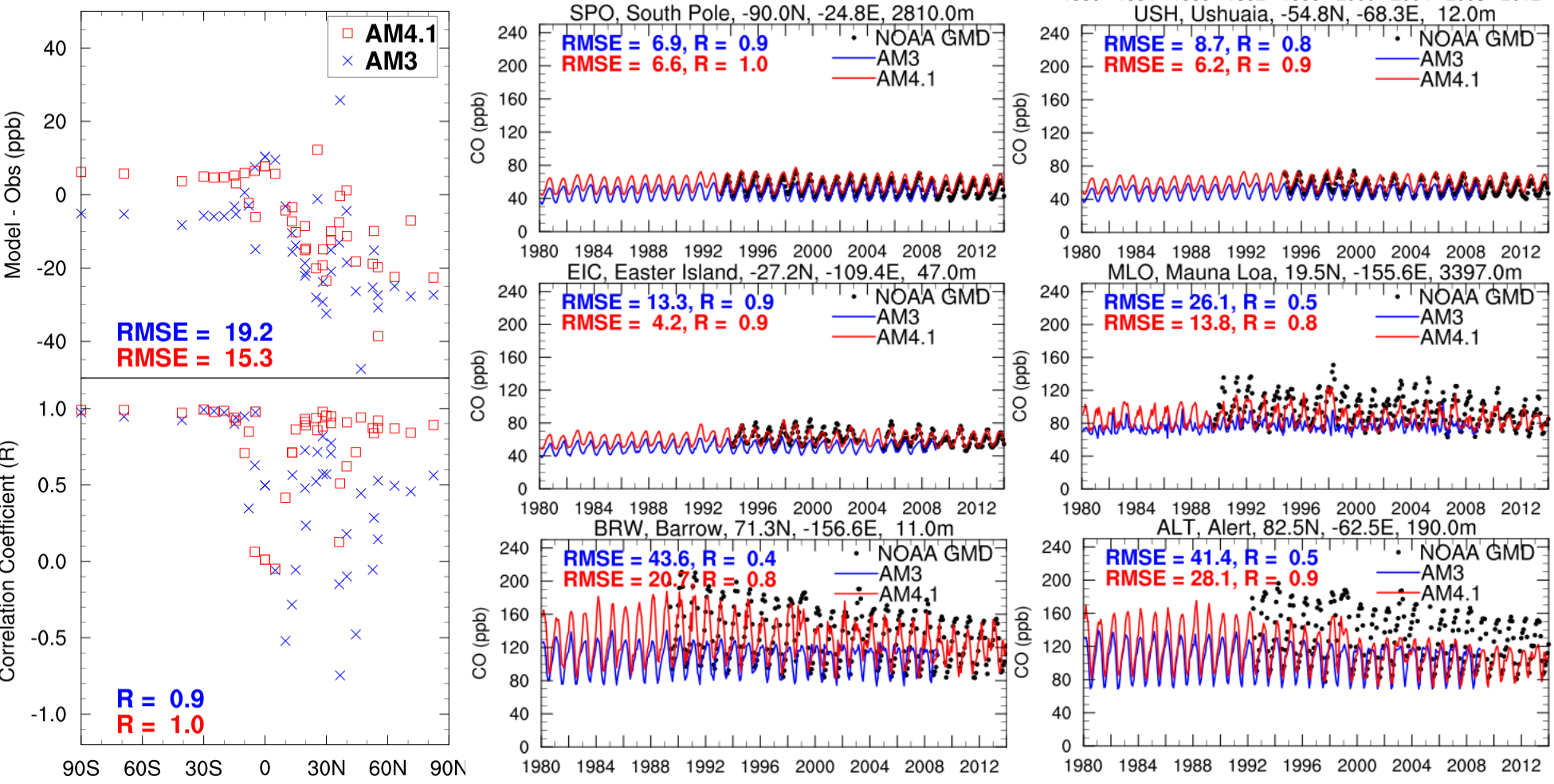


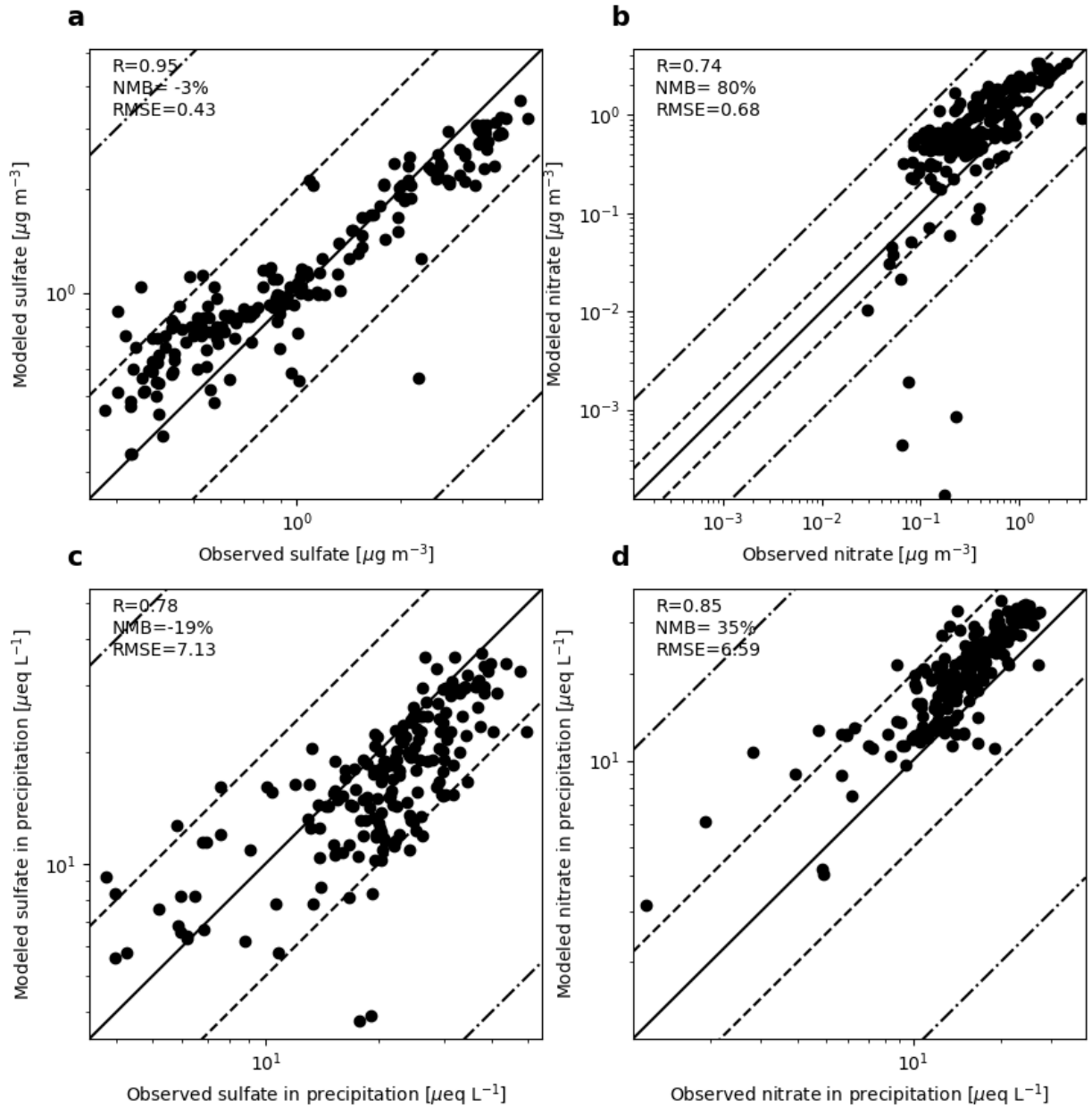
AM3 - MOPITT (JJA)

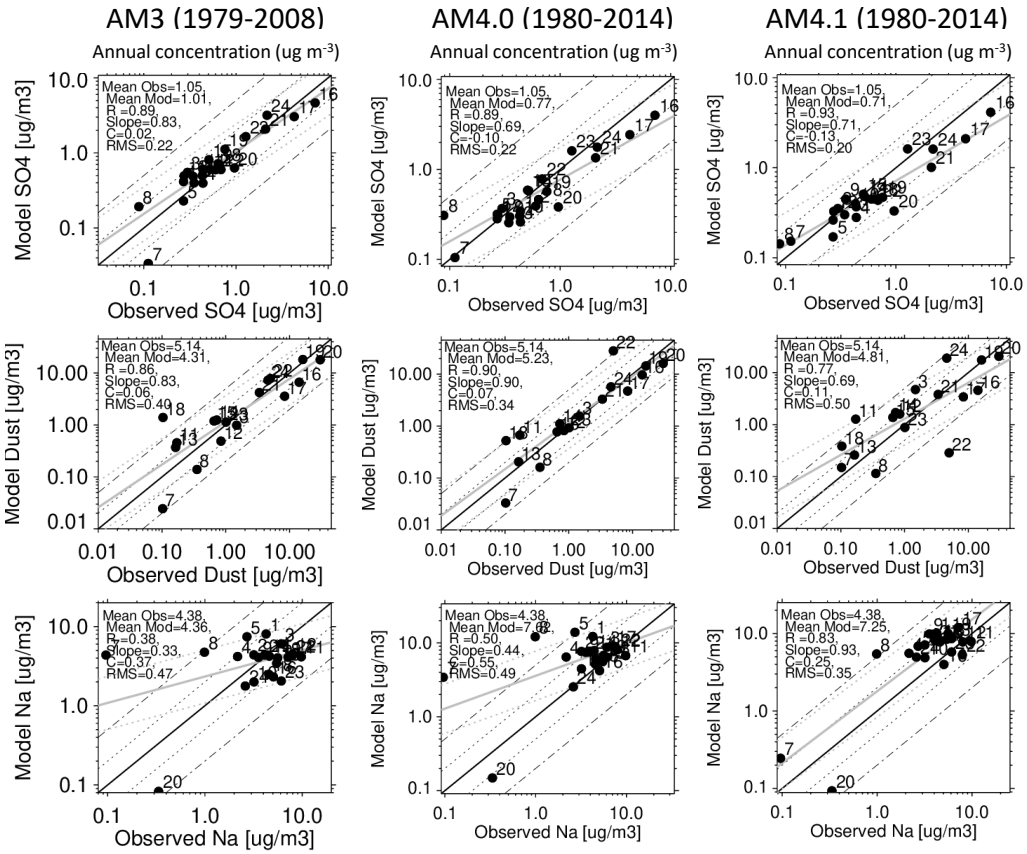


AM4.1 - MOPITT (JJA)

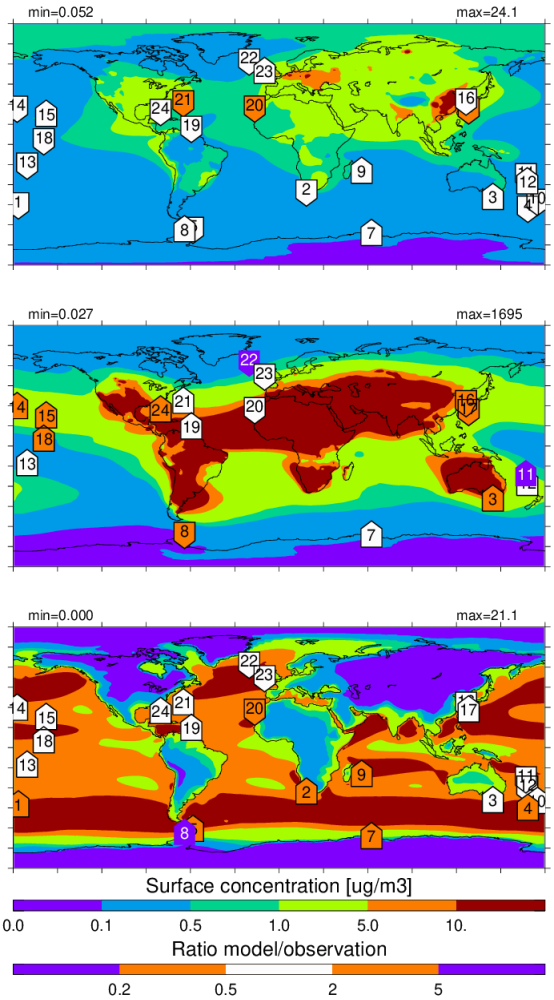


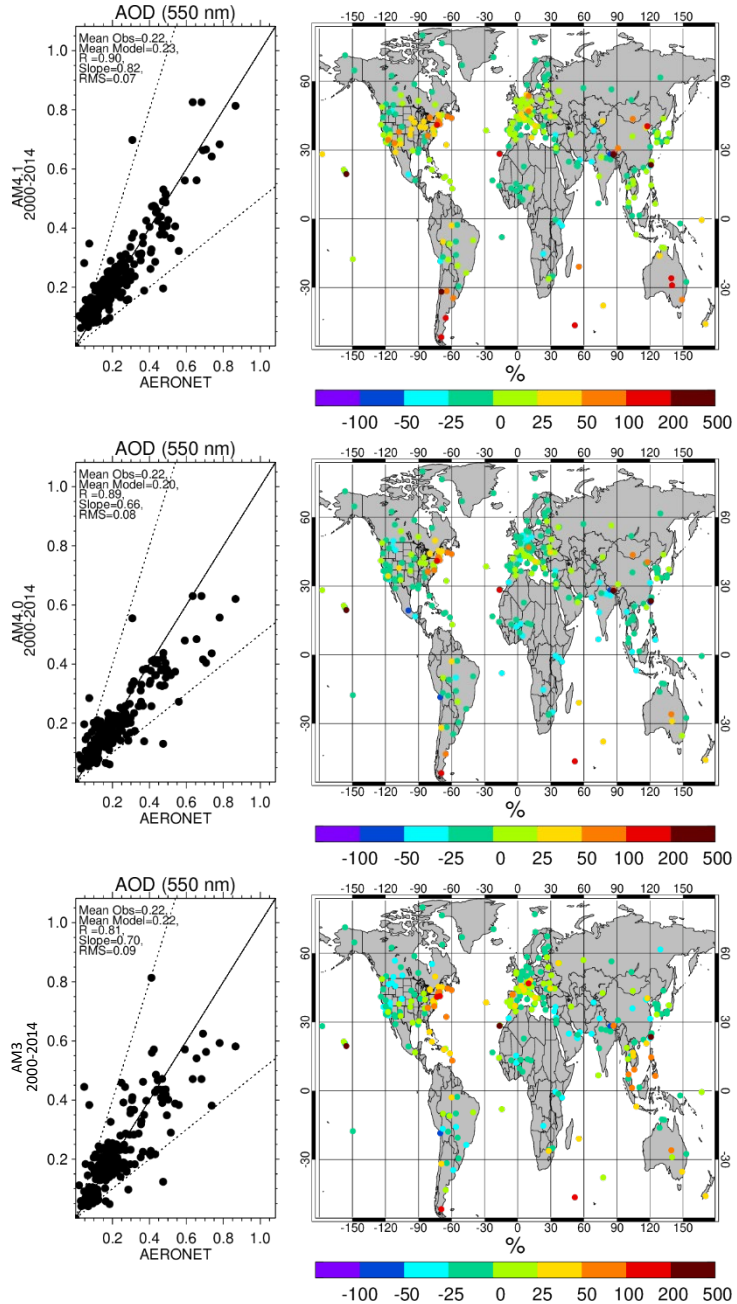

AM4.1
AM3
Model - Obs (ppb)
RMSE = 19.2
RMSE = 15.3
Correlation Coefficient (R)
R = 0.9
R = 1.0
90S
60S
30S
0
30N
60N
90N
SPO, South Pole, -90.0N, -24.8E, 2810.0m
RMSE = 6.9, R = 0.9
RMSE = 6.6, R = 1.0
USH, Ushuaia, -54.8N, -68.3E, 12.0m
RMSE = 8.7, R = 0.8
RMSE = 6.2, R = 0.9
EIC, Easter Island, -27.2N, -109.4E, 47.0m
RMSE = 13.3, R = 0.9
RMSE = 4.2, R = 0.9
MLO, Mauna Loa, 19.5N, -155.6E, 3397.0m
RMSE = 26.1, R = 0.5
RMSE = 13.8, R = 0.8
BRW, Barrow, 71.3N, -156.6E, 11.0m
RMSE = 43.6, R = 0.4
RMSE = 20.7, R = 0.8
ALT, Alert, 82.5N, -62.5E, 190.0m
RMSE = 41.4, R = 0.5
RMSE = 28.1, R = 0.9
NOAA GMD
AM3
AM4.1
CO (ppb)

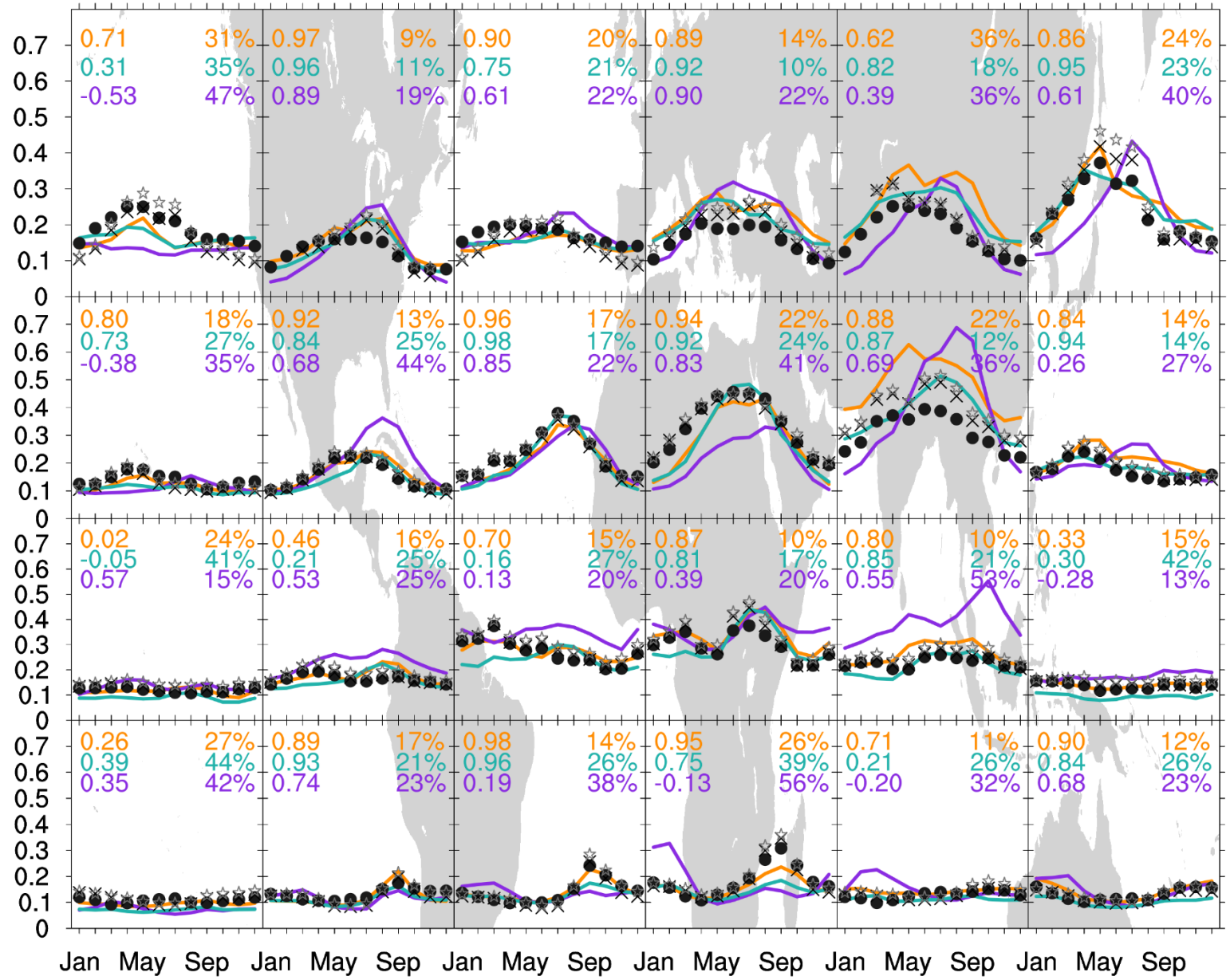

a
R=0.95
NMB= -3%
RMSE=0.43
Modeled sulfate [μg m−3]
Observed sulfate [μg m−3]
b
R=0.74
NMB= 80%
RMSE=0.68
Modeled nitrate [μg m−3]
Observed nitrate [μg m−3]
c
R=0.78
NMB=-19%
RMSE=7.13
Modeled sulfate in precipitation [μeq L−1]
Observed sulfate in precipitation [μeq L−1]
d
R=0.85
NMB= 35%
RMSE=6.59
Modeled nitrate in precipitation [μeq L−1]
Observed nitrate in precipitation [μeq L−1]

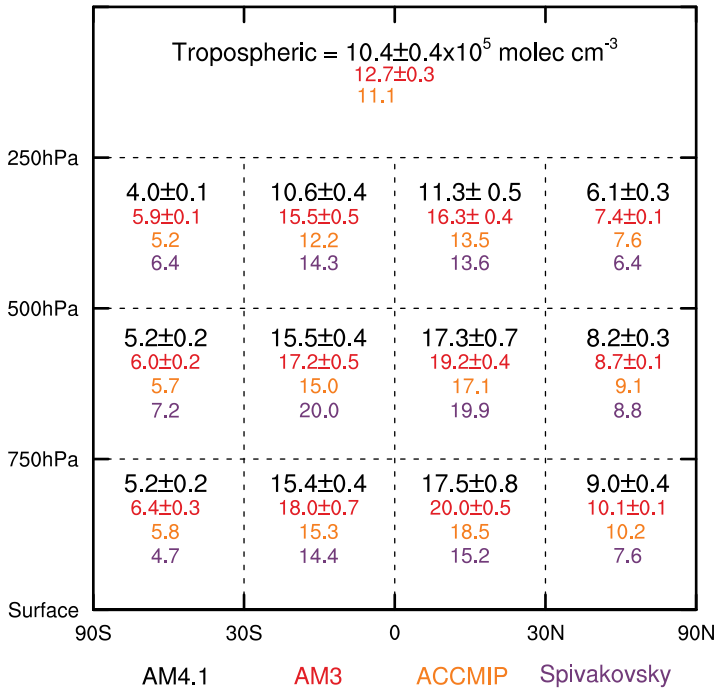


ESM4_amip_D1 1980-2014




AOD (550 nm)
Mean Obs=0.22,
Mean Model=0.23,
R =0.90,
Slope=0.82,
RMS=0.07
AM4.1
2000-2014
AERONET
%
-100 -50 -25 0 25 50 100 200 500
AOD (550 nm)
Mean Obs=0.22,
Mean Model=0.20,
R =0.89,
Slope=0.66,
RMS=0.08
AM4.0
2000-2014
AERONET
%
-100 -50 -25 0 25 50 100 200 500
AOD (550 nm)
Mean Obs=0.22,
Mean Model=0.22,
R =0.81,
Slope=0.70,
RMS=0.09
AM3
2000-2014
AERONET
%
-100 -50 -25 0 25 50 100 200 500


0.71 31%
0.31 35%
-0.53 47%
0.97 9%
0.96 11%
0.89 19%
0.90 20%
0.75 21%
0.61 22%
0.89 14%
0.92 10%
0.90 22%
0.62 36%
0.82 18%
0.39 36%
0.86 24%
0.95 23%
0.61 40%
0.80 18%
0.73 27%
-0.38 35%
0.92 13%
0.84 25%
0.68 44%
0.96 17%
0.98 17%
0.85 22%
0.94 22%
0.92 24%
0.83 41%
0.88 22%
0.87 12%
0.69 36%
0.84 14%
0.94 14%
0.26 27%
0.02 24%
-0.05 41%
0.57 15%
0.46 16%
0.21 25%
0.53 25%
0.70 15%
0.16 27%
0.13 20%
0.87 10%
0.81 17%
0.39 20%
0.80 10%
0.85 21%
0.55 53%
0.33 15%
0.30 42%
-0.28 13%
0.26 27%
0.39 44%
0.35 42%
0.89 17%
0.93 21%
0.74 23%
0.98 14%
0.96 26%
0.19 38%
0.95 26%
0.75 39%
-0.13 56%
0.71 11%
0.21 26%
-0.20 32%
0.90 12%
0.84 26%
0.68 23%
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
Jan May Sep Jan May Sep Jan May Sep Jan May Sep Jan May Sep Jan May Sep


Tropospheric = 10.4±0.4x10^5 molec cm^-3
12.7±0.3
11.1
250hPa
500hPa
750hPa
Surface
4.0±0.1
5.9±0.1
5.2
6.4
10.6±0.4
15.5±0.5
12.2
14.3
11.3± 0.5
16.3± 0.4
13.5
13.6
6.1±0.3
7.4±0.1
7.6
6.4
5.2±0.2
6.0±0.2
5.7
7.2
15.5±0.4
17.2±0.5
15.0
20.0
17.3±0.7
19.2±0.4
17.1
19.9
8.2±0.3
8.7±0.1
9.1
8.8
5.2±0.2
6.4±0.3
5.8
4.7
15.4±0.4
18.0±0.7
15.3
14.4
17.5±0.8
20.0±0.5
18.5
15.2
9.0±0.4
10.1±0.1
10.2
7.6
90S
30S
0
30N
90N
AM4.1
AM3
ACCMIP
Spivakovsky