Supplement of

How well do Earth system models reproduce the observed aerosol response to rapid emission reductions? A COVID-19 case study

Ruth A. R. Digby et al.

Correspondence to: Ruth A. R. Digby (ruth.digby@ec.gc.ca)

Table S1. Mean AOD over reference period (2015-2019) from different satellites, and the spread in these estimates. The relative spread is calculated with respect to the mean of the highest and lowest AOD estimates in each region.

Region	MODIS Aqua	MISR	CALIOP ASN	ACROS-C	absolute spread	relative spread
N. Hemis	0.231	0.190	0.178	0.181	0.053	26%
E. China	0.483	0.331	0.445	0.451	0.152	30%
India	0.402	0.365	0.459	0.477	0.112	27%
Europe	0.157	0.125	0.127	0.124	0.033	23%

S1   Observational Uncertainties and Sampling Differences

This work is based on comparisons between observed and simulated anomalies of AOD. In order to draw meaningful conclusions from these comparisons we need to assess how much of the difference between our datasets comes from differences in sampling, and how much can be attributed to uncertainties or biases in the products themselves. Here we assess the role of sampling differences in explaining the considerable spread between AOD estimates from different satellites, which can differ by 20-30% (Table S1).

Our main analysis used monthly and regional mean data products. For the simulations, these means are spatiotemporally complete; the observed means use retrievals obtained at the satellite's particular overpass time, in clear-sky conditions (for the passive sensors), when the retrieval was successful and not prevented by a myriad of potential limitations such as sun glint or complex terrain. Here we conduct a systematic intercomparison between pairs of observational data products, with each pairing selected to isolate the effect of a particular sampling effect. Discrepancies that cannot be attributed to sampling differences are then used to estimate the degree of uncertainty in our observed AOD responses. Throughout, figures that do not require gridcell-to-gridcell comparisons (spatial-mean timeseries, histograms) are generated at the products' native resolutions, and figures that do require gridcell-to-gridcell comparisons (scatter plots, calculation of correlation between datasets) are generated on fields that have been interpolated to a common $1° \times 1°$ resolution.

S1.1   Temporal Sampling

We first assess the effects of temporal sampling on our comparison between satellite observations, which are sampled only at a particular overpass time, and model simulations, which are integrated over the full diurnal cycle. Figure S1 compares three-hourly, daily, and monthly AOD values for MAM 2004 from a sample CanAM5 simulation. Timeseries plot the evolution of AOD over the three-month period in each of our analysis regions, with blue, orange, and green lines corresponding to the different averaging timescales. A dotted black line indicates the diurnally-integrated, MAM-mean AOD. For all regions except the Northern Hemisphere, blue horizontal lines additionally show the mean of AOD for MAM 2004 sampled only at our satellites' overpass times: dash-dot for the Aqua daytime overpass (MODIS-Aqua and CALIOP daytime product), dotted for the Aqua nighttime overpass (CALIOP nighttime product), and dashed for the Terra daytime overpass (MODIS-Terra, MISR). Local satellite overpass times were estimated from Figure 5 of Levy et al. (2018). These overpass times were converted to UTC and the closest 3-hour timestamp was selected from the model output. The MAM-mean AODs obtained by sampling at each overpass time, and their biases relative to the diurnally-integrated values, are summarized in Table S2.

The MAM-mean AODs of the diurnally-integrated and overpass-sampled products are very similar, especially compared to the spread in AOD from different instruments (Table S1). This fact is unsurprising given that the three-hourly and daily time series demonstrate that the magnitude of the simulated diurnal cycle is negligible compared to the weekly- and monthly-scale variability. Subject to the accuracy of CanAM5's simulated diurnal cycle, then, we do not expect differences between the observed and simulated AOD samples to be caused by differences in the temporal sampling frequency.

Even though the above results indicate that simulated AOD does not vary substantially throughout the day, other factors affecting the retrieved AOD may. These factors include the possibility of a diurnal cycle in cloud cover (which would introduce sampling differences between the products), or in retrieval uncertainty (such as the increased uncertainty in daytime CALIOP retrievals due to the solar background). We assess this possibility by comparing sets of observations taken by instruments with
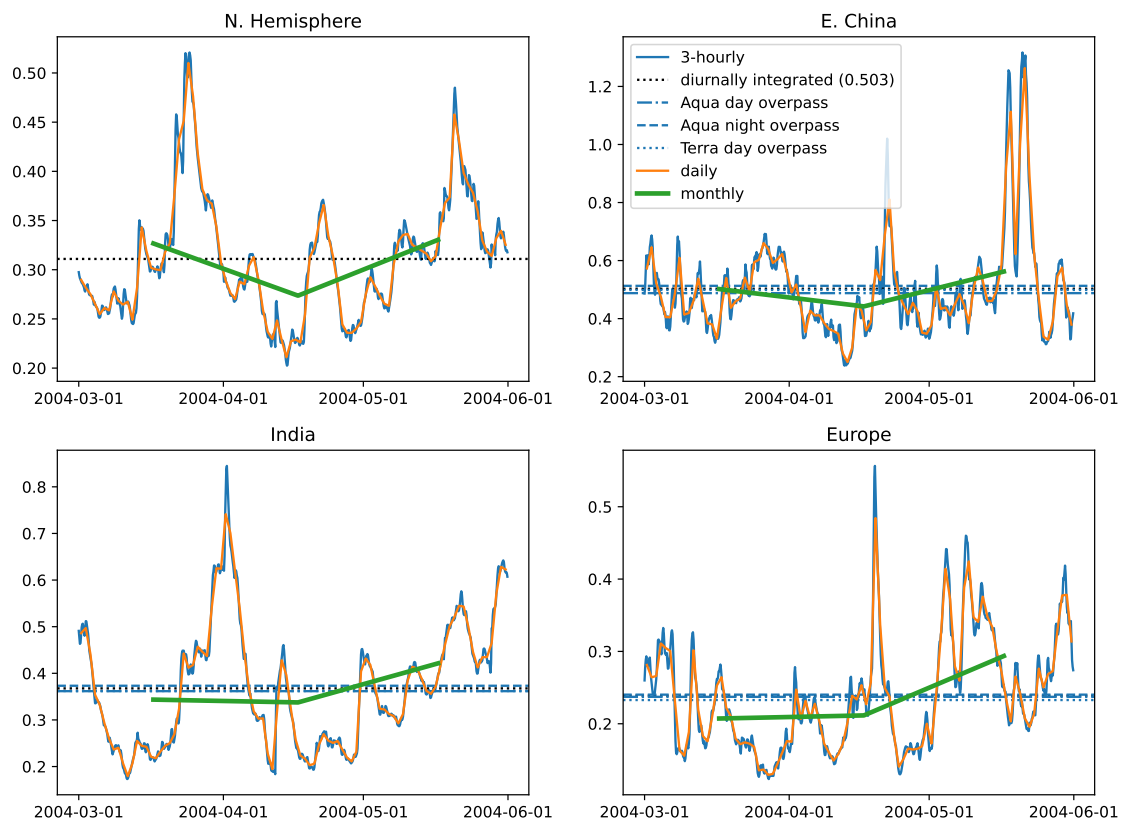
Figure S1. The impacts of temporal averaging and sub-daily sampling demonstrated with CanAM5. Timeseries show 3hr, daily, and monthly AOD (blue, orange, and green lines respectively) through MAM 2004 in our 4 analysis regions. Horizontal lines show the mean AOD when averaged over the full time period from March 2004 to June 2004 (black) and when sampled only at satellite overpass times (blue); these values are tabulated in Table S2.

Table S2. Mean AOD values obtained from 3-hourly CanAM output, when averaged over all of MAM 2004 ("full day") and when sampled only at satellite overpass times. Values in parentheses indicate the percent change in AOD when sampled at a particular overpass time as compared to the diurnally-integrated value. These relative differences are an order of magnitude smaller than the spread bewteen region-mean AOD from different satellites (Table S1).

Region	Full Day	Aqua Day Overpass	Aqua Night Overpass	Terra Day Overpass
N. Hemis	0.311	-	-	-
E. China	0.503	0.488 (-3.0%)	0.513 (+2.0%)	0.499 (-0.8%)
India	0.368	0.362 (-1.6%)	0.373 (+1.3%)	0.362 (-1.6%)
Europe	0.239	0.237 (-0.8%)	0.240 (+0.4%)	0.233 (-2.5%)

the same algorithms but different overpass times: MODIS Aqua and MODIS Terra in Figure S2 and CALIOP AllSkyDay and CALIOP AllSky Night in Figure S3.

MODIS Terra crosses the equator from North to South at 10:30 local time, and MODIS Aqua crosses from South to North at 13:30. For regions near the equator (e.g. India), their overpass times thus differ by roughly three hours; further poleward (e.g. Europe) the difference is larger. There is a known offset of ∼0.02 (∼13%) between the global, monthly-mean AOD values retrieved by the two MODIS instruments, with Aqua in better agreement with AERONET measurements and Terra biased higher (Levy et al., 2018; Schutgens et al., 2020). A detailed assessment of the Level 2 daily aerosol products, incorporating ground-based AERONET observations and the Goddard Earth Observing System (GEOS-5) Earth system model, has demonstrated that this offset is not due to diurnal cycles in either aerosols or clouds, but likely to a calibration or algorithmic bias (Levy et al., 2018).

The Levy et al. assessment investigated global AOD biases. It is possible that diurnal cycles could lead to significant biases on a regional scale. Overall, however, Figure S2 suggests that the Aqua and Terra products are in good agreement and the effect of differing overpass times on MAM-mean AOD is small.

Figure S3 undertakes a similar comparison between CALIPSO's all-sky day and night products. Because these products are produced by a single instrument, calibration drift is not a concern. Using the all-sky product eliminates the sampling effects that could be introduced by diurnal cycles in cloud cover. It is worth noting, however, that CALIPSO's daytime products have higher uncertainty due to the presence of a solar background which reduces the signal-to-noise ratio (Young et al., 2018). There are appreciable differences between the distributions of daytime and nighttime AOD retrievals over East China and Europe, with the nighttime product being skewed to larger values than the daytime one. Reassuringly, however, the distributions of anomalies are in good agreement. The low correlations in the scatter plots of these anomalies are not concerning, since aerosols may have advected over the 12 hours between overpasses, so we would not necessarily expect gridcell-by-gridcell agreement between the two products. The important metric in this comparison is that the aggregate distribution of anomalies is similar.

In principle, instrument-to-instrument spread in the amount of dependence on overpass time could indicate the presence of diurnally-varying biases in the various products. However, because there is different spacing between the overpass times of the two instrument sets considered here, the existence of these biases cannot be disentangled from a true diurnal cycle in AOD. For the purposes of this analysis, however, the impact of temporal sampling on spatiotemporally-aggregated AOD anomalies is found to be small and is unlikely to be a dominant contributor to the spread between datasets seen in the main analysis.

S1.2 Cloud Masking

We next compare the all-sky and cloud-free CALIPSO products to isolate the impacts of cloud masking, another difference between the sampling of the models and the passive satellites. We use the daytime CALIPSO product as it is expected to be most similar to the passive sensors, which can only observe in cloud-free conditions.

As Figure S4 demonstrates, the differences between all-sky and cloud-free AOD retrievals measured simultaneously by the same instrument are small when spatially aggregated and averaged over a three-month time period. Cloud-masking-induced spatiotemporal sampling differences between the observed and simulated datasets are thus unlikely to contribute substantially to the differences between datasets in our analysis.

S1.3 Instrument-to-Instrument Differences

Having established that overpass time and cloud masking have small effects on the 3-month spatiotemporal means considered in this work, particularly when considering anomalies, we next compare the AOD retrieved by two different instruments that are orbiting on the same satellite. Comparing MODIS-Terra and MISR, both onboard the Terra satellite, is the closest we can come to isolating the effect of retrieval algorithm on AOD estimates. We use "algorithm" here in the broadest sense, to include not only differences in the AOD retrieval process but also differences in cloud masking, differing susceptibility to effects like sun glint, and the impacts of product resolution. Note that despite being mounted on the same satellite, the two instruments are not viewing the same scene: MISR's field of view is much smaller than, and embedded within, MODIS's.

Numerous comparisons between MODIS Terra and MISR exist in the literature. These include both retrieval-level assessments and comparison of the resulting spatiotemporally aggregated products. When MODIS Terra and MISR conduct si-
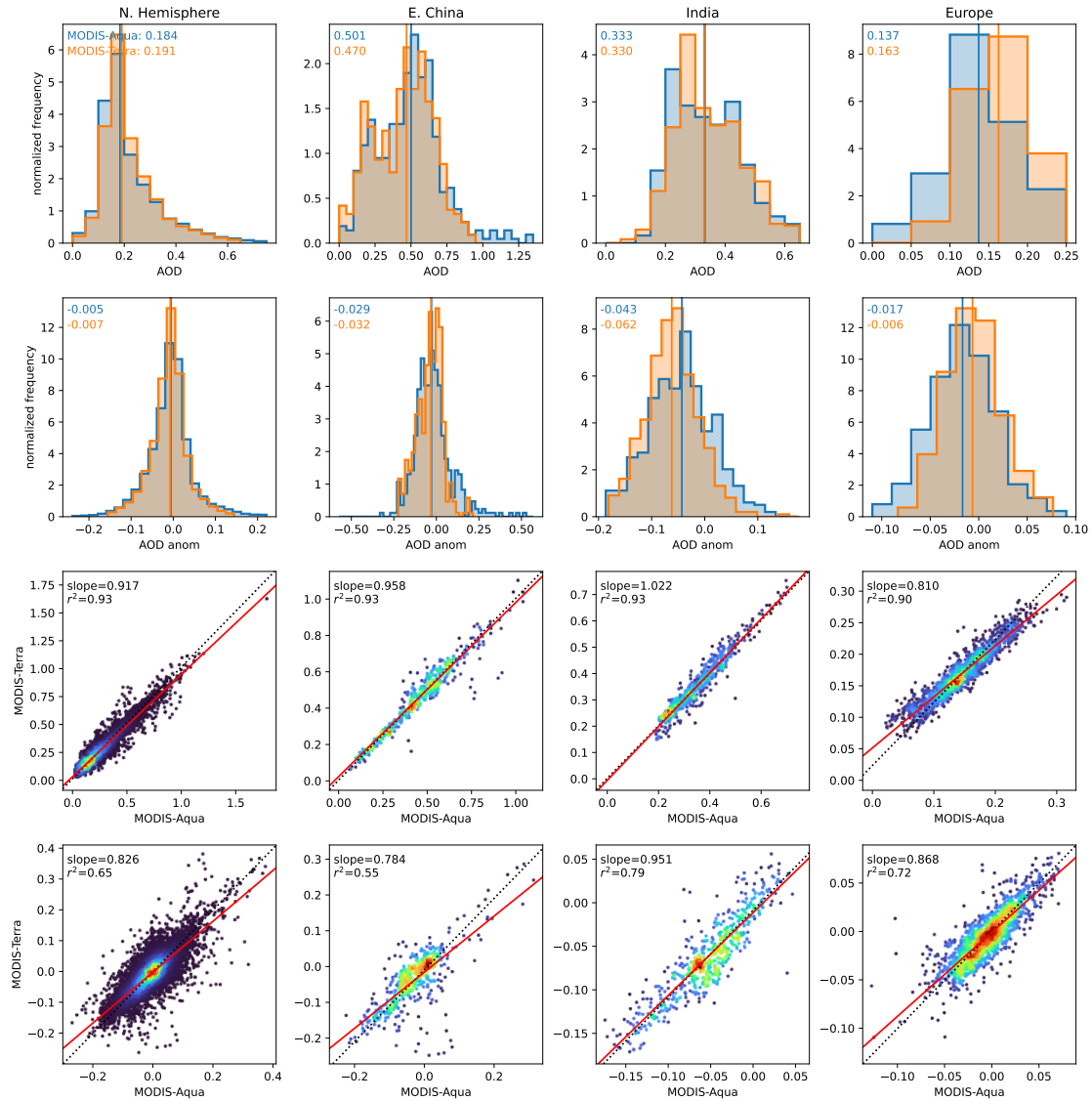
Figure S2. Investigation of the effects of overpass time on observed AOD, by comparing retrievals from MODIS-Terra (equatorial overpass 10:30 local time descending) and MODIS-Aqua (equatorial overpass 13:30 local time ascending). Histograms show the distributions of MAM 2020 AOD values (top row) and anomalies (bottom row) at the grid cell level in our 4 analysis regions, with blue and orange corresponding to MODIS- Aqua and MODIS-Terra respectively. Note that the bin width in each panel is the same: 0.05 for raw AOD values, and 0.02 for AOD anomalies. Text in the upper left corner of each panel lists the median value in each distribution (blue and orange for Terra and Aqua respectively). Scatter plots show these same distributions plotted against each other, again with AOD values in the top row and anomalies in the bottom. Red lines show a linear fit to the scatter, dotted black lines show 1:1 to guide the eye, and colour indicates normalized point density. Because these are both MODIS instruments they share the same algorithm, cloud mask, etc. and only differ in overpass time.
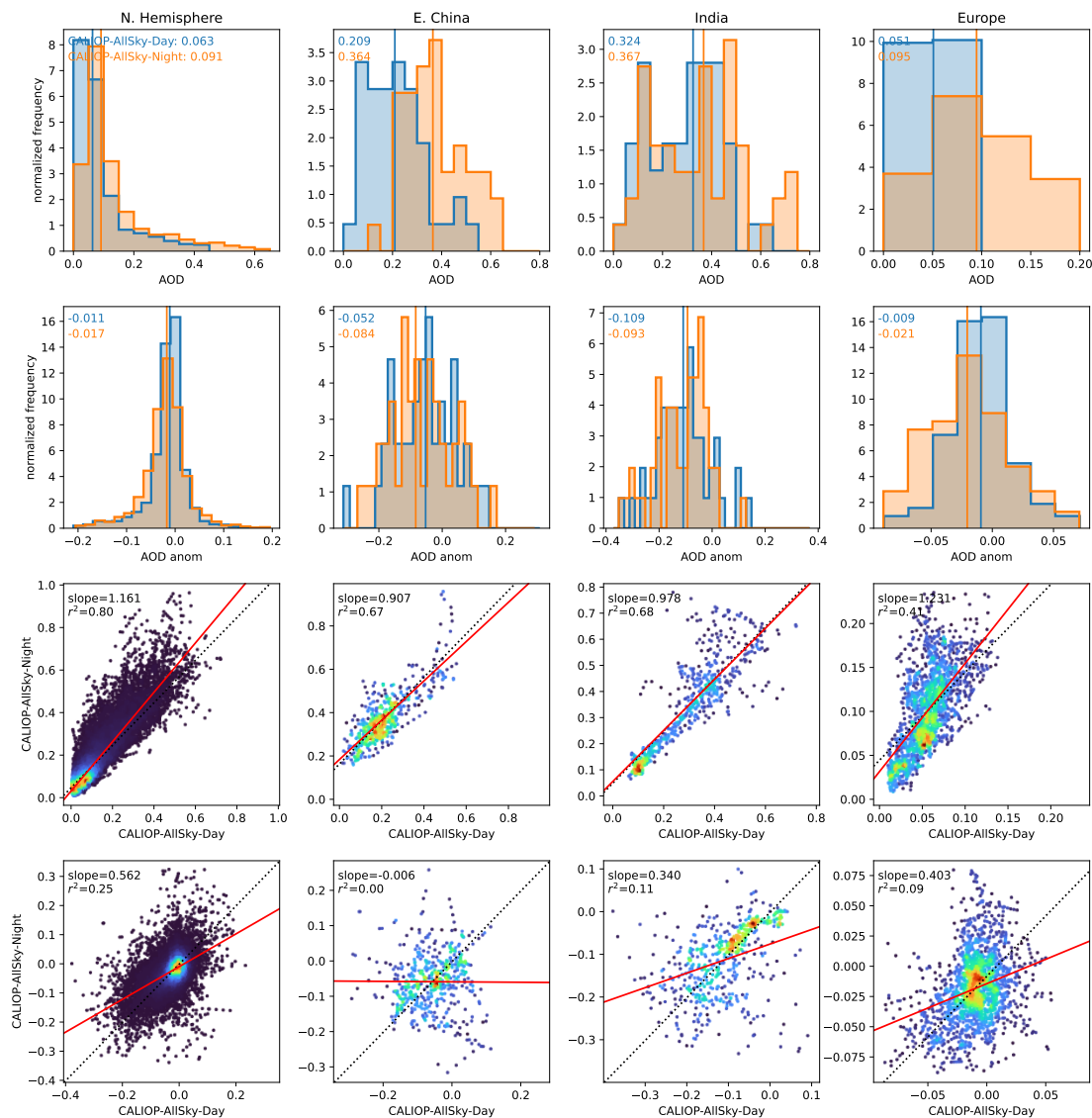
Figure S3. As Figure S2 but comparing CALIPSO all-sky measurements made during the day and night overpasses (blue and orange histograms; 13:30 and 01:30 equatorial crossings respectively). Because CALIPSO is a single satellite it is not subject to calibration drift the way that MODIS-Aqua and -Terra are. Further, using the all-sky product eliminates the potential sampling differences that could be induced by diurnal cycles in cloud cover. Although differences are seen in the distributions of daytime and nighttime AODs, particularly over East China and Europe, the distributions of anomalies are in good agreement. The low correlations may be the result of advection, since aerosols may have moved about the region during the 12 hours between overpasses. Of greater importance in this comparison is the overall distribution shapes, which we find to be similar.
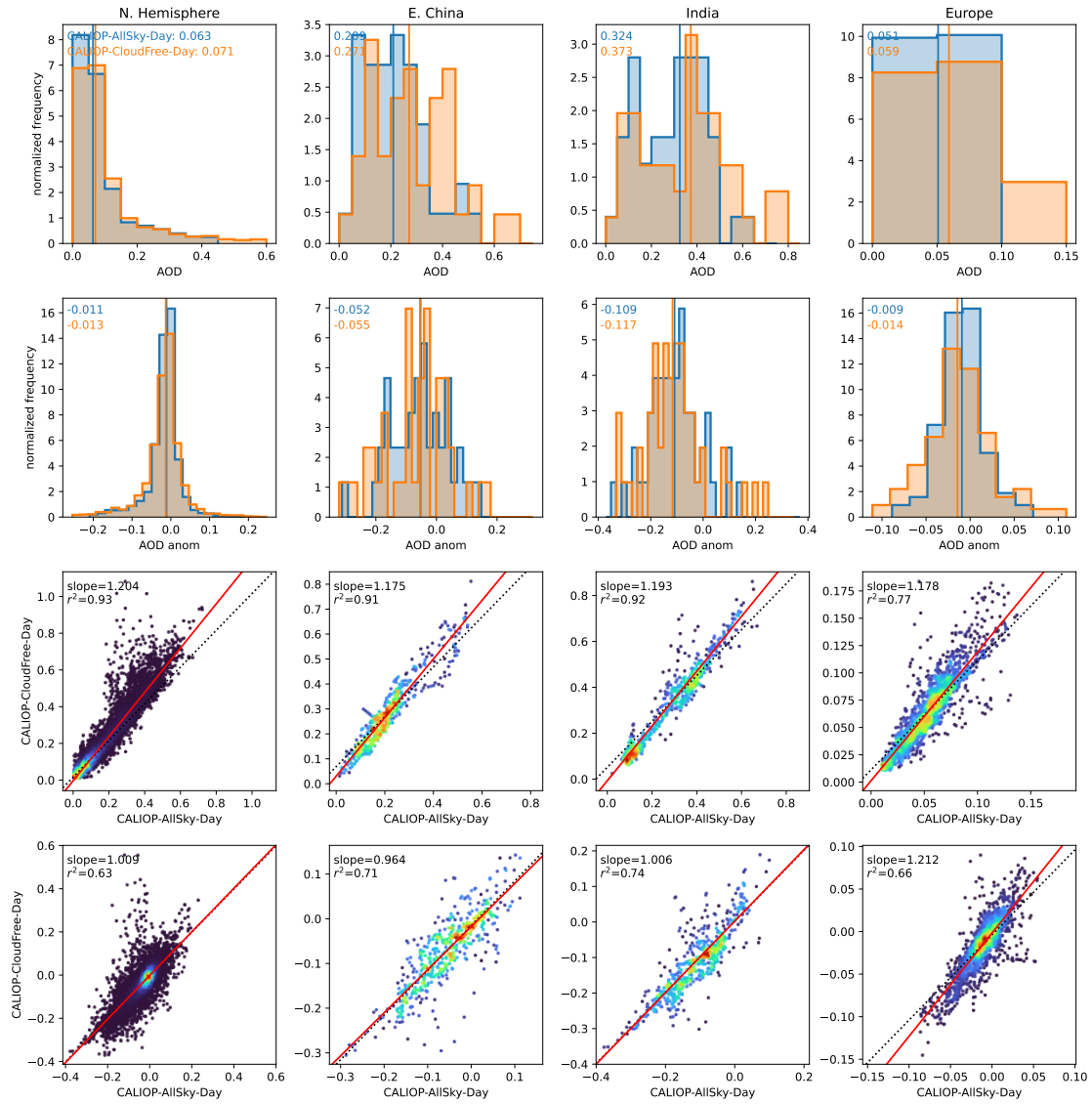
Figure S4. As Figure S2 but comparing CALIPSO all-sky (blue) and cloud-free (orange) data products to isolate the sampling effects of cloud masking.

multaneous retrievals of the same scene, the resulting optical depth measurements are in reasonably good agreement. In an assessment of previous versions of the algorithms (MODIS C5 and MISR v22; our analysis uses MODIS C6.1 and MISR v23), Kahn et al. (2009) report correlations of ~ 0.9 over ocean and ~ 0.7 over land. However, these coincident retrievals are comparatively rare. Factors limiting the attainment of coincident retrievals include differences in the instruments' cloud masking algorithms, and the susceptibility of MODIS to sun glint which MISR can avoid using its off-axis cameras. When AOD retrievals are aggregated into gridded monthly products, the agreement is noticeably worse, with MODIS Terra biased high and MISR biased low relative to ground-based AERONET measurements (Wei et al., 2019; Sogacheva et al., 2020; Vogel et al., 2022; Mangla et al., 2020). These biases vary regionally and seasonally, with MISR biased especially low over East Asia and MODIS Terra biased especially high over North America and the Middle East (Wei et al., 2019).

For our analysis regions, the MAM 2020 AOD of these two products differ far more than any of the preceding comparisons between products with shared algorithms (Figure S3 - Figure S5). Notably, the differences are characterized not only by systematic offsets (e.g. with MODIS consistently reading higher than MISR) but by different distribution shapes. Furthermore, the characteristics of these discrepancies vary from region to region, supporting previous reports of spatially-dependent systematic differences between the datasets.

Comparing Figure S5 with the preceding figures suggests that the bulk of the difference between different data products' raw AOD values results from systematic biases rather than to differences in sampling. Reassuringly, however, the anomalies in Figure S5 are substantially more robust than the raw values, and exhibit agreement comparable to any of the pairs of datasets with shared algorithms. Deseasonalizing effectively removes a majority of the systematic difference between datasets. This observation largely motivates our use of deseasonalized products in the main analysis.

S2 Extended Figures: Dust-Subtracted Aerosol Optical Depth

This section includes supplementary figures to support the discussion in Sections 5.2 and 6. These figures include a timeseries showing the magnitude of dust-subtracted AOD in the CovidMIP models over 2015-2020 (Figure S6), and maps of the observed (Figure S7) and simulated (Figure S8) DSAOD anomalies in MAM 2020.

S3 Sensitivity Tests: Technical Details and Further Analysis

S3.1 Construction of Aerosol Emission Inventories

We construct updated aerosol emission inventories for both anthropogenic and biomass-burning emissions of black carbon (BC), organic carbon (OC), and sulfur dioxide (SO$_2$), for the period 2013-2020. Gridded monthly anthropogenic emissions for 2013-19 were taken from the 2021-04-21 release of the Community Emissions Data System (CEDS) emission inventory (Smith et al., 2021); the 2016-07-16 release, described in Hoesly et al. (2018), was used for the historical period in CMIP6. For 2020, we set control emissions to repeat the 2019 emissions, and create a COVID-19-perturbed dataset by applying the same percentage reduction to these control emissions as was applied to the SSP2-4.5 emissions in the original CovidMIP experiment. For simplicity we do not perturb the aviation emissions. Although aviation was dramatically reduced during COVID-19, the aviation emissions simulated in CanAM5 contain minimal aerosol content. We also do not account for the reduction in sulfur oxide emissions from shipping which followed regulations introduced by the International Maritime Organization (IMO) in 2020 (IMO). Although the signal of this reduction has been identified in ship tracks (Watson-Parris et al., 2022), a corresponding AOD signal is not evident in our data.

Biomass burning emissions for 2013-2020 were taken from the Global Fire Emissions Database (GFED). GFED observations were used in CMIP6 from 1998 to 2014 (van Marle et al., 2017), and extending the dataset to 2020 provides a more realistic estimate than the linear projection used in SSP2-4.5.

The change in accumulated aerosol emissions over the 2015-2019 reference period between the original and updated inventories is shown in Figure S9. Reductions in anthropogenic emissions are dominated by changes prior to 2015 [1]. Following

---

[1] https://github.com/JGCRI/CEDS/blob/master/documentation/Version_comparison_figures_v_2021_02_05_vs_v_2016_07_16(CMIP6).pdf
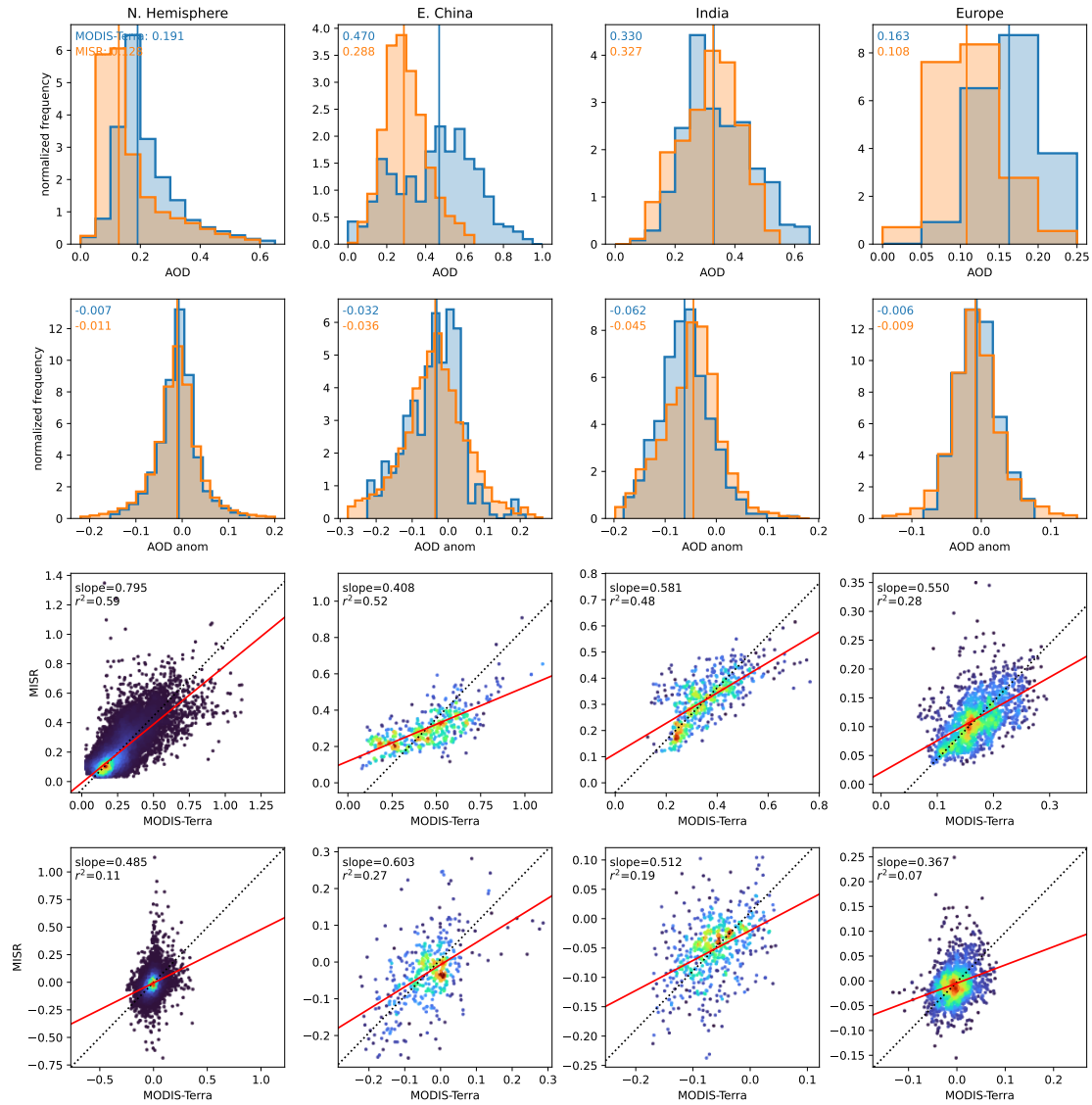
Figure S5. As Figure S2 but comparing MODIS-Terra (blue) and MISR (orange). MODIS and MISR are both on board the Terra satellite. Differences between the products therefore result from differences in detection and retrieval algorithms, as well as in the field of view.
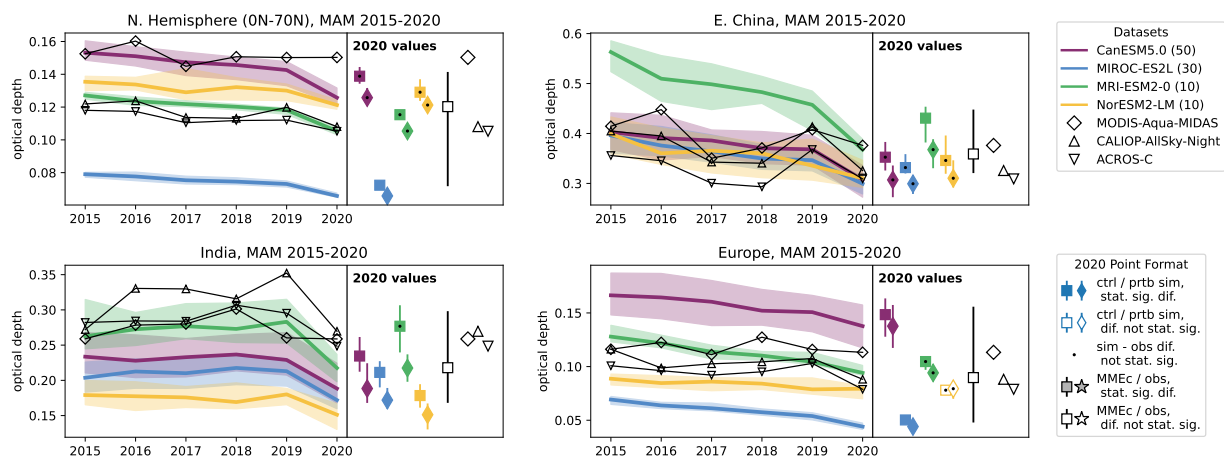
Figure S6. As Figure 2 but showing the magnitudes, not anomalies, of the dust-subtracted aerosol optical depth.

2015, emissions of anthropogenic BC and OC follow similar or somewhat less negative trends to those from SSP2-4.5, albeit at a lower magnitude. In East China, SO2 emissions decrease more rapidly over 2015-2017 in the updated inventory than in SSP2-4.5, but then flatten out over the remainder of the reference period. Because all three species' emission changes are dominated by reductions in China, this same behaviour is seen in Northern Hemispheric and global mean SO2 emissions. The impacts of these reduced emissions on total AOD and aerosol radiative forcing over 1990-2019 have been recently investigated by Lund et al. (2023), who identify decreased AOD and improved agreement with observations over this time period when using the more recent inventories.

S3.2 Model Description: CanAM-new-emis and CanAM-new-emis-ndgd

CanAM-new-emis and CanAM-new-emis-ndgd were set up identically except that CanAM-new-emis-ndgd was nudged to ERA5 reanalysis (details below). The two sets of simulations are in most ways identical to the original CanESM5 experiment, except for ensemble size (10 realizations instead of 50), the updated emissions inventories, a tracer-transport tuning correction which prevents the formation of spurious dust storms (**?**), and the fact that CanAM5 is not a coupled model. Instead, sea surface temperatures and sea ice concentrations are specified using the NOAA OI SST V2 High Resolution Dataset (Reynolds et al., 2002), and sea ice masses for the Northern and Southern Hemispheres are taken from PIOMAS (Zhang and Rothrock, 2003) and ORAP5 (Zuo et al., 2017) respectively.

The control ensembles were initialized in 2013 and run until 2020, with the first two years discarded as spin-up, and the COVID-19-perturbed simulations were branched in January 2020 as was done in the original CovidMIP experiment.

In CanAM-new-emis-ndgd, the simulated temperature, wind, and humidity fields were nudged toward the corresponding reanalysis fields using first order relaxation with a nudging timescale of 24 hours. As a result, the simulation closely reproduces the observed variation in these variables (although other parameters continue to evolve freely) and meteorologically-driven variability between the ensemble members is substantially reduced.

We additionally consider a 3-member ensemble of atmosphere-only, free-running CanAM simulations with the original SSP2-4.5 emissions ("CanAM-old-emis") in order to disentangle the effects of model configuration and emission inventory when comparing CanESM5 and CanAM-new-emis (Figure S10).

The effects of model configuration, emissions inventory, and nudging to ERA5 reanalysis on DSAOD magnitudes are illustrated in Figure S11 (MAM means) and Figure S12 (annual means). We do not show the change in total AOD due to the
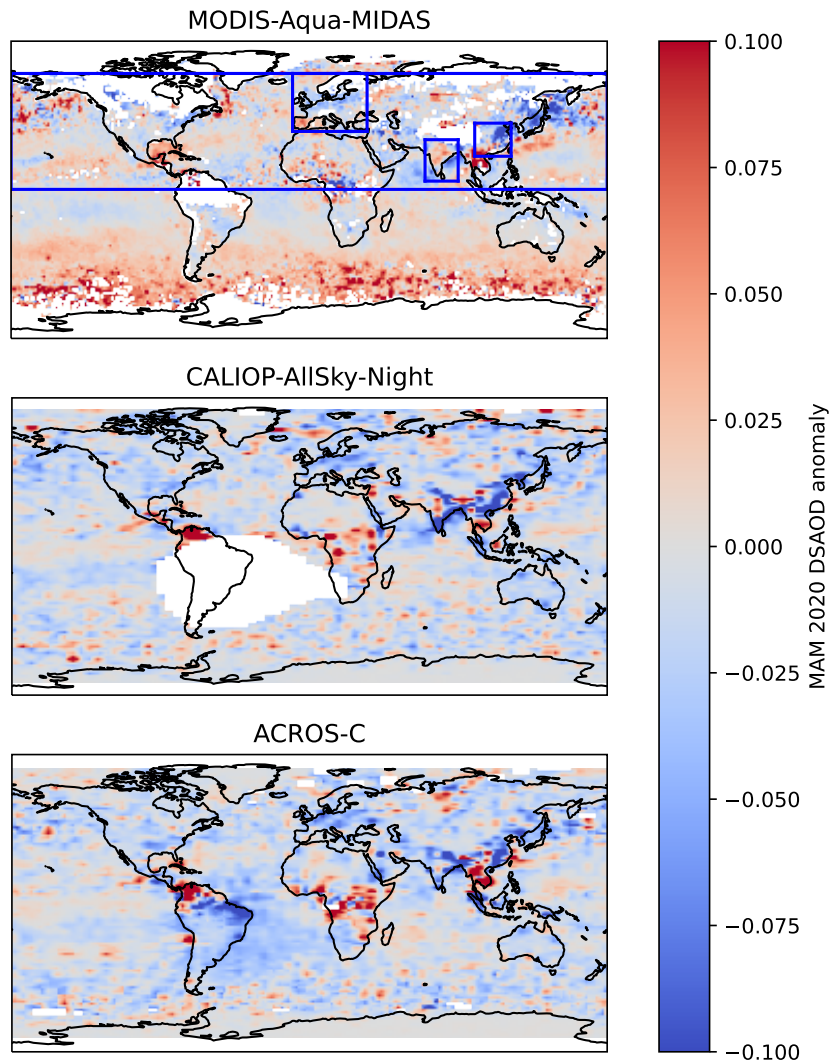
Figure S7. Observed MAM 2020 dust-subtracted AOD anomalies, relative to the 2015-2019 mean. *Top:* Total AOD from MODIS Aqua, dust from MIDAS. *Middle:* Total and dust optical depths from CALIOP AllSky Night. *Bottom:* Total and dust optical depths from ACROS-CALIOP. All maps have been interpolated to 1x1° resolution. Blue boxes in the first panel indicate the boundaries of our four analysis regions. The positive AOD anomalies over the Southern Hemisphere which are visible in the MODIS observations are due, at least in part, to the 2019/2020 Australian bushfires. CALIOP is less sensitive than MODIS to elevated smoke.
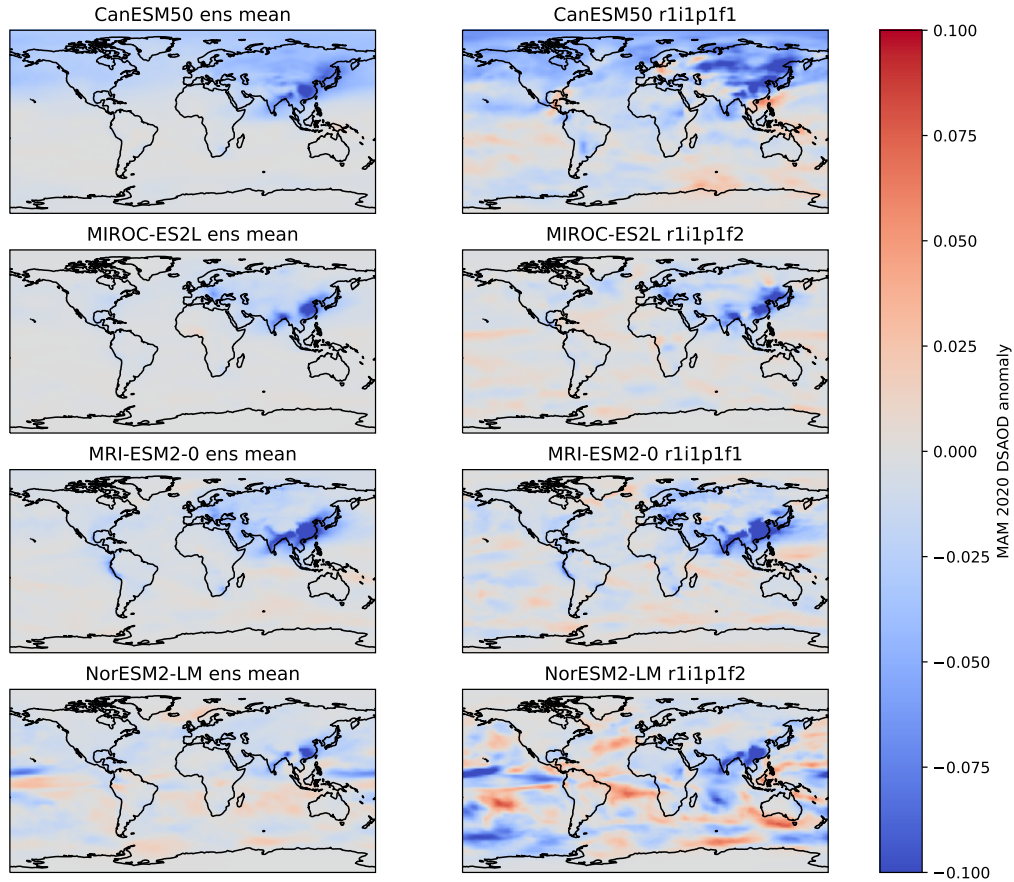
Figure S8. Simulated MAM 2020 dust-subtracted AOD anomalies, relative to 2015-2019 mean. Left-hand column shows ensemble means and right- hand column shows the first realization of each ensemble. Rows correspond to the four CovidMIP models from our main analysis. All maps have been interpolated to $1 \times 1°$ resolution. Note that even single realizations are smoother than the observed optical depth fields (Figure S7), as discussed in Section 6.
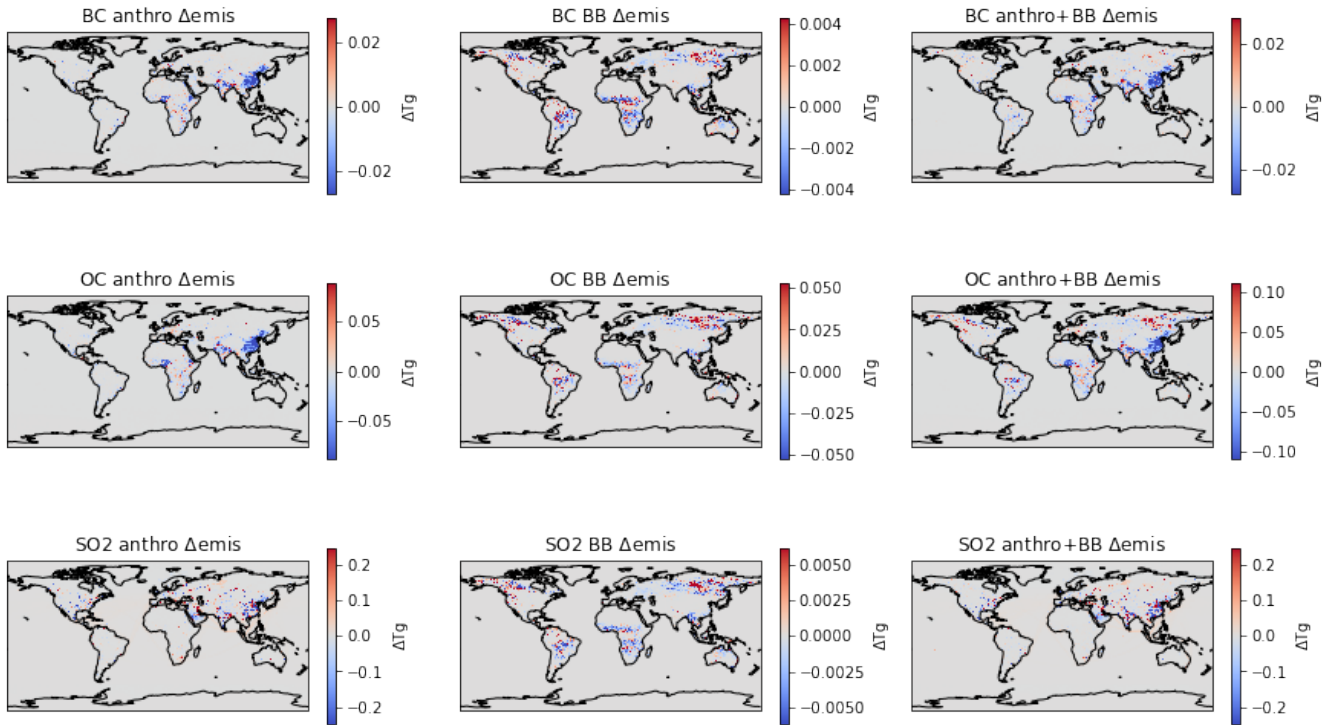
Figure S9. Changes in the accumulated emissions of BC, OC, and SO2 over 2015-2019 when updating from SSP2-4.5 emissions to the latest CEDS historical emissions (v2021-04-21) and GFED biomass burning emissions, decomposed into the changes in anthropogenic, biomass burning, and total emissions. Of our four analysis regions, only in the Northern Hemisphere and only for OC are biomass burning emissions comparable in magnitude to those from anthropogenic sources.

aforementioned issue with mineral dust in CanESM5. Because this issue was corrected in the CanAM5 runs shown here, there
150 is a substantial difference between the total AOD simulated by CanESM5 and CanAM-new-emis which is unrelated to the updated emissions inventories.

In all regions, both springtime and annual-mean DSAOD are lower in CanAM-new-emis than in the original CanESM5 CovidMIP ensemble. In China the emission reductions dominate the difference, whereas model configuration has a larger effect in Europe. The changes are overall small in India in the spring, resulting in the similarity between 2020 anomalies simulated
155 by CanESM5 and CanAM-new-emis that was seen in Figure 3, but emission reductions are more important when considering annual means. When the Northern Hemisphere is considered as a whole, emission reductions and model configuration have similar impacts on DSAOD magnitude. Analysis of both these simulations and a larger ensemble of historical simulations (averaged over 2000-2014, not shown) indicates that the reduction caused by model configuration is systematic, and so AOD anomalies can be directly compared between model configurations. The results of nudging to ERA5 reanalyses are mixed, but
160 generally the magnitude of the simulated AOD is either unaffected or somewhat increased relative to CanAM-new-emis. The interannual variability tends to increase as well, in some cases beyond the ensemble spread simulated by CanAM-new-emis.

Whether changes to the inventory and/or simulated meteorological conditions improve agreement between the magnitude or variability of the observed and simulated optical depth fields depends on the region in question and, for the Northern Hemi-
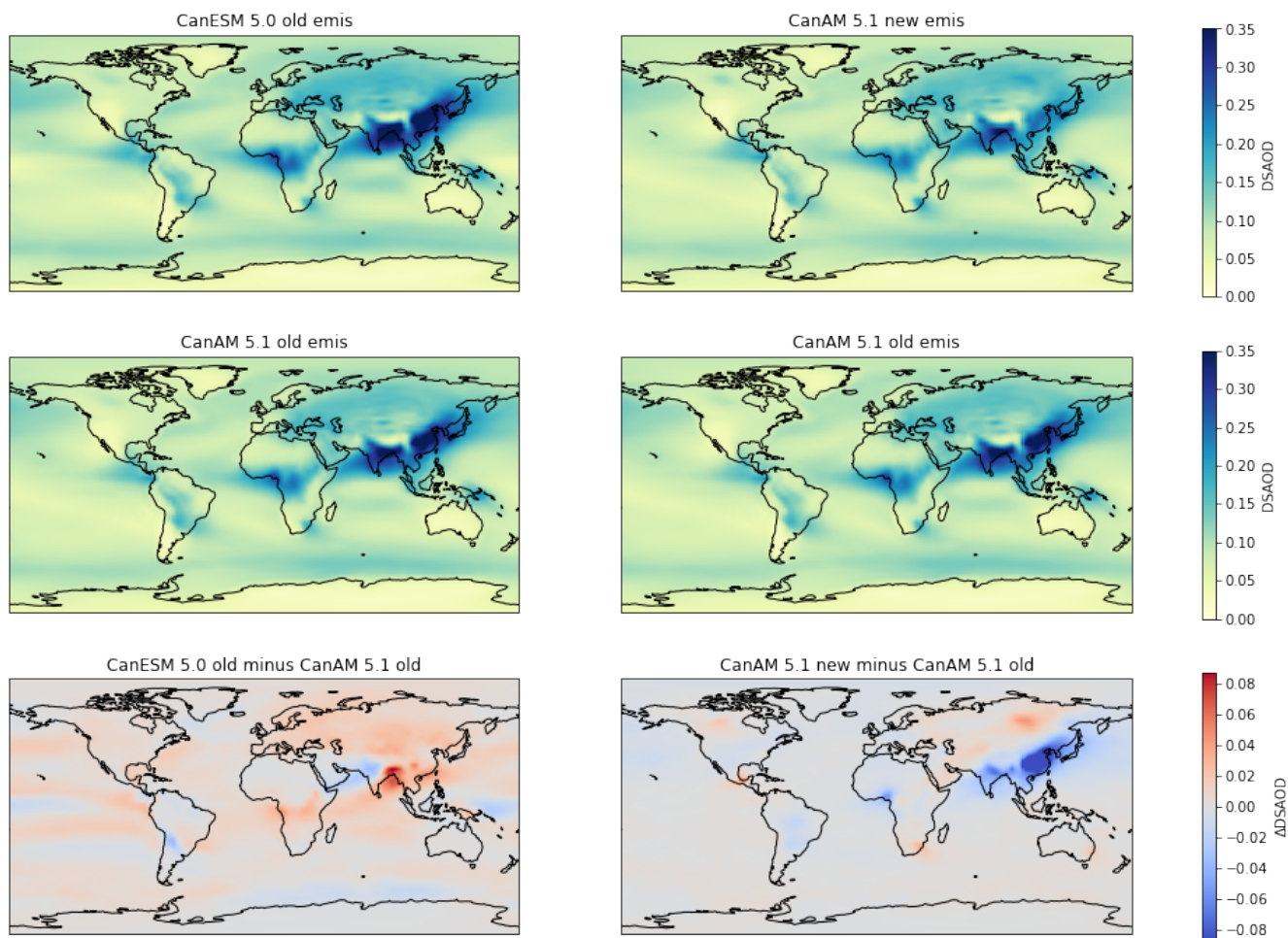
Figure S10. Comparison between ensemble-mean DSAOD, averaged over the 2015-2019 reference period, in three sets of simulations: CanESM5 (upper left), CanAM-new-emis (upper right), and CanAM-old-emis (centre row, both columns). The bottom row shows the DSAOD differences between pairs of ensembles, demonstrating the effects of model configuration (coupled vs atmosphere-only, left column) and emission inventory (SSP2-4.5 vs CEDS v2021-04-21 and GFED biomass burning, right column) on simulated DSAOD.

sphere, on which satellite is considered. The existing discrepancies likely stem from in both the observed and simulated fields, and the relative magnitudes of these biases will change from region to region. It is also important to note that CanAM5 does not have a separate set of parameter tunings for use with nudged simulations. Model parameters are tuned to produce realistic climatological aerosol properties in the free-running simulation. If there are typically biases in the free-running temperature, wind, or humidity fields, the tuning will compensate for these. As a result the model may simulate unrealistic aerosol properties when presented with a more realistic meteorological field.
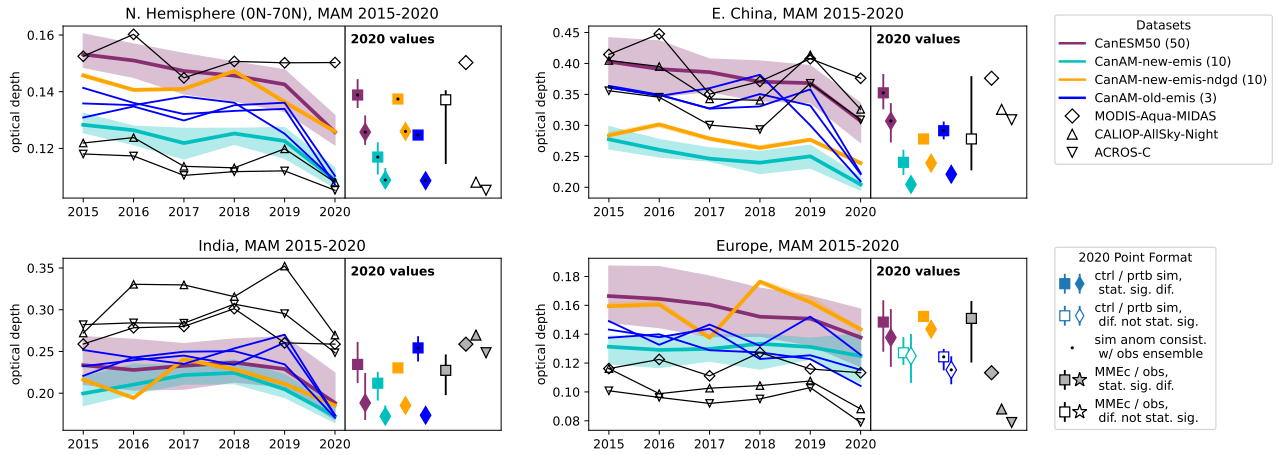
13

Figure S11. Comparison between the dust-subtracted AOD values (not anomalies) over MAM 2015-2020 as simulated by CanESM5, CanAM-old-emis (atmosphere-only, original emissions inventories), CanAM-new-emis (atmosphere-only, updated emissions inventories), and CanAM-new-emis-ndgd (atmosphere-only, updated inventories, and nudged to ERA5 temperature, wind, and humidity fields). Because there are only 3 realizations of CanAM-old-emis, they are shown as individual lines instead of an ensemble median and 5-95% spread as for the other simulations.
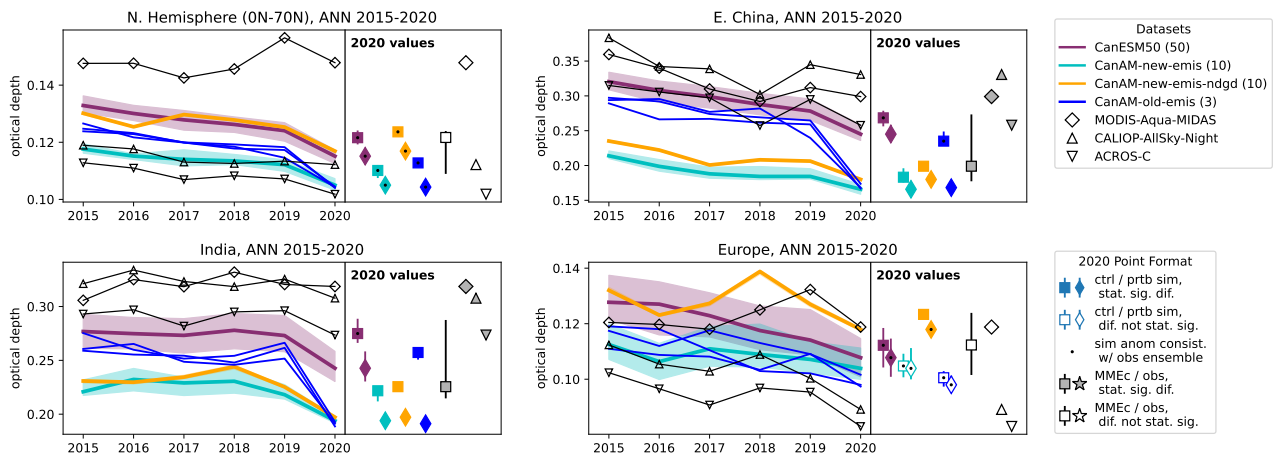

Figure S12. As Figure S11 but for annual means.

References

IMO 2020 – cutting sulphur oxide emissions, https://www.imo.org/en/MediaCentre/HotTopics/Pages/Sulphur-2020.aspx.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-i., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) Anthropogenic Emissions of Reactive Gases and Aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V., Paradise, S. R., Hansen, E. G., and Remer, L. A.: MISR Aerosol Product Attributes and Statistical Comparisons With MODIS, IEEE Trans. Geosci. Remote Sens., 47, https://doi.org/10.1109/TGRS.2009.2023115, 2009.

Levy, R. C., Mattoo, S., Sawyer, V., Shi, Y., Colarco, P. R., Lyapustin, A. I., Wang, Y., and Remer, L. A.: Exploring Systematic Offsets between Aerosol Products from the Two MODIS Sensors, Atmospheric Meas. Tech., 11, https://doi.org/10.5194/amt-11-4073-2018, 2018.

Lund, M. T., Myhre, G., Skeie, R. B., Samset, B. H., and Klimont, Z.: Implications of Differences between Recent Anthropogenic Aerosol Emission Inventories for Diagnosed AOD and Radiative Forcing from 1990 to 2019, Atmospheric Chem. Phys., 23, https://doi.org/10.5194/acp-23-6647-2023, 2023.

Mangla, R., J, I., and S.s., C.: Inter-Comparison of Multi-Satellites and Aeronet AOD over Indian Region, Atmospheric Research, 240, https://doi.org/10.1016/j.atmosres.2020.104950, 2020.

Reynolds, R., Rayner, N., Smith, T., Stokes, D., and Wang, W.: An improved in situ and satellite SST analysis for climate, Journal of Climate, 15, 2002.

Schutgens, N., Sayer, A. M., Heckel, A., Hsu, C., Jethva, H., de Leeuw, G., Leonard, P. J. T., Levy, R. C., Lipponen, A., Lyapustin, A., North, P., Popp, T., Poulsen, C., Sawyer, V., Sogacheva, L., Thomas, G., Torres, O., Wang, Y., Kinne, S., Schulz, M., and Stier, P.: An AeroCom–AeroSat Study: Intercomparison of Satellite AOD Datasets for Aerosol Model Evaluation, Atmos. Chem. Phys., 20, https://doi.org/10.5194/acp-20-12431-2020, 2020.

Smith, S., Ahsan, H., and Mott, A.: CEDS v. 2021.04.21 gridded emissions data. Community Emissions Data System (CEDS)., https://doi.org/https://doi.org/10.25584/PNNLDataHub/1779095, 2021.

Sogacheva, L., Popp, T., Sayer, A. M., Dubovik, O., Garay, M. J., Heckel, A., Hsu, N. C., Jethva, H., Kahn, R. A., Kolmonen, P., Kosmale, M., de Leeuw, G., Levy, R. C., Litvinov, P., Lyapustin, A., North, P., Torres, O., and Arola, A.: Merging Regional and Global Aerosol Optical Depth Records from Major Available Satellite Products, Atmospheric Chem. Phys., 20, https://doi.org/10.5194/acp-20-2031-2020, 2020.

van Marle, M. J. E., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A.-L., Field, R. D., Arneth, A., Forrest, M., Hantson, S., Kehrwald, N. M., Knorr, W., Lasslop, G., Li, F., Mangeon, S., Yue, C., Kaiser, J. W., and van der Werf, G. R.: Historic Global Biomass Burning Emissions for CMIP6 (BB4CMIP) Based on Merging Satellite Observations with Proxies and Fire Models (1750–2015), Geosci. Model Dev., 10, https://doi.org/10.5194/gmd-10-3329-2017, 2017.

Vogel, A., Alessa, G., Scheele, R., Weber, L., Dubovik, O., North, P., and Fiedler, S.: Uncertainty in Aerosol Optical Depth From Modern Aerosol-Climate Models, Reanalyses, and Satellite Products, J. Geophys. Res. Atmospheres, 127, https://doi.org/10.1029/2021JD035483, 2022.

Watson-Parris, D., Christensen, M. W., Laurenson, A., Clewley, D., Gryspeerdt, E., and Stier, P.: Shipping Regulations Lead to Large Reduction in Cloud Perturbations, Proc. Natl. Acad. Sci., 119, https://doi.org/10.1073/pnas.2206885119, 2022.

Wei, J., Peng, Y., Mahmood, R., Sun, L., and Guo, J.: Intercomparison in Spatial Distributions and Temporal Trends Derived from Multi-Source Satellite Aerosol Products, Atmospheric Chem. Phys., 19, https://doi.org/10.5194/acp-19-7183-2019, 2019.

Young, S. A., Vaughan, M. A., Garnier, A., Tackett, J. L., Lambeth, J. D., and Powell, K. A.: Extinction and Optical Depth Retrievals for CALIPSO's Version 4 Data Release, Atmospheric Meas. Tech., 11, https://doi.org/10.5194/amt-11-5701-2018, 2018.

Zhang, J. and Rothrock, D.: Modeling Global Sea Ice with a Thickness and Enthalpy Distribution Model in Generalized Curvilinear Coordinates, Monthly Weather Review, 131, https://doi.org/10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2, 2003.

Zuo, H., Balmaseda, M., and Mogensen, K.: The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals, Climate Dynamics, 49, https://doi.org/10.1007/s00382-015-2675-1, 2017.