Atmos. Chem. Phys. Discuss., 12, C12468–C12479, 2013
www.atmos-chem-phys-discuss.net/12/C12468/2013/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Host model uncertainties in aerosol radiative forcing estimates: results from the AeroCom prescribed intercomparison study" *by* P. Stier et al.

P. Stier et al.

philip.stier@physics.ox.ac.uk

Received and published: 10 February 2013

Response to comments of referee 1

We would like to thank the reviewer for the helpful comments that substantially improved the manuscript. We very much appreciated the detailed remarks and hope to have addressed all raised issues.

This paper describes a comparison of nine global models subjected to a prescribed direct aerosol forcing. The models were subject to two different cases, a purely scattering aerosol and an absorbing aerosol, both with the same wavelength dependence of the optical depth, a fixed asymmetry factor, and the same

height distribution. In addition, the authors attempt to separate the uncertainties due to the model background states from the uncertainties due to the aerosol response. This is a timely and interesting paper, given the current high level of concern about aerosol forcing of climate and the need to determine the contributing factors to model uncertainties regarding that forcing.

General comments

1. While comparison studies such as this are only possible through the willing participation of investigators, it is unfortunate that none of the models here are exactly comparable to those currently being used in the IPCC AR5. However, the CAM5.1, ECHAM5.5, and LMDZ models are close. It would be helpful to have a brief statement regarding the similarities of these models to the versions used in the AR5.

We entirely agree. Unfortunately, it is not always possible to coordinate different intercomparison studies. We should point out that this intercomparison study primarily aims to attribute forcing uncertainties to host model effects in models and setups currently used to estimate aerosol radiative forcing and that we tried to ensure that models correspond as closely as possibly to the AeroCom Direct models used in the Myhre et al. (2013) paper, providing a significant contribution to the AR5 radiative forcing estimates. To further improve the comparability with the AeroCom Direct study (Myhre et al., 2013) we have now included three more models (HadGEM2, GEOS-CHEM and SPRINTARS) contributing to both studies in the revised manuscript.

However, we believe that the main conclusion of our study, that host model effects are a significant contribution to uncertainty in current radiative forcing calculations, applies equally to the range of current CMIP5 models used in AR5.

Specifically, both surface albedo and cloud fraction depend on the exact simulation setup and will differ between coupled CMIP runs and the nudged/fixed SST setup used in this study, even if the models were otherwise identical. We added the following

ACPD

12, C12468–C12479, 2013

Interactive Comment

paragraph to the manuscript to caution readers of over-interpretation

The used models are up-to-date configurations and some of the host models are almost identical to model versions used in the Coupled Model Intercomparison Project study CMIP5 (Taylor et al., 2012). However, the representation of host model parameters such as clouds and surface albedos, depends on the exact model setup and will differ from CMIP5 runs, even if they were conducted with the same model version. Nonetheless, we believe that the conclusions of our study equally apply to the range of models used in CMIP5.

2. Throughout the paper there are references to differences in results that may be due to differences in radiative transfer codes. It would be very helpful to have a brief comparison of the column radiative transfer (RT) forcing for each model due to the prescribed aerosol. A comparison as a function of solar zenith angle and atmospheric profile (primarily water vapor and ozone) would help identify the error expected simply due to the differences in RT codes. There is a reference to a paper by Randles et al. but I could not find a copy of this paper (only an abstract for a talk). I think it would be very helpful to have a longer discussion here of how the RT model results compare with the global results presented here and what uncertainties arise from the use of less accurate RT models.

We apologise that the Randles et al. paper was not available at the time of the submission of this manuscript. However, this study is now out in ACPD:

Randles, C. A., S. Kinne, G. Myhre, M. Schulz, P. Stier, J. Fischer, L. Doppler, E. Highwood, C. Ryder, B. Harris, J. Huttunen, Y. Ma, R. T. Pinker, B. Mayer, D. Neubauer, R. Hitzenberger, L. Oreopoulos, D. Lee, G. Pitari, G. Di Genova, J. Quaas, Fred G. Rose, S. Kato, S. T. Rumbold, I. Vardavas, N. Hatzianastassiou, C. Matsoukas, H. Yu, F. Zhang, H. Zhang, and P. Lu, Intercomparison of shortwave radiative transfer schemes in global aerosol modeling: results from the AeroCom Radiative Transfer Experiment, Atmos. Chem. Phys. Dis-

ACPD

12, C12468–C12479, 2013

cuss., 12, 32631-32706, http://www.atmos-chem-phys-discuss.net/12/32631/2012/ acpd-12-32631-2012-discussion.html, 2012.

This companion paper is devoted to discuss the implications of differences in the transfer codes in detail and its Section 3.4 directly compares the findings in the AeroCom Prescribed study to the results obtained in the Radiative Transfer study with a summary of the findings presented in Fig. 8 of Randles et al. (2012).

In this manuscript, we discuss the implications of the Randles et al. (2012) study in section 3.2.1, which we now extended to:

Results from the AeroCom Radiative Transfer Intercomparison (Randles et al., 2012) provide further insights: in this study, line-by-line benchmarks models show stronger atmospheric absorption in the purely scattering case than most schemes used in GCMs and CTMs, except Oslo-DISTORT used in Oslo-CTM2, which supports the higher values reported by some models here. The second highest absorption enhancement in the scattering case of the AeroCom Radiative Transfer study is simulated by another multi-stream model GSFC-FL, highlighting potential structural limitations of radiative transfer schemes used in GCMs. As the aerosol extinction is generally low in spectral regions of strong water vapour absorption, differences in the treatment of Ozone could be a contributor to these differences. This could in turn be affected by the spectral resolution of the models.

While it is not trivial to explicitly quantify the impact of radiative transfer codes in the framework of AeroCom Prescribed, we have now also added an estimate of the unexplained forcing errors (not explained by surface albedo or clouds), which includes (but is not limited to) radiative transfer effects. See Fig 19 and the additional description added to section 3.3:

If we assume that the errors in radiative forcing due to either surface (E_s) or cloudy (E_{cld}) albedo variation are independent, the remaining unexplained error can be defined as the $E^2 - E_s^2 - E_{cld}^2$, where E is the total error in aerosol radiative forcing

ACPD

12, C12468–C12479, 2013

Interactive Comment

(standard deviation in aerosol radiative forcing across the models). This unexplained error is shown alongside the errors due to surface and cloud albedo in Fig. 19. While the unexplained forcing errors show spatial correlation with either the surface albedo or cloudy albedo errors, potentially due to the co-variability of error sources or limitations of this analysis, it is important to note that their magnitude are significantly lower than the errors due to surface albedo or cloud effects. However, the light hatching in the plots of the unexplained forcing errors suggests that these are nevertheless not zero.

3. All integrated numbers presented in the paper are for the entire globe. However, much of the model diversity arises from variations in non-ocean albedo, particularly in polar latitudes. It would be very useful to have a comparison table similar to Table 3 for ice-free ocean (or ocean equatorward of 60 degrees) for both aerosol forcing cases. This comparison would remove a large degree of uncertainty due to surface albedo variability. I strongly encourage the authors to add this to the paper.

We entirely agree that land surface albedo, in particular in high-latitude regions with snow or ice cover is the dominant driver of the simulated forcing variability owing to surface albedos. On the other hand, anthropogenic aerosol sources are concentrated over land and the large spatio-temporal variability makes it hard to define a single representative averaging region (we could define a large number of alternatives to the proposed ice-free ocean ones). For this reasons, we believe that this variability is best expressed in lat/lon plots and the most realistic weighting is obtained through realistic aerosol distributions. This had been already addressed in the conclusions of the manuscript:

From our analysis it becomes clear that host model effects have a significant spatio-temporal variability, that may not match the aerosol perturbation in question so the derived global mean diversities may not be directly comparably to AeroCom Direct. However, recalculating the simulated forcing diversity as weighted average, using the ECHAM5-HAM2 anthropogenic optical depth as weighting factor, only slightly changes

ACPD
12, C12468–C12479, 2013

the global mean RF_{TOA}^{all} from -4.51 to -4.89 Wm^{-2} and reduces the inter-model absolute (relative) standard deviations from 0.70 to 0.54 Wm^{-2} (15% to 11%).

In addition, this has also implicitly been resolved (for a subset of the models) through the more realistic FIX1 experiment (in which the prescribed anthropogenic aerosols perturbation automatically gives high-latitude areas a low weight).

To illustrate the spatial distribution of diversity we have now added multi-model mean and standard deviation plots to each panel of lat/lon plots (Fig. 2 to 15), as now described in Section 2.1:

Results are generally summarised by mean values and standard deviations (SD) of global annual mean model values. In addition, we provide also plots of multi-model mean fields and standard deviations. To allow for comparability in the "diversity" of different parameters, we also report the relative standard deviation $RSD = \frac{|\sigma|}{|\mu|}$, where σ is the standard deviation and μ the mean value of the respective parameter. We should caution that the sample size across the models is very limited so that standard deviation is used here simply as a measure of the inter-model spread and should not be interpreted based on the underlying assumption of a Gaussian distribution, e.g. in the sense of confidence intervals.

4. Some aspects of the specified forcing lead to rather peculiar results that are inconsistent with reality. One of the most obvious is the strong forcing results over Antarctica (and over the Arctic to a somewhat lessor degree) in the absorbing aerosol clear-sky case (Figure 8). This is caused by the assumption of a high concentration of absorbing aerosol in an area where this does not actually occur. While I have no problem with the prescription of the aerosol for these runs, I think it would be very helpful to identify such features in the discussion and warn the casual reader that these features do not represent reality. I have similar concerns about the rather strong above-cloud absorption in ocean stratus regimes, although I cannot tell how much of the aerosol burden actually ends up

ACPD

12, C12468–C12479, 2013

above these clouds.

As you state, this is an idealised experiment and the resulting forcing fields (of experiments FIX 0,2,3) should not be mistaken for realistic aerosol forcings, which are considered in the FIX1 experiment and, less constrained, in the AeroCom Direct companion study Myhre et al. (2013). To further caution readers, we now have re-iterated this point in Section 2.1 as follows:

The setup of the AeroCom Prescribed simulation is outlined in Table 2. Readers are cautioned that experiments FIX0, FIX2, FIX3 are highly idealised, with unrealistic globally constant aerosol radiative properties, and that the resulting forcing fields of these experiments should not be mistaken with realistic aerosol forcings.

5. The authors define a relative standard deviation (standard deviation divided by the mean, both in absolute value). I understand the rationale but I wonder about its useful- ness when the mean values have small magnitudes. It is not clear that an RSD going from 10% to 90% is particularly important when it is driven by the mean dropping from 4 or 5 W/m2 to 1 W/m2 or less, rather than be significant changes in the standard deviation.

This is a good point and exactly for this reason we always cite absolute and relative standard deviations together. Unfortunately, the calculation of the “real” anthropogenic aerosol radiative forcing suffers from exactly this problem: the mean net forcing is small (and may be of comparable magnitude as host model uncertainties, leading to large RSDs of a small net forcing).

Specific comments

1. P25494, L5ff: I think this paragraph could do with a little more explanation. I assume from the statement here that all models other than ECHAM allow the aerosol forcing to affect the diabatic heating, which in turn may alter dynamical responses. If this is not the case, it would be helpful to state the situation more

ACPD

12, C12468–C12479, 2013

Interactive Comment

clearly. If ECHAM is relaxed to a pre-determined ECMWF reanalysis, then how much influence can the aerosol have? It seems to me that it must be quite limited.

Thanks for pointing this out. Indeed, this section needed some more explanation and has been changed as follows. Please note that the ECHAM FIX1 submission has been re-run in the same setup as the other models: using diagnostic instantaneous forcing, with no feedback of aerosol radiative properties to the meteorological fields so that the paragraph in question has been removed. Changed sections:

Section 2.1:

To maximise comparability, all simulations are performed in the same setup as used by the models in the submission to the AeroCom Direct experiment (Myhre et al., 2012) with model-specific aerosol distributions and radiative properties. The simulations were performed for one simulation year. Models report diagnostic instantaneous radiative fluxes, i.e. host model components other than aerosol radiative properties, such as clouds, are identical for all simulations by each model. Radiative forcing (RF) is calculated from monthly mean flux difference between the respective simulations.

Section 2.2:

Out of 12 submitted configurations, 6 are General Circulation Models to (GCMs), 5 are Chemistry Transport Models (CTMs) and 1, MPI-2stream, is an offline radiative transfer scheme. MPI-2stream prescribes cloud derived from the International Satellite Cloud Climatology Project as described in (Kinne et al.,2013. All models report diagnostic instantaneous radiative forcing, i.e. aerosol radiative effects do not feed back to the model meteorology, which remains identical for the radiative transfer calculations of the different simulations. The shortwave spectral resolution varies from 2 to 19 SW bands, and the complexity of the radiation parameterisations varies considerably. Model spatial resolutions vary from $4^\circ \times 5^\circ$ *in the horizontal and 19 vertical levels to* $1^\circ \times 1^\circ$ *in the horizontal and 72 vertical levels.*

ACPD

12, C12468–C12479, 2013

2. P25494, L22: It would be helpful to reference Table 1 when you first mention FIX0 since it is defined in Table 1.

Done. We now reference the experiment abbreviations in the list right at the start of Section 2.1 and also improved the structure of Table 1.

3. Same paragraph: This is a minor point as far as the paper is concerned, but cloud fractions less than 0.5 are simply too low. The use of simulators in AR5 models shows that cloud amounts in GCMs are generally too low, but most are still higher than 0.5.

We agree. However, such configurations are currently in use for a range of applications. The purpose of this study is to highlight the implications of such host model effects for aerosol radiative forcing.

4. P25495, L2: It appears that there are equally large cloud problems in the southern oceans and northern hemisphere storm tracks.

We agree that the variability of cloud fractions is large in various regimes. This is now also evident in the added mean and standard deviation plots (Fig. 2 and other panels). We have changed this sentence to:

The simulation of persistent high cloud-fraction stratocumulus decks off the western coasts of the Americas and Africa as well as in the storm-tracks differs significantly among models.

5. P25495, L21: Figure 16 is very difficult to read. I suggest that it be expanded, especially since there is no penalty for extra space in an on-line article. The caption inTable 3 should identify the symbols.

We agree. This was an artefact of the ACPD online view formatting and should have significantly improved in the ACP print format.

6. P25496, L16ff: As noted earlier, this is a particularly important comparison

ACPD

12, C12468–C12479, 2013

and should be expanded. Ozone absorption is probably not well represented by a low- spectral resolution model and may be the source of the discrepancy.

We agree and have addressed this issue in more detail in the Randles et al. (2012) companion paper, Section 3.4 and Fig. 8, cross-referenced in the relevant section. See also our response to comment 2 above.

7. P25501, L6ff. Figure 19 is really pretty complicated. I think you need to spend a little more time explaining what you are doing and how you are doing it. Also, I got confused by the concept of "shading" on a color plot. Presumably, the authors are referring to the hatching, but it is certainly not obvious, especially because the plots are tiny as presented here. Please clarify this situation.

Thanks for pointing this out. We have changed "shading" into "hatching". We have also extended the description caption of Fig. 19, the overal description in this section. Additions/changes:

We define sensitivity as in Eq. (2) as the slope of a regression of TOA SW forcing against either cloud or surface albedo (Fig. 19), with each data pair representing a different model. We use dark hatching to indicate regions where the sign of the sensitivity may change (due to removing a model). We use a light hatching to indicate where the variation (due to removing a model) in errors is less than 30%.

In the slope plots, the dark hatching tends to occur for small (absolute) values. These are the areas where sign of the slope is uncertain. In the error plots, the light hatching occurs for large values. These are the areas where the uncertainty in the error is small. The conclusion for both types of plot is essentially the same: wherever we see a strong signal, it also tends to be a reliable signal.

New caption of Fig. 19:

*Decomposition of the impact of host model errors on aerosol radiative forcings. The top row shows the **standard deviations** in surface and cloud albedo among the mod-*

ACPD

12, C12468–C12479, 2013

*els. The **sensitivity** (linear regression coefficient) of aerosol radiative forcing to either surface or cloud albedo in the models is shown in the second row (for the pure scattering FIX2-FIX0 case) and the fourth row (for the absorbing FIX3-FIX0 case). The **forcing error** due to either surface or cloud albedo (standard deviation in albedo times sensitivity) is shown in rows 3 and 5. Assuming independent errors, the (remaining) **unexplained error** in the radiative forcing is shown in the right most column, row 3 and 5. Here the white areas denote regions where the sum of squared errors due to surface or cloud albedo is larger than the deviation in aerosol radiative forcing across the models. Dark hatching indicates that the sign of the sensitivity may change (due to removing a model). The light hatching in the error plots indicates low uncertainty ($<$ 33%) in the error. Based on annual averaged fluxes for the models CAM-PNNL, GOCART-GEOS, GOCART-MERRA, LMDZ, LMDZ-39L, MPI-2stream, IMPACT, OsloCTM2, HadGEM2-ES, SPRINTARS, GEOS-CHEM remapped to a resolution of 1.875° x 1.875°.*

Technical comments

There are a number of annoying spelling and grammatical errors in the manuscript. I don't intend to identify them all but here are a few examples:

Apologies and thanks for pointing those out. We have proofread the revised submission carefully.

1. P25491, L1: this phrase is just dangling in space following the list. It should be written as a complete sentence.

We have moved this up to the begin of the sentence:

Aerosol radiative effects depend on a wider range of atmospheric parameters and their representation in host models used in the forcing calculation, henceforth collectively referred to as "host model effects", in particular on:

2. P25496, L18: "benchmark", no"s"

Done.

ACPD

12, C12468–C12479, 2013

Interactive Comment

3. P25496, L22: ozone is not a proper noun – no capitalization

Done.

4. P25498,L9: insert "albedo" after "single scattering

Done.

5. P25499, L17: "effect", no"s"

Done.

6. P25502, L28: "conducted"

Done.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 25487, 2012.