Response to review of **"***Concentrations and radiative forcing of anthropogenic aerosols from 1750-2014 simulated with the OsloCTM3 and CEDS emission inventory"* by Marianne T. Lund, Gunnar Myhre, Amund S. Haslerud, Ragnhild B. Skeie, Jan Griesfeller, Stephen M. Platt, Rajesh Kumar, Cathrine Lund Myhre, Michael Schulz

We thank the anonymous referee for the careful and thorough review of our paper, and the useful suggestions. Responses to individual comments are given below.

Anonymous Referee #2
Received and published: 18 September 2018

In this paper, the authors describe updates made to aerosol representations in the Oslo Chemistry Transport Model, compare the simulated aerosol to various observations, and use that model with a recent emission dataset to derive the time series of radiative forcing of aerosol-radiation interactions (RFari) over the industrial era. Their estimate of RFari is at the weaker end of latest AeroCom and IPCC assessments, and the authors work suggest that globally-averaged RFari has weakened since the 1980s. The paper is nicely written and to the point. Figures are well chosen and illustrate the discussion well. Sensitivity studies are interesting. There are a few aspects that can be clarified however: The justification for model parameters is often lacking and there is also a need to better link section 3.3 on RFari timeseries with model evaluation. Below I list the clarifications that I would like to see ahead of the paper's publication. Addressing those comments should not require additional analysis, so represent minor revisions.

1 Comments

• Abstract, line 38: Be more specific about what is new about the treatment of black carbon in OsloCTM3
Clarified. The text now reads:
*"The treatment of black carbon (BC) scavenging in OsloCTM3***"**

• Abstract, line 47: Aren't the AR5 forcing estimates for 2011?
Yes, this is a typo, thanks for noticing.

• Introduction, line 76: Are those the CMIP5 historical emissions documented by Lamarque et al., doi:10.5194/acp-10-7017-2010, 2010? If so, then reference the paper here.
Yes, reference included.

• Introduction, line 77: I seem to remember that historical finished in 2005, not 2000, in CMIP5.
Corrected. In fact, this sentence confused the inventory documented by Lamarque (1850-2000) with the CMIP5 experimental design with historical period 1850-2005. For clarification, we have rephrased slightly, adding referring to both Lamarque et al. 2010 and Taylor et al. 2012. The paragraph now reads:

"The Community Emissions Data System (CEDS) recently published a new time series of emissions from 1750 to 2014, which will be used in the upcoming CMIP6 [Hoesly et al., 2018]. CEDS includes several improvements, including annual temporal resolution with seasonal cycles, consistent methodology between different species, and extending the time series to more recent years, compared to previous inventories and assessments [e.g., Lamarque et al., 2010; Taylor et al., 2012]."

• Introduction, lines 114–116: Is that sentence still talking about scavenging?

Scavenging and other processes. Text clarified**:**
"(…) few studies have focused on impacts of scavenging and other processes on a broader set of aerosol species or the combined impact in terms of total aerosol optical depth (AOD)."

• Section 2.1, line 157: What are the finer spatial and temporal resolutions of the new transfer rates?
Monthly instead of seasonal and latitudinal rates established based on simulations of 10 instead of four emission source regions, covering not only the northern hemisphere, but giving near global coverage.
Now specified in the text:
"Specifically, the latitudinal transfer rates have been established based on experiments with 10 instead of four emission source regions and with monthly, not seasonal resolution."

• Section 2.1, lines 161–163, line 203, and Table 2; and Section 2.4, lines 284–285: The numbers selected for OC-OM conversion, hydrophilic fractions of BC and OM, aerosol solubility fractions, and absorbing OM fractions need to be justified. Where do they come from?
References to OC-OM conversion and hydrophilic fractions of BC and OM added.

On the OM fraction we have added the following:
"*Organic matter has a large variation in the degree of absorption [e.g., Kirchstetter et al., 2004; Xie et al., 2017], from almost no absorption to a strong absorption in the ultraviolet region. Here, we have implemented absorbing organic matter according to refractive indices from Kirchstetter et al. [2004]. The degree of absorption varies by source and region and is at present inadequate quantified: Here we assume 1/3 of the biofuel organic matter and ½ of the SOA from anthropogenic volatile organic carbon (VOC) precursors. The remaining fractions of biofuel, fossil fuel and marine POA and SOA (anthropogenic and all natural VOCs) are assumed to be purely scattering organic matter. As these fractions are not sufficiently constrained by observational data and associated with significant uncertainty, we also perform calculations with no absorption by OA."*

• Section 2.1, lines 169–170: How is the scaling of marine OM emissions done in practice? Sea spray is dependent on wind speed, so do you need to run the model a first time to get OM emissions, then scale them, then run again?
The scaling is done by multiplying with a fixed factor, determined from previous test runs, which depends on the sea salt production scheme used (here the factor is 0.5). The exact scaling depends on meteorology and resolution, found to result in inaccuracies of 3-5%. The marine OM emissions is checked after the run to make sure that no additional tuning of the scaling factor is needed beyond these uncertainties. We have specified the scaling factor in the text based on comments by both referees. The following has been added:
"The scaling factor depends on the chosen sea salt production scheme (see below) and to some degree on the resolution; here we have used a factor of 0.5. "

• Section 2.1, line 195: Why does the DEAD scheme need energy budget calculations?
Radiative fluxes are needed for calculation of boundary layer properties related to the dust mobilization.
For clarification we have added:
"As a minor update, radiative flux calculations, required for determination of boundary layer properties in the dust mobilization parameterization [Zender et al., 2003], now uses radiative surface properties and soil moisture from the meteorological fields."

• Section 2.1, paragraph starting line 198: Does the model account for reevaporation of precipitation? And re-suspension of dry-deposited aerosol?

The model treats evaporation (reversible and irreversible), but not re-suspension. Following this comment, and comment from anonymous referee #1, we have expanded the description of removal in Sect. 2.1, including these specifications.

• Section 2.3, line 249: Why fix the meteorology to 2010? Once reanalysis data is available (from 1979 for ERA, I believe), then it would be good to use the meteorology that corresponds to the actual year. There is a sensitivity study dedicated to the impact of meteorology, and that impact is fairly sizeable (lines 556–558). So using 2010 for all years does not seem ideal.
Because re-analysis data is only available for a limited period of the historical time series considered here, we maintain that keeping meteorological data constant is a reasonable experimental design for a study with our objective, for consistency and in order to isolate the effect of change in the emissions. Moreover, RF is calculated relative to pre-industrial, where we have no way of using "real" meteorological data. As recent study by Myhre et al. 2017 looked at the effect of emission changes over the 1990-2014 period, and in this case meteorology for each year was used in some of the models.
Our sensitivity test gives an indication of the order of magnitude impact on the aerosol burden and, as the direct RF scales quite linearly with aerosol burden, on the consequent RF. Using the rather extreme case of opposite ENSO phases, we find burden changes of 1-10%, suggesting that using constant meteorological input data is unlikely to change conclusions about the overall trends
Use of correct meteorological data is of course important for accurate distribution of concentrations and model evaluation, which is why we limit our validation to the year 2010.

• Section 2.3, paragraph starting line 254: It would be good to have more insight into the choice of sensitivity studies of removal. Why no SOLINC sensitivity study, for example?
The following text has been added for clarification:
"To modify the scavenging, we tune the fixed fractions that control aerosol removal efficiency in the model (see Sect. 2.1). Table 2 summarizes fractions used in the baseline configuration and the three sensitivity tests. A decrease and increase in efficiency of 0.2 is adopted for scavenging of all aerosols by liquid clouds (except hydrophobic BC and OM) and ice clouds, respectively. Note that there is no test with increased removal by liquid clouds, as, with the exception of hydrophobic BC, OM and SOA, 100% efficiency is already assumed. For ice clouds, we also reduce the efficiency to a fraction of 0.1, or 0.001 if the value is 0.1 in the baseline configuration. We note that these changes do not represent realistic uncertainty ranges based on experimental or observational evidence, as there are limited constraints in the literature, but are chosen to explore the impact of a spread in the efficiency with which aerosols act as ice and cloud condensation nuclei."

• Section 2.3, line 258: I do not think that the resolution of the control simulation has been given at this stage, but I may have missed it.
The resolution is given in the paragraph before.

• Section 2.4, paragraph starting line 270: It would be useful to have a Table summarizing the assumptions made about size distribution and refractive index (or MEC/MAC if that is what the model uses) for all modelled species, with references where appropriate. That information is crucial to understanding differences in radiative forcing efficiencies between models, yet is rarely given in model description papers.
We believe that this information is readily enough available from the references (Myhre et al. 2007, Myhre et al. 2009) in Sect. 2.4, with updates subsequently described.A table with this information has been included in the supplementary material. However, we have added a table in the supplementary material with changes in burden, AOD, AAOD, RFari, and normalized RF over the period 1750-2010 given for individual aerosol components, as well as the net RFari. This information has been commonly presented in previous studies and will allow the reader to better understand differences compared to the present analysis. The text has been updated accordingly.

• Section 2.5, line 297: “frameworks” – did you mean networks?
Yes, error corrected.

• Section 2.5, lines 321–322: What does the AeroCom validation tool do?
Probably poor/unclear chose of wording: The comparison between the model and measurement is done through the AeroCom data base, i.e., model data is interpolated and extracted for each location of the AERONET stations (this method is already described in a later paragraph). We have clarified the sentence:
“The comparison with AERONET data was done using processed through the validation tools available from the AeroCom data base hosted by Met Norway.”

• Section 3.1 starting line 354: It would be most useful to also include residence times for each species in Table 3. Perhaps in parentheses after the burden? Residence times are also crucial to understanding differences in aerosol distributions among models.
Good suggestion, we agree that this is useful additional information. We have included the residence times for the baseline simulation in Table 3, with remaining simulations in a separate table in the supplementary material (for readability).

• Section 3.1, paragraph starting line 356: The authors try to explain differences with the previous version of the model, which is great, but that is not done consistently. Why is sulfate burden 35% higher than in OsloCTM2? (Line 360.) Is that understood? And what explains the 25% higher OA burden? According to lines 369–370, the marine POA contribution is too small to explain that large difference.
We have added the following for sulfate:
“While the total SO_2 emission is only 5% higher in the present study than in the OsloCTM2 AeroCom III simulation, the atmospheric residence time of sulfate is 50% longer, suggesting that the burden difference is mainly attributable to changes in the parameterization of dry and large-scale wet deposition (Sect. 2.1).”
And modified the OA discussion:
“The OsloCTM3 estimate includes the contribution from marine OA emissions (Sect. 2.1), which may explain part of the difference as marine OA was included in some of the AeroCom II models, but not OsloCTM2. Additionally, the residence time of OA of 5.3 days is longer than in the OsloCTM2 AeroCom II experiment.”

• Section 3.1, line 377: “against output” – it is the other way around, model versus observations.
The sentence is technically correct as it stands, I think, but I can see that it would be more as expected to start with the model output and have switched it around for clarification:
“Figure 1 shows results from the baseline OsloCTM3 simulation against annual mean measured surface concentrations of EC, OC, sulfate and nitrate in Europe, North America and Asia.”

• Section 3.1, line 384: What correlation? Need to rephrase to clarify that sentence.
With CAWNET observations. Rephrased for clarification:
“In contrast, the correlation with CAWNET observations is generally similar to, or higher than, other regions/networks.”

• Section 3.1, line 414: Is being close to the AeroCom nitrate multi-model mean a good thing? How do models compare to observations in Bian et al. 2017?
Good point. We have added the following:
“Results showed that most models tend to underestimate ammonium concentrations compared to observations in North America, Europe and East Asia, with a multi-model mean bias and correlation of

0.886 and 0.47, respectively. Tthe OsloCTM3 shows good agreement with ammonium measurements in North America, but has a bias and correlation close the model average in the other two regions."

• Section 3.1, lines 426–428: Erroneous emissions have a big impact on model performance. Can't you re-run at all? That would be a shame.
We did re-run the year 2010 simulation for the assessment of impact described in this paragraph. In the revised manuscript, the documentation of the impact of the updated CEDS emissions on model performance is also extended to more species and the following text has been added:
"While repeating all simulations require more resources than available, we have performed an additional run for the year 2010. Figure S4 shows the comparison of modeled concentrations against IMPROVE measurements with the two emission inventory versions, CEDSv16 and CEDSv17. In the case of BC, the comparison shows a 5% higher correlation and 15% lower RMSE with the CEDSv17 than CEDSv16. A similar improvement is found for nitrate, with 26% higher correlation and 12% lower RMSE, while in the case of OC and sulfate, the difference is small (< 5%). Smaller differences of between 2-10% are also found in the comparison against measurements in Europe and Asia (not shown). Hence, using the updated version of the emission inventory has an effect on the model performance in terms of surface concentrations, but without changing the overall features or conclusions."

As global and national emission totals are unchanged between the two inventory versions, and we consider only the direct radiative effect, we believe that the implications for the RF estimates is small. We confirm this by also repeating the 1750 simulation as part of the revision: the RFari (2010-1750) is 2% stronger with the CEDSv17 inventory.

• Section 3.1, Figure 2: Could you show numbers in each slice of the pie chart in Figure 2? That would help make the comparison more quantitative.
Yes, done.

• Section 3.1, paragraph starting line 453: Figure 3d suggests that the model overestimates low optical depths but underestimates large optical depths. Is that correct? From experience, it is a common model deficiency, It may have even been noted in an AeroCom paper.
Yes, the model overestimates the very lowest measured AOD values. The most pronounced underestimation is seen for AOD around 0.1 to 1.0, but with a tendency towards more underestimation than overestimation also for the highest measured values. However, we are not aware of any AeroCom papers or multi-model studies documenting similar results that we could base a discussion around.

• Section 3.1, line 468: Small addition for clarity: "there are notable difference in model performance"
Included.

• Section 3.2, line 502: The analysis should point to Table 2 to make sense of sensitivities to wet removal. Species with the same scavenging fractions are likely to behave similarly in such sensitivity studies.
Good point. Included.

• Section 3.2, lines 519–520 and Figure 4: What is happening in North America? The change is in the Midwest, whereas I would have put the source regions more to the East.
Upon closer inspection of the burden of individual components, we see that this increase in total AOD in Central North America is mainly driven by an increase in ammonium nitrate. There is actually a dipole pattern in the ammonium nitrate change with different emission inventories: a lower burden is found in the Easternmost US (i.e., east of approx. 90W) and a notable increase in burden in the Midwest when ECLv5 and CMIP5 emissions are used. The former is likely due to the higher SO2 emissions in these inventories than in CMIP6, which is mainly confined to the east and south along the Mexican Gulf. The increased burden likely results from the increase in NOx emissions, which extends into the states of

Kansas, Oklahoma and Nebraska, and beyond. At the same time there is slight decrease in SO2 emissions in the western part of the US, which reduces sulfate concentration and hence competition for available ammonia. In the ECLv5 inventory, there are also higher emissions of NH3. We have added the following to the discussion:
"Over central North America the AOD is higher, mainly due to more ammonium nitrate, whereas the higher AOD over Eastern Europe and part of Russia is a result on higher sulfate concentrations."

• Section 3.2, lines 526–527 and Figure 4: What is happening with dust in METDTA? Also a teleconnection with ENSO?
Presumably, this refers to the ratio of total AOD in figure 4 and the area with values above one off the coast of North Africa, where the dust influence on AOD is prominent. The influence of met-data on dust occur through both production and transport/scavenging. There are some studies suggesting impacts of ENSO on precipitation in Sahel, and given that we have selected years of opposite ENSO phase, it is likely that we see an impact. However, other meteorological differences will also play a role and cannot be distinguished. While we cannot provide a quantitative answer, we have added the following to the discussion:
"There is also a notable change in the Atlantic Ocean, where mineral dust is a dominating species. The meteorological data can affect production, deposition and transport of dust directly, as well as indirectly through ENSO-induced teleconnections as suggested by e.g., Parhi et al. [2016]."

• Section 3.2, lines 530–531: Why such a strong dependence of nitrate on emission inventory? The vertical profile of ammonium nitrate formation is determined by temperature and competition from ammonium sulfate, so the role of emission inventories is not obvious. Which emission differences matter: ammonia or sulfur dioxide?
The change in nitrate formation is a complex interplay of changes in several gases (SO2, NH3 and NOx) and the differences in these, in turn, vary between inventory, as well as region. Hence, a significant impact of emission inventory is not unexpected. However, it is interesting to note the stronger effect on emissions, and less impact of changes to the scavenging, on the burden of nitrate, and the contrast to the other species. Based on this comment and the one above, we have added the following:
"In general, the burden of BC, OA and dust is significantly affected by changes in the scavenging assumptions, while nitrate responds more strongly to different emission inventories, likely due to the complicated dependence on emissions of several precursors and competition with ammonium-sulfate. We also note that at higher altitudes the absolute differences in the burden of nitrate are small."

With the data available from the simulations in the present study, we cannot quantify the roles of the respective emission difference. There are also likely strong seasonal and geographical variations. However, an inspection of the spatial distribution of differences in emissions, zonal mean concentrations and burden, can shed some light on the driver. It also important to point out that at high altitudes, nitrate values are small, as are the absolute differences.

For RCP/CMIP5 emissions, the global mean vertical profile of nitrate is lower than then using CMIP6 emissions throughout the atmosphere, mainly determined by a decrease in concentration between 10 and 40N. Compared to the CMIP6 inventory, RCP/CMIP5 emissions of both NH3 are lower in Asia, North Africa and South America. NOx is also lower, in particular in Asia. (SO2 emissions in India are somewhat lower as well, but we do not know to what extent less completion for available ammonia offsets an overall lower amount of precursors.) NOx is also lower in South Asia in ECLv5, while sulfate is higher, the combined effect being a lower nitrate burden. However, for nitrate there is a higher concentration around 15-40S and 35-50N, especially 800 and 400 hPa. This seems to arise, at least in part, from higher NH3 emissions, and areas with lower SO2, in China, South America and central Africa. So in summary, changes to all three gases seem to contribute, but with different roles in different regions.

• Section 3.2, lines 537–540: Again, should link to Table 2.
Done.

• Section 3.3 starting line 618: It would most interesting to link specific model deficiencies to errors in RFari estimates. This is done briefly in the discussion, but could be done more clearly in section 3.3. For example, he model generally underestimate surface concentrations. Does that bias cancel out when taking differences over the industrial era? Another example is that the model underestimates aerosol in Asia. Does that mean that recent RFari is underestimated? Or that similar underestimations would have happened over US/Europe when sulfate was high in those regions in the 1950-1980s? A final example is about high latitudes. There is a confident conclusion about RFari changes north of 70_N (lines 679–681), but do you really have confidence in your model in that region?
It is not straightforward to investigate indications of biases from comparisons to observations and the implication for RF. A bias compared to surface measurements may not necessarily imply that the total column has the same bias. Moreover, even if there is an underestimation of the total AOD, the influence of such a bias on RF depends on the mix of absorbing and scattering aerosol. Similarly, the aerosol mix is an important complicating factor the effect of underestimated or missing emission sources, for instance in Asia, where the uncertainties in the emissions are larger than in other regions. At high latitudes, the model does show an underestimation of BC concentrations, both at the surface and at higher altitudes. However, this is mainly an issue during winter and early spring, when the direct aerosol effect is small due to lack of sunlight. For other species and AOD, the availability of observations is a limiting factor. Additionally, we cannot perform a detailed evaluation of the simulated pre-industrial concentrations.

As the referee notes, the manuscript already discusses some of these issues. For the reasons noted above, we are reluctant to speculate too much beyond these more general points or attempt too specific links between model deficiencies and RF estimates. However, we have restructured Section 4 slightly, to make this discussion easier to follow and included a couple of additional issues. The paragraph now reads:

"A significant range from -0.6 to -0.13 W m-2 surrounds the central RFari estimate of -0.35 W m-2 from IPCC AR5 [Boucher et al., 2013], caused by the large spread in underlying simulated aerosol distributions. Deficiencies in the ability of global models to reproduce atmospheric aerosol concentrations can propagate to uncertainties in RF estimates. As shown in Sect. 3, the OsloCTM3 generally lies close to or above the multi-model mean of anthropogenic aerosol burdens from recent studies and is found to perform reasonably well compared with observations and other global models, with improvements over the predecessor OsloCTM2. In particular, recent progress towards constraining the vertical distribution of BC concentrations has resulted in improved agreement between modeled and observed vertical BC profiles over the Pacific Ocean with less of the high-altitude overestimation seen in earlier studies. However, as shown by Lund et al. [2018], there are discrepancies compared to recent aircraft measurements over the Atlantic Ocean. A remaining challenge is the model underestimation of Arctic BC concentrations. However, this is seen mainly during winter and early spring, when the direct aerosol effect is small due to lack of sunlight. In contrast, the higher emissions in the CEDSv16 inventory also results in an improved agreement with BC surface concentrations over Asia. In general, we find lower surface sulfate concentrations in the model compared with measurements. This could contribute to an underestimation of the sulfate RFari, which is weaker in the present study than in IPCC AR5. An underestimation of observed AOD in Asia is also found, however, the implication of this bias on RF is not straightforward to assess, as it is complicated by the mix of absorbing and scattering aerosols. (…)"

• Section 3.3, line 638: What is "OsloCTM3fast"?
This is a typo, there is no fast version of the model. Corrected.

• Section 3.3, line 675: “clearly” – Not that strong a signal really. Did the radiative forcing efficiency change with the change in emitting region? Combining by eye the two panels of Figure 8 would suggest that the change is rather small.

We think that the decline in net RFari north of 40°N is quite clearly seen in this figure, as is the strengthening between 10-30°N. However, the peak RFari is indeed significantly lower in 2014 than it was in 1970-90s. This is in large part due to the different evolution of scattering and absorbing aerosols, where sulfate burden is lower than in the late 20th century, while the BC burden has contributed to increase. There may also be an influence of different forcing efficiencies, however, since we only consider direct radiative effects, it is likely to be smaller than the impact of changing relative importance of scattering and absorbing aerosols. To clarify, we have modified the text:

“Over the past decades, there has been shift in emissions, from North America and Europe to South and East Asia. This is also reflected in the zonally averaged net RFari over time in Fig. 8b. RFari declined in magnitude north of 40°N after 1980, with particularly large year-to-year decreases between 1990 and 1995, and from 2005 to 2010, and strengthened in magnitude between 10°-30°N. The RFari also strengthened in the Southern Hemisphere subtropical region, reflecting incresing emission in Africa and South America after 1970. However, the peak net RFari is considerably weaker in 2014 than the peak in 1980. This mainly is due to fact that simultaneously with the southwards shift, the sulfate burden has declined, while the BC burden has increased steadily at the same latitudes, resulting in a weaker net forcing.”