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 #1
Received and published: 30 August 2018

General comments:
This paper evaluates the present-day aerosol distributions and radiative forcing in the latest OsloCTM3 chemical transport model. Surface concentrations, vertical distributions as well as aerosol optical depth are evaluated for the year 2010 against an extensive set of ground-based and satellite measurements. Uncertainties in the aerosol distributions are also assessed through a number of sensitivity studies assessing the uncertainty in the aerosol emissions, removal processes, meteorology and model resolution. The model is then used to assess the historical aerosol direct radiative forcing (RF) from 1750-2010. The net radiative forcing as well as the RF due to individual species are compared against the most recent published literature and CMIP5 estimates. Aerosol-cloud interactions are not assessed in this study.

This is a well-written, clearly presented, model evaluation manuscript. This paper does not document the latest updates to the OsloCTM3 model, but appropriate references are provided for this. Instead it documents the latest aerosol simulations with this model using the latest CMIP6 emissions inventory and subsequently derives the latest direct aerosol RF estimate. So I believe it is within the GMD scope as a Model Evaluation paper. This paper is timely given the CMIP6 project is now well underway and will provide a useful update to direct aerosol RF estimates in this regard. I recommend this paper to be published subject to a number of revisions I detail below.

Specific comments:
I struggled to appreciate the motivation for the additional sensitivity simulations conducted in this study. While these studies are useful and worth reporting the motivation for conducting these studies and link with the rest of the paper needs to be made clearer. This is probably most easily achieved in the Introduction.
Parts of the introduction has been rewritten in order to address this in a better and more explicit way. In particular, the following paragraph has been added:

"As accurate representation of the observed aerosol distributions in global models is crucial for confidence in estimates of radiative forcing (RF), these issues emphasize the need for broad and up-to-date evaluation of model performance.
The diversity of simulated aerosol distributions, and discrepancies between models and measurements, stem from uncertainties in the model representation aerosol processing. Knowledge of the factors that control the atmospheric distributions is therefore needed to identify potential model improvements and need for further observational data, and to assess how remaining uncertainties affect the modeled aerosol abundances and, in turn, estimates of RF and climate impact. A number of recent studies have investigated the impact of changes in aging and scavenging processes on BC distribution, focusing on aging and wet scavenging processes (e.g., [Bourgeois and Bey, 2011; Browse et al., 2012; Fan et al.,

2012; Hodnebrog et al., 2014; Kipling et al., 2013; Lund et al., 2017; Mahmood et al., 2016]), resulting in notable improvements, at least for specific regions or observational data sets. However, with some notable exceptions (e.g., [e.g., Kipling et al., 2016]), fewer 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)."

In Sect 2.3 Please report the length of each time slice simulation and how this impacts the signal to noise in resulting RF estimates. Are you running 20-30 years for each time slice?
Each simulation is one year, with six months spin up. Text has been clarified and now reads:
"All simulations are one year with six months spin-up. "

In Sect 3.3 please describe clearly either explicitly or through appropriate references how you calculated the RFari. Have you calculated an effective radiative forcing or used the more traditional radiative forcing metric? It is currently unclear. Also, how are the individual species RFari determined - through species only runs or is the RT code able to output this?. Please provide this detail in the manuscript.
Regarding the first half of this comment: The radiative forcing calculations described in detail in Sect. 2.4 of the paper, including a definition of RFari, i.e., forcing due to aerosol-radiation interactions, not including any rapid adjustments as for the effective RF. To clarify, we refer to back to the description in Sect. 3.3. The species-specific RFari is obtained by individual runs, where the concentration of the respective species is set to the pre-industrial level. This has been clarified in the text:
"The RFari of individual aerosols is obtained by separate simulations, where the concentration of the respective species is set to the pre-industrial level."

Use of older version of CEDS emission inventory: I am not convinced that the historical RFari will not be impacted at all by the choice of CEDS emission inventory. Changes in spatial distribution of emission could impact aerosol removal, transport and thus lifetime, temporal shifts in the distribution could also potentially impact the historical evolution of the ERF. While I understand the computational burden of repeating all tests, two runs using the new CEDS version with 1750 and for instance year 2010 emissions would allow you to quantify the impact on the RF fairly easily to allow you to justify it in the text. It looks like you may have the 2010 simulation already from you Fig S4 plots.
We have repeated the 1750 simulation with the new CEDS version 2017 inventory and calculate a 2% stronger net RFari (2010 relative to 1750) than with the 2016 version. Global burdens differ by 2-10% between runs with the two inventories, showing that there is some impact through removal and transport. The impact on net RFari is a combination of a slightly higher BC burden and lower sulfate and OA burden. Here limit the analysis to direct aerosol effects only. RFari scales more with the total burden changes, which are small. As the reviewer points out, the effect of spatial changes in the aerosol distribution could be more important for the ERF, i.e., also considering aerosol-cloud interactions.

Furthermore limiting your evaluation to BC is also questionable as I would expect notable differences in for instance SO2. I would request that you extend this evaluation over the USA to all aerosol species where you have observations.
We agree and have extended the comparison to organic aerosol, nitrate and sulfate, as well measurements over in Asia and Europe. Figure S4 has been updated and the following text is 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. The net RFari in 2010 relative to 1750 is 2% stronger with the CEDSv17 inventory, a combined effect of slightly higher global BC burden and lower burdens of sulfate and OA.”

Changes to the large-scale ice scavenging efficiency have a large control on the global aerosol distribution. The paper would benefit from a more detailed description of how the LS ice (and liquid) scavenging is parameterized in the model in Sect 2.1
The paragraph on removal has been expanded with more details and now reads:

“Aerosol removal includes dry deposition and washout by convective and large-scale rain. Rainfall is calculated based on European Center for Medium-Range Weather Forecast (ECMWF) data for convective activity, cloud fraction and rain fall. The efficiency with which aerosols are scavenged by the precipitation in a grid box is determined by a fixed fraction representing the fraction of this box that is available for removal, while the rest is assumed to be hydrophobic. The parameterization distinguishes between large-scale precipitation in the ice and liquid phase, and the OsloCTM3 has a more complex cloud model than OsloCTM2 that accounts for overlapping clouds and rain based on [Neu and Prather, 2012]. When rain containing species falls into a grid box with drier air it will experience reversible evaporation. Ice scavenging, on the other hand, can be either reversible or irreversible. For further details about large-scale removal we refer the reader to Neu and Prather [2012]. Convective scavenging is based on the Tiedtke mass flux scheme (Tiedtke 1989) and is unchanged from the OsloCTM2. The solubility of aerosols is given by constant fractions, given for each species and type of precipitation (i.e., large-scale rain, large-scale ice, and convective) (Table 2). Dry deposition rates are unchanged from OsloCTM2, but the OsloCTM3 includes a more detailed land use dataset (18 land surface categories at 1°x1° horizontal resolution compared to 5 categories at T42 resolution), which affects the weighting of deposition rates for different vegetation categories. Re-suspension of dry deposited aerosols is not treated.”

Technical corrections:
Line 70: 2011 - this should be 2010
IPCC AR5 give the forcing in 2011 relative to 1750. However, we see that there is a typo in our abstract, where 2010 is used. This has been corrected.

Line 170: It would be worthwhile to report here what global scaling factor you have used in the Gantt parametersation of marine OM. Gantt et al. 2015 I think use a global scaling factor of 6 but this is believed to be highly model dependent.
This is good point. Using the scaling factor of 6 from Gantt et al. produces too high emissions in our model. The scaling factor is also dependent on the sea salt production scheme and, to some degree, on resolution. With the sea salt scheme used here, we use a factor of 0.5. The text has been updated and reads:

“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. “

Line 281: Explicity state the aerosol concentration threshold below which you apply the Bond and Bergstrom method. Why do you change approaches? Again state the motivation for these different approaches in different regimes and why they need to be made. Does Zanatta lead to too high a MAC in

the low aerosol regimes? Has this been constrained by observations - if yes, provide appropriate reference here. Some discussion of the uncertainties in both approaches as this links to your uncertainties in BC RFari discussed in Sect 3.3

We have included the following description in the manuscript:

"The measurements in Zanatta et al. [2016] represent continental European levels. For very low concentrations of BC, the formula given in Zanatta et al. [2016] provides very high MAC values. We have therefore set a minimum level of BC of 1.0e-10 g m^{-3} for using this parameterization, and for lower concentrations we use [Bond and Bergstrom, 2006]. In addition, we have set a maximum value of MAC of 15 m^2 g^{-1} (637 nm) to avoid unrealistic high values of MAC compared to observed values."

Line 329-330: It would be useful here to be clear on what model variables you've sampled at 3 hour resolution and what ones you've sampled monthly and why you needed to do this.

Text has been clarified:

"the model data is linearly interpolated to the location of each station using annual mean, monthly mean (concentrations) or 3-hourly output (AOD), depending on the resolution of the observations."

Line 452: EBAS - I think this should refer to EMEP/ACTRIS. EBAS has not been previously mentioned.

Yes, thank you for pointing this out.

Line 468: Remove "as for surface concentrations,"

Removed.

Table 3: It would be useful to have the % change in burden listed alongside the burdens of all the sensitivity experiments. Perhaps in parentheses next to the burdens.

We tried this, but we think that it exacerbated the readability of the table without adding much useful information as these numbers are easily calculated from the absolute values. We have therefore chosen to leave this table as is.

Line 541-544: Can you actually make this statement? You do not show any impact on surface concentrations. I do not think this paragraph adds any value to the manuscript and would remove.

The evaluation against surface concentration measurements does suggest an improvement in model performance with the newer emission inventories, but similar performance in the other sensitivity tests. However, as this is first covered in the following section, we see that this paragraph seems a bit out of context here and we have removed it.

Line 555: Could the higher correlation and lower bias found in the 1x1RES test not just be a consequence of improved spatial sampling in your comparison and not actually due to an improved distribution of AOD? As you do not evaluate the species whose emissions would depend more strongly on resolution (ie: dust, sea salt) it's hard to quantify the benefit here.

This is a good point, and based on previous studies of the impact of sampling and the coarseness of typical model horizontal grids on comparisons with (in particular AOD), we would expect an improved sampling to play a role here. This point is now made in the manuscript:

"For both observables, the improvement in the 1x1RES simulation may result from a better sampling at a finer resolution, improved spatial distribution or a combination. "

Line 638: OsloCTM3fast - what is this?? There is no prior mention of a fast configuration!!

Thanks for pointing this out, it is an error and has been corrected.

Line 688: "These results emphasize the importance of assumptions related to the BC absorption" –> these results emphasize the importance of assumptions and uncertainties related to the BC absorption.

Corrected.

Figure 6: It would be more informative to show obs +/- 1 standard deviation instead of just + 1 std.
After trying different versions, we have rather added the 25/75 percentiles as a shading to show spread in both directions.

Table S2: Column 2 you use MNB instead of NMB
Corrected, this is a typo and should be NMB here as well.