Response to Referee #2 for the manuscript: “Assimilating aerosol optical properties related to size and absorption from POLDER/PARASOL with an ensemble data assimilation system”

Dear Editor & Reviewers,

Thank you for reviewing our submitted manuscript. Your comments helped us highlight and clarify some of our results better. Below you can find our responses for all of the raised questions.

Best regards,
Athanasios Tsikerdekis

Format:

Question
Answer
Quoted text, **added/changed text**, ~~removed text~~

Minor revisions:

Figure 1a: Why are there no AOD retrievals available over India?
The uncertainties shown in Figure1 are averaged over the period 20th of July to 28th of August 2006, which is during the summer monsoon in India where high precipitation and cloudiness occur. Therefore, the lack of retrievals over India is caused by the cloud screening of the algorithm. Note that multi-angle multi-wavelength photopolarimetric measurements have the ability to distinguish scattering caused by aerosol particles and cloud droplets (Stap et al., 2015) which facilitates cloud screening.

Figure 1a: Why are the uncertainties larger in the Southern Ocean?
The uncertainty of POLDER observations is defined after evaluating it with AERONET (FigureA1). Low AOD values (<0.05) have a very high relative uncertainty (~100%), but small absolute uncertainty. Southern ocean has very low AOD (<0.05) values over the whole course of the evaluated period. As noted in the manuscript, AERONET is a spatially sparse ground-based network of stations some generalization had to be made as far as the performance of POLDER in remote areas. The southern ocean is definitely one of these remote areas.

Lines 110-112: Can you provide an estimate of how many data points you gain by using the L1.5 AERONET retrievals rather than L2 retrievals. What is the effect of using L1.5 AOD retrievals on the POLDER uncertainty estimates?
POLDER uncertainty estimates were calculated using the AERONET Inversion L1.5 V3 dataset in cases were all variables were available or could be calculated (AOD_{550}, AOD_{865}, $AE_{550-865}$, $AAOD_{550}$, SSA_{550}). There are two major differences between L2.0 and L1.5. (i) Improved cloud screening in L2.0 and (ii) AAOD and SSA calculation use only cases of AOD>0.4 in L2.0. If L2.0 data were used instead of L1.5 then we would have ~4 times less available collocated data points between POLDER and AERONET. Our POLDER uncertainty estimates are conservative because all POLDER – AERONET differences are attributed to POLDER while in some cases half of it comes from AERONET (for other properties than AOD, for AOD itself the AERONET uncertainty is 0.01).

Line 184: Since the model resolution is (1.875◦ x 1.875◦) and the POLDER resolution is 1 x 1 degrees, do you use some kind of super observation approach for the assimilation?

The current version of the data assimilation system does not account for representations errors. We are planning to include a method that account for that. However it is noted that cloudiness is the biggest driver in representation errors in AOD. Thus if the clouds are represented correctly in the model the representation error would be zero. It is noted that ECHAM-HAM calculates aerosol optical properties on cloud free part of each grid and POLDER retrieves only in non-cloudy conditions.

Line 343-344: Since you assume the same level of uncertainty for both the natural and anthropogenic aerosols, does this approach not underestimate the background error covariance?

Thank you for this comment. Ideally we should have picked different uncertainty for natural and anthropogenic sources, though the version of the data assimilation used in the paper could perturbed emissions only by species and not by emission sector (e.g. Industrial, ships, fires etc), thus the distinction of anthropogenic and natural sources for some species (e.g. OC, BC, SO4, SO2) was not possible at the moment. In our follow-up research we are planning to perturb with high ensemble standard deviation at least the purely natural species (DU and SS).

Line 382: What is Section 0?

Corrected to subsection 3.2.

Line 386: I am not sure if I understood the daily assimilation set-up correctly. Do you run the daily forecast from 00 to 23 hour first and then call the LETKF code for the assimilation of POLDER observations at 00, 06, 12, and 18 hours? Does the next day forecast use initial conditions from the 18 h assimilation? If this is correct, what is the benefit of assimilation at 00, 06, ad 12 hours because we are not accumulating the benefits of assimilation at these times in the forecasts.

Thank you for giving me the opportunity to clarify and adjust the text. Background forecast is run from Day1 00:00:00 to Day2 00:00:00. POLDER observations are assimilated for 00, 06, 12, 18 and they adjust the mixing ratio at Day2 00:00:00 (analysis). Then the system is restarted using as initial condition the analysis conditions at Day2 00:00:00. The manuscript text was adjusted accordingly:

The daily cycle of data assimilation involves daily forecasts **(Day$_t$ 00 UTC to Day$_{t+1}$ 00 UTC)** of all perturbed ensemble members. Upon completion of these simulations, the LETKF code is called which performs a spatial collocation of the simulated (ECHAM-HAM) and the retrieved (POLDER) observations for four temporal time-steps (00, 06, 12, 18 UTC). Subsequently LETKF computes a new analysis state vector (ECHAM-HAM aerosol mixing ratio) at ~~the last time step of the day~~ **Day$_{t+1}$ 00 UTC**, which serves as initial conditions for the next day's forecast. The process is repeated till the end of the data assimilation experiment.

Figure 15: The correlation coefficient for AE decreases from the Control to Total experiment. What is the reason for that? Is it because POLDER AE has a lower correlation coefficient compared to control?
Thank you for this comment. This may indeed be the reason. However, we consider a difference between 0.593 and 0.606 not significant.

Figure 16: Should the title of y-axis be MASS AAOD in panels b and f?
Indeed that is true, thank you for pointing this out. Figure was updated.

Changes:

Line 49: Change “disentangles” to “disentangle”.
Line 67: Change “colour, polarization” to “colour, and polarization”
Line 161: Change “~0.03 is” to “~0.03 in”.
Line 353: Change “initially” to “initial”.
All of the abovementioned changes have been addressed.