Dear Reviewer,
we thank for your suggestions, comments, questions, and for your criticism. Here, we would like to clarify and address your main criticisms (blue). This is then followed by revisions of our paper (red) made in light of your comments (in black).
Yours sincerely,
Christine Pohl, on behalf of all co-authors

1) Uncertainty in Ext can exceed 100 %.
→ The reviewer bases his/her comments and conclusions in particular on Fig. 2, which shows uncertainties between -45 and 140 % in the aerosol extinction coefficient. We thank the reviewer for motivating us to focus on these numbers because in rechecking them we noticed a mistake in this Figure. The wrong number density profiles that we used to calculate the extinction coefficients led to the large uncertainties. After the correction, the uncertainties in the extinction coefficients reduce to -15 - + 25 % (see revision below). Although this does not change the other results of the manuscript or the reviewer's comments, the magnitude of the range of uncertainty is much smaller than that addressed by the reviewer. In particular, we disagree with the reviewer's statement that the uncertainties of the extinction coefficients are greater than 100% (see revision below).

2) The number density N has to be assumed.
→ No remote sensor can measure or retrieve N without making any assumption. Even for the occultation technique, where the assumptions are not so strong, the overall shape of the particle size distribution (typically unimodal lognormal) and shape and composition of the aerosol particles (typically spherical particles consisting of sulfate acid) needs to be assumed.
In addition the lack of sensitivity for particles smaller than ~ 10 nm might be an issue.

In general, there are only a few in situ instruments that can measure N, one balloon-borne and 2-3 air-borne. Remote sensors using the scattering technique can be used to infer N by assuming some other parameter of the particle size distribution, such as median radius and/or distribution width. This, however, does not eliminate the uncertainty related to one fixed parameter. In terms of the used assumptions, the SCIAMACHY retrieval is not significantly different from any other retrieval for remote sensing data. Thus disqualifying SCIAMACHY based on this limitation would also disqualify many other published and well-recognized retrieval algorithms.

3) The number density N of the Wyoming OPC can vary by up to a factor of 10.
→ The measurements made indicate that N does not vary by a factor of 10. This is clearly shown in the attached plots (see p. 10-11). A variation of 0.5 - 2.0 is a better assessment of the observed range. A paper was recently submitted to JGR which reviews the history of total aerosol measurements from Wyoming (Norgren et al., in review: Measurements of total aerosol concentration in the stratosphere: a new balloon-borne instrument and a report on the existing measurement record). A figure in that paper shows that at 500 K +/- 20 K, except for 5 measurements out of 132, there were no

changes > about a factor of 2, including the Pinatubo period. For the SCIAMACHY period there were 3 such measurements between 2007 and 2009, but these profiles provide “only“ 4 times greater number densities than assumed. Uncertainties in the retrieved (r_g, σ_g) and calculated (Ext, r_{eff}) aerosol quantities due to such an underestimation of N are now discussed in the manuscript (see revision below).

4) Without knowing N following volcanic eruptions the retrieval is not very useful.
→ The values of N after volcanic eruptions are not known. This is because it is far impossible for instruments to be launched in time to make the measurements. The timing of measurements is very critical. If new particles are formed at high concentrations they quickly coalesce or are scavenged by larger particles, so their lifetime is very short. At Laramie two new particle layers have been measured in August of 1991. Neither of these layers appeared in any further measurement. Overall the N profile returned to pre-Pinatubo levels in about a year after peaking about a factor of 2 above pre-Pinatubo levels. Similar perturbations were not observed following any further eruption in the record. Of course this doesn't rule out an effect when one is much closer than Laramie to one of these volcanoes. But again the issue or handicap for the retrieval algorithms to infer the PSD parameters from SCIAMACHY observations is similar for the measurements of all other remote sensing observations. See point 2.

Major Issues

This methodology is highly sensitive to the assumed, a priori, number density profile as demonstrated in Fig. 2. Here, the authors showed that changing the a priori N profile (by factors of and 2) changes r_g by ≈ +-30%, sig_g by ~ +-6% (no big deal), r_e by ~ +- 15-20 %, and extinction by -45% to + 140 % (it is interesting to note that this scaled nearly linearly).

We apologize here, a mistake was made when calculating the extinction coefficients. Instead of using the N_ECSTRA profile (blue line) which was assumed for the retrieval, the „true“ N profiles (coloured lines) were mistakenly used to calculate the retrieved Ext. Ext is the product of the particle number density and the particle scattering cross section integrated over the number size distribution. This leads to the linearly scaled uncertainties in the retrieved Ext.

Granted, all bodes well when the a priori N matches current conditions (as the authors demonstrated in Fig. 5 and elsewhere). However, under volcanically active conditions it is entirely reasonable that the a priori is more than a factor of 2 different (looking at the Wyoming OPC record I see changes in excess of a factor of 10 after eruptions within the SCIAMACHY time period). Taken to an extreme, how would this method perform after Pinatubo or Hunga Tonga?

N does not vary by a factor of 10. This is clearly shown in the attached plots (p. 10-11). A variation of 0.5 - 2.0 is much more indicative of the record. So we are not sure which part of the record the reviewer is referring to to make that statement. A paper was recently submitted to JGR which reviews the history of total aerosol measurements from Wyoming (Norgren et al., in review: Measurements of total aerosol concentration in the stratosphere: a new balloon-borne instrument and a report on the existing measurement record). A figure in that paper shows that at 500 K +/- 20 K, except for 5 measurements out of 132, there were no changes > about a factor of 2, including the Pinatubo period. For the SCIAMACHY period there were 3 such measurements between 2007 and 2009, but these profiles provide „only“ 4 times greater number densities than assumed.

While the authors evaluated the influence of an incorrect a priori (via Fig. 2) there remains 1 glaring shortcoming of the method: the “real” N profile is unknown therefore we do not know how much uncertainty this introduces to the retrieval and we cannot quantify the uncertainty of the inferred PSD values and the derived extinction coefficients. What we do know is that this uncertainty can be substantial. If the authors were to limit their analysis to conditions when N is stable then they could make reasonable guesses for their a priori (that’s basically what they do here, using the OPC record). However, that is not interesting. The interesting bits are in the post-eruption atmosphere when the N profile and PSD parameters are most dynamic! In short, we know that the PSD parameters and extinction coefficients as derived from SCIAMACHY data are wrong. . . but we have no gauge for how wrong they are.

Yes, but in these conditions all instruments are struggling with limited information to help. And because of the limited range of variability of N, the retrieval errors will be smaller than anticipated by the reviewer (see new sensitivity tests in the current version of the manuscript).

Unfortunately it’s not just the quantitative results that are suspect but we must also suspect and qualitative interpretation of the data as well.
I really like this paper so this leaves me with a dilemma. I am left questioning how I would use this data and what is the ultimate purpose of this paper (i.e., what does the community now know that we did not before). I think what we now is this: PSD parameters can be inferred from SCIAMACHY data and these inferred parameters have modest sensitivity to the a priori N. The calculated extinction is much more sensitive. However, given the extreme range of N in the Wyoming OPC record after major eruptions within your time period I have to conclude that this methodology is useful only during stable/background conditions and not reliable in the aftermath of volcanic eruptions.

As already mentioned above, most of the balloon-borne recorded N is quite stable. Accordingly, the methodology is not only applicable in stable conditions, but also in the aftermath of some volcanic eruptions and biomass burning events.
We investigated the influence of volcanic eruptions on the SCIAMACHY PSD retrieval based on a newly created synthetic data set. We added the following

text to the manuscript:

To conclude, the retrieved (r_g, $σ_g$) and calculated aerosol characteristics (r_{eff}) depend on the assumed a priori N profile. The closer this assumption is to reality, the more precisely the aerosol characteristics can be retrieved. However, the number density varies in the reality to an unknown extent. Therefore, it is impossible to quantitatively estimate the retrieval uncertainty caused by the number density assumption. We can only provide uncertainty limits.
Regarding the balloon-borne OPC measurements from 2002 to 2012, a variation of N by a factor between 0.5 and 2.0 encloses ≈ 80% of the variation observed in the Wyoming record. This means the above mentioned uncertainties are representative for most of the SCIAMACHY record. In the remaining 20% of cases, the balloon-borne OPC record reveals number densities in around 18 km altitude that are about four times larger than the N_{ECSTRA} profile assumed in the SCIAMACHY retrieval. They originate from major volcanic eruptions of Tavurvur (2006), Kasatochi (2008), and Nabro (2011).
In the aftermath of volcanic eruptions, an underestimation of the a priori N profile by a factor of four does not necessarily lead to a doubling of the above mentioned uncertainties. The latter rather depend on the true PSD of the aerosol plume, i. e., the interplay between r_g, $σ_g$, and N. We repeated the simulations using the altitude-dependent aerosol profiles shown in Fig. 2 (black lines) but perturbed the profiles below 25 km altitude according to OPC measurements and SAGE III/ISS retrievals (Wrana et al., 2021) in post-eruption periods. The number density at 18.4 km altitude was four times larger than the assumed a priori number density N_{ECSTRA}. We considered cases with increasing and decreasing volcanic particle size. In the best case, the retrieval uncertainties at 18.4 km are at 30% (r_g), -8% ($σ_g$), and 18% (r_{eff}). In the worst case, they are twice as large. Uncertainties in Ext are between -70 and 50% (not shown).
…
To summarize […] the spatiotemporal distribution of the stratospheric aerosol number density is unknown. The SCIAMACHY retrieval uses assumptions that lead to errors in the derived aerosol characteristics. Radiative transfer simulations where the a priori and true N differ by up to a factor of four imply uncertainties of +-30% (r_g), +-8% ($σ_g$), and +-20% (r_{eff}) during volcanically quiescent and some post-eruption periods. However, uncertainties in volcanic plumes can also double depending on the PSD. Ext has an uncertainty of +-25% during volcanically quiescent periods and of -70 - 50% during post-eruption periods.

This limitation is systemic throughout the paper and is inherent within the methodology itself and I currently see no path to salvaging it. It is for this reason that I cannot recommend this paper for publication.
I recognize that my view may be in the minority and, should the editor decide to allow publication, then I fully support him in this decision.

We hope we have explained the intrinsic need for a priori knowledge to constrain the retrievals of the aerosol PSD. We have quantified the size of the error and use the best available knowledge from in situ climatologies to constrain the retrievals.

Specific and Minor Comments

– page 7, line 197: “unambiguity” should be “ambiguity”?
Corrected.
– page 9, lines 268–269: “the retrieved PSD parameters and the assumed number density are used to calculate the effective radius (Eq. (5)) and the extinction coefficient (Eq. (6)) of the aerosol particles” Earlier in the manuscript you stated that N can be fixed through space/time because N plays a minor role in the retrieval process (fair enough). However, here you see how N plays a crucial role in calculating some derived parameters (especially extinction). Perhaps a statement regarding this dependence is appropriate here.

We have added the following text:
Note that both the retrieved (r_g, sigma_g) and the calculated parameters (r_eff, Ext) may slightly depend on the choice of the a priori number density profile. However, it will be shown in Sect. 6 that the strong correlation between the PSD parameters can compensate for retrieval errors in the calculated parameters, provided that the a priori number density profile does not deviate considerably from the true profile.

– page 10, line 283: “classes” should be “bins”?
Corrected.
– page 17, lines 475–479: “A quantitative error estimation of both assumptions, the Lambertian surface and the a priori N profile, is therefore not possible for real retrievals. We can only point to these sources of uncertainty.” This is an accurate statement and it is highly unfortunate. In my view, this is the dominant shortcoming of this method: you know the numbers are wrong but you don’t know by how much. This uncertainty will be more pronounced immediately after major events (N can change by a factor of 10 or more, which is FAR more than the ”doubling” you modeled). It may be necessary to explicitly tell the reader of this shortcoming post-eruptions.

We do not agree with this statement. This is because our analysis of the Wyoming CN record shows that a doubling of the CN concentration accounts for 70-90% of the variation observed, and there are almost no observations that exceed a factor of 10 of the mean, especially during the SCIAMACHY years.

We have added the following text based on new simulations that have been done:
The SCIAMACHY retrieval uses assumptions that lead to errors in the derived aerosol characteristics. Radiative transfer simulations where the a priori and true N differ by up to a factor of four imply uncertainties of +-30% (r_g), +-8% (σ_g), and +-20% (r_eff) during volcanically quiescent and some post-eruption periods. However, uncertainties in volcanic plumes can also double depending on the PSD. Ext has an uncertainty of +-25% during volcanically quiescent periods and of -70 - 50% during post-eruption periods.

– page 18, lines 483–484: In what way is this “striking”?

We have clearly identified the volcanic signatures in the aerosol characteristics. We have consistent results for different volcanoes. The aerosol characteristics have a clear dependence on the eruption strength, as shown in the comparison of, e.g., Soufrière Hills with Nabro. Such eruptions are also unlikely to cause large perturbations in the total particle number density. Few eruptions do. So the uncertainty in number density is not a large and significant factor.

We have changed the sentence to:
The changes in the aerosol characteristics after volcanic eruptions are readily identified.

Fig. 2 showed that underestimating number density (yellow line) results in over estimation of particle size and under estimation of distribution width. This results in an over estimation of approx 100% in the extinction coefficient and an over estimation of approx 20% for r_e. Undoubtedly, the a priori N value in your model is too low after these eruptions, which puts you squarely in the situation I just described (i.e., over estimation of r_g, over estimation of extinction, etc.).

We apologize for a coding error. After correction, the overestimation of Ext is ~ 25 %. The a priori N value in our model is too low after volcanic eruptions. This does not necessarily result in a larger overestimation of Ext. Uncertainties in Ext depend not only on the underestimation of N but on uncertainties of all three PSD parameters. Underestimation of Ext is also possible as can be seen in, e.g., Fig. 6.

I don't doubt that extinction increased, I don't doubt that particles became bigger, and I don't doubt that distribution width decreased. However, particles do not always get bigger after eruptions as some of your co-authors have demonstrated (https://acp.copernicus.org/articles/23/9725/2023/). Therefore, this leaves the reader wondering how much of the variabilty shown in Fig. 3 is a by-product of a wrong number density. Given the level of uncertainty in this method I do not believe that you can say that your data unambiguously proves (much less quantified) these changes occurred. Undoubtedly changes are expected, but at this point I think that the most defenceable statement that can be made, based on your product, is that things changed. . . by some amount.
– page 19, Figure 4 caption: The OPC record reports 2 modes (fine and coarse mode). Did you use both modes in calculating the extinction coefficient? The coarse mode can have a disproportionate impact on extinction.

We assumed a single mode lognormal PSD to calculate Ext. This assumption does account for the coarse mode because the entire distribution is used to optimize the fitted three lognormal parameters. We could also compare the Ext with those calculated using the full bimodal distributions, but we doubt this would change our understanding significantly.

– page 19, Figure 4: Panels (d) and (i) do not make sense. Why do the red and blue lines cross each other at ≈18 km at not at ≈20 km (i.e., where the 2 a priori N lines cross each other in panels (c) and (h))? All other panels have the red/blue intersection at the same altitude so why are (d) and (i) different?

Thank you for being thorough. We have checked the Ext profiles (panels d and i) again, their calculation is correct. Calculating the normalized extinction coefficient ($\hat{E}xt$ = Ext/N with N as the total number density) will result in red and blue lines crossing each other at ~20 km. Below 20 km, the red line ($\hat{E}xt$ using N_Wyoming) is lower than the blue line ($\hat{E}xt$ using N_ECSTRA). These profiles are then multiplied by the respective N profiles. Where N_Wyoming > N_ECSTRA. The absolute values depend on the size of factors leading to an intersection of the red and blue lines at ~18 km.

– page 19, Figure 4 caption: What about the light red and light blue colors? Can you define those here so the reader need not search the text for the explanation?

We have revised Figure 4. Light red and light blue colors have been eliminated by shading areas.

– page 20, lines 522–524: "Remarkable are the similar profile shapes from SCIAMACHY and OPC on 7 November 2009 in case of r_g and σ_g by assuming the a priori N based on balloon-borne measurements. This is due to the similarity of the SCIAMACHY-assumed and OPC-measured N profiles." This is no surprise (as the authors state, this is due to the extreme similarity between the current OPC N profile and the climatology, that was based on OPC data). What this tells me, yet again, is that the profile is entirely dependent on the a priori N profile.

This is certainly true as it has been pointed out in the paper already. No remote sensing instrument provides the true N profile, in fact there are only a few in situ instruments that can do that, one balloon-borne and 2-3 aircraft borne. Consequently, the SCIAMACHY retrieval algorithm requires an assumed number density profile.

– page 20, lines 526–527: "good agreement of the extinction coefficient from OPC and SCIAMACHY, regardless of the assumed a priori number density (Fig. 4(d,i))" I disagree on the interpretation of this figure, but I admit that I am struggling to find an interpretation and more information would be helpful. Panel (d) certainly looks promising (SCIAMACHY and OPC are in good agreement) and I would be surprised if it were bad. However, panel (i) is less impressive. While the shapes are in good agreement the OPC extinction is larger by a factor of 2-3 (is this what is meant by good agreement)?

We disagree with the statement that SCIAMACHY and OPC Ext differ by 200 – 300 %. The difference is < 33 % in Fig 4d and < 43 % in Fig 4 i. However, it could indeed be argued that this is not a „good" agreement. We have revised the relevant sentence:

One central statement of Fig. 4 is the agreement between the extinction coefficients from OPC and SCIAMACHY, with deviations within 33 % (Fig. 4d) and 43 % (Fig 4i), regardless of the assumed a priori number density.

Undoubtedly all of this variability is driven by differing number densities. However, it is important for the reader to understand how the OPC extinction coefficients were calculated here: did you only use the first mode or both modes? If you only used the first mode then, especially after an eruption, we can reasonably expect the second/coarse mode to be enhanced and have a disproportionate impact on extinction. Here is the point: currently the difference between SCIAMACHY ext and OPC ext is 200%-300%, if you include the second OPC mode in calculating ext the difference will become larger and this does not qualify as ”good agreement”.

We use monomodal PSDs derived from OPC data to calculate Ext. We add the following sentence in Sect. 5.1:
The unimodal PSDs are used for comparison with SCIAMACHY-derived aerosol characteristics. From these, the aerosol extinction coefficients are calculated according to Eq. 6.

– page 20, line 528: “. . . the three PSD parameters remain consistent with each other.” I do not understand what is meant by “consistent.” Can you please clarify?

We have changed the sentence:
It can be explained by the strong correlation of the PSD parameters. For example, if N is overestimated, r_g and σ_g change accordingly and to a certain extent. This enables the correct calculation of aerosol characteristics such as Ext from the combination of all three PSD parameters.

– page 21, Figure 5 caption (and corresponding text): “. . . relative errors. . . ” This assumes that the OPC is correct, which it is not. You stated on the previous page that disagreement between OPC/SCIAMACHY was driven, in part, to differing sampling volumes and because the OPC only sampled the edge of the aerosol plume while some of the SCIAMACHY profiles were collected within the plume. Should be reworded to ”percent difference” (or something comparable) here and throughout the text.

Done. We have replaced „relative errors“ by „relative differences“.

– page 21, lines 539–540: “Since SCIAMACHY is not sensitive to stratospheric aerosols with r_g ! 0.06 µm (Malinina et al., 2019), corresponding OPC PSDs are excluded from the comparison” This is not correct and these smaller r_g’s should be included. They will not contribute much to extinction at 750 nm, so no big deal, but, as written, you are not doing a valid comparison.

SCIAMACHY is not sensitive to aerosols smaller than r_g = 0.06 um. This fact is considered in the retrieval algorithm by using a lower threshold in the particle size retrieval: The mode radius cannot be smaller than 0.05 um. Accordingly, for the retrieved sig_g, this corresponds to a lower r_g limit of 0.06 um.
If OPC-measured aerosol characteristics with r_g lower than 0.06 um are compared with corresponding SCIAMACHY-retrievals, there will be disrepancies which are only caused by the user-defined retrieval constraints. This comparison will be meaningless and could be misleading for the reader. For this reason, we exclude the small aerosols.

We have revised the corresponding sentence:
Radiation scattered by aerosols with PSDs of r_g<0.06 um is below the sensitivity limit of SCIAMACHY (Malinina et al., 2019). Corresponding OPC measurements and collocated SCIAMACHY retrievals are therefore not in the

comparison.

– page 21, lines 547–548: “. . . Ext depend only slightly on the choice of the a priori N profile.” This is not what Fig. 2 tells me; N is highly important.

This statement just refers to Fig. 5i) where it is obvious from the overlaying of the red and blue areas and lines that the initial choice has little impact.

– page 22, lines 566–567: “. . . to obtain a sufficient number of collocations.” Sufficient for what? You now have 4255 coincident profiles. . . how many did you have when the maximum distance was only 200 km?

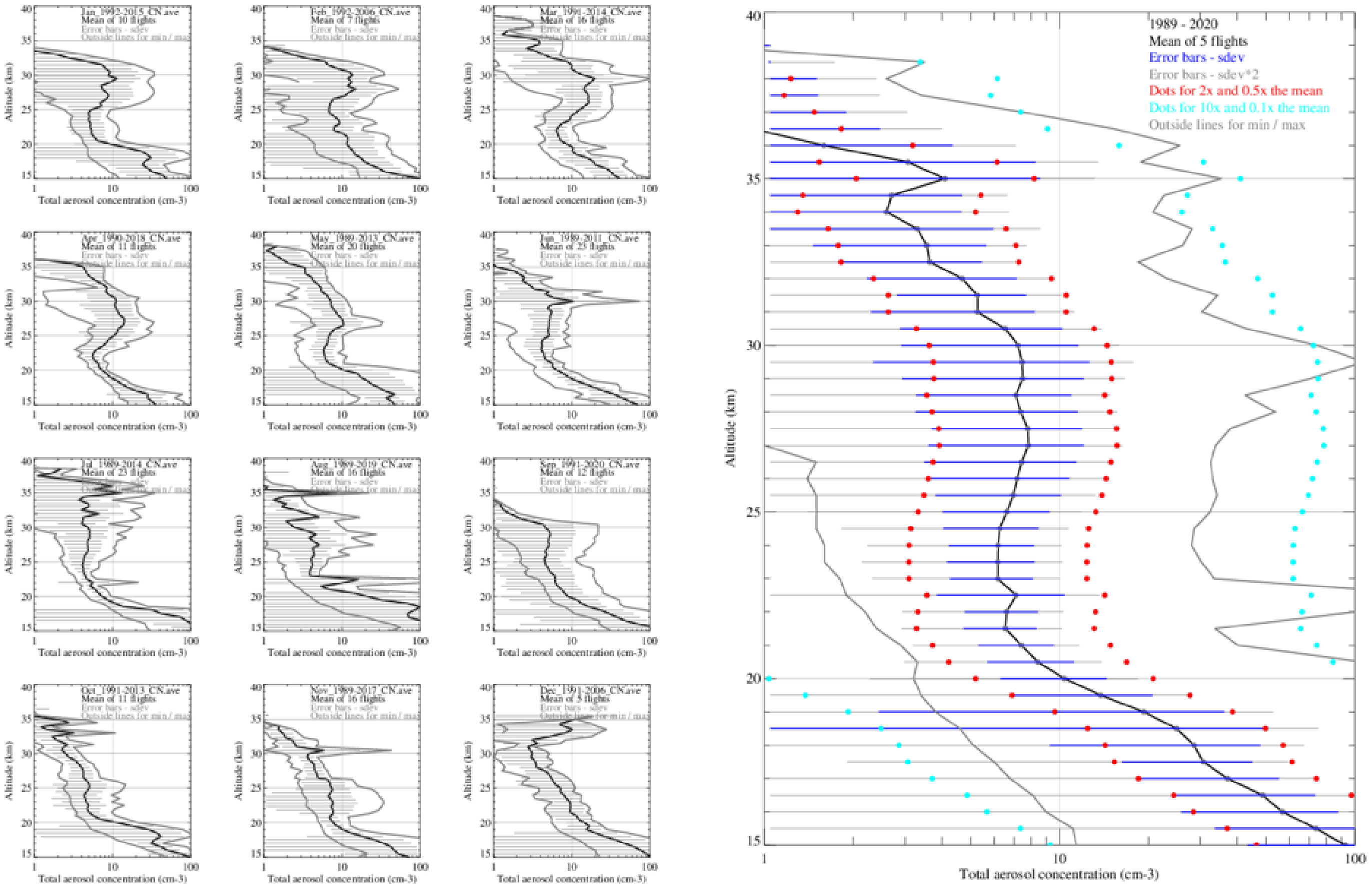
If the maximum distance was only 200 km, there will be ~ 1500 collocated SAGE II and SCIAMACHY profiles.
We have revised the respective sentence:
For SAGE II, the maximum distance from SCIAMACHY is increased to 500 km. This yields 4 255 collocations...

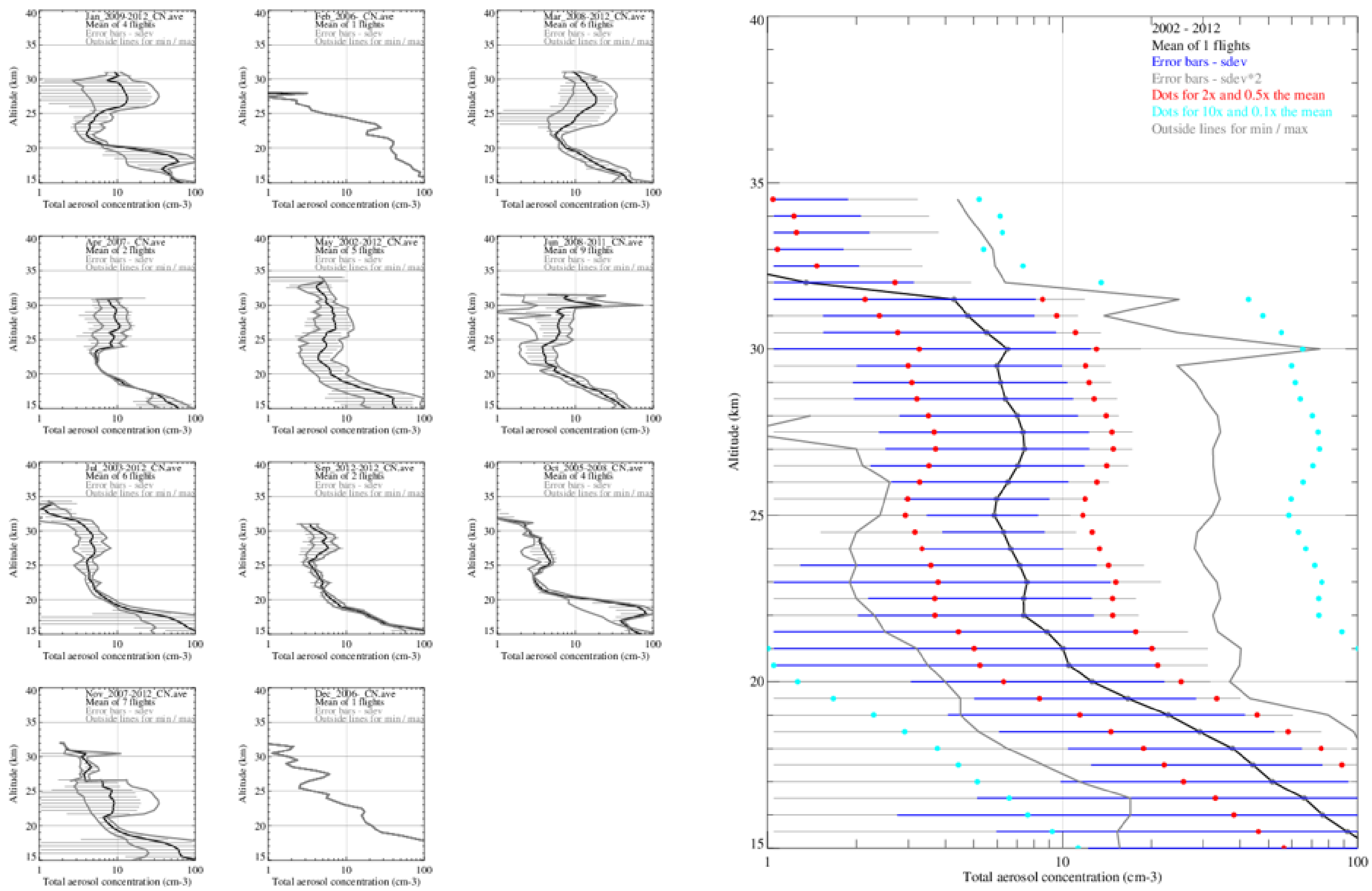
– page 23, Figure 6: This behavior is exactly what I expected based on Figure 2.
– page 24, line 585: “Discrepancies are smaller in the middle latitudes. . . ” Yes, this makes sense because these latitudes are the least impacted by volcanic activity. Therefore, your a priori N profile is more similar to the N profile of these latitudes.

We added the explanation:
The differences are smaller in the middle latitudes (~30-50°N/S) below 25 km altitude because the stratospheric aerosols in these altitudes and regions are impacted to a lesser extent by volcanic activity than in other latitudes.

– page 34, lines 774–775: “The retrieved median radii and geometric standard deviations should therefore be considered with caution in areas with high aerosol loading.” This is true for extinction as well.

We revised the sentence:
The retrieved and calculated aerosol characteristics should therefore be considered with caution in areas with high aerosol loading.


Jan_1992-2015_CN.ave
Mean of 10 flights
Feb_1992-2006_CN.ave
Mean of 7 flights
Mar_1991-2014_CN.ave
Mean of 16 flights
Apr_1990-2018_CN.ave
Mean of 11 flights
May_1989-2013_CN.ave
Mean of 20 flights
Jun_1989-2011_CN.ave
Mean of 25 flights
Jul_1989-2014_CN.ave
Mean of 23 flights
Aug_1989-2019_CN.ave
Mean of 16 flights
Sep_1991-2020_CN.ave
Mean of 12 flights
Oct_1991-2013_CN.ave
Mean of 11 flights
Nov_1989-2017_CN.ave
Mean of 16 flights
Dec_1991-2006_CN.ave
Mean of 5 flights
1989 - 2020
Mean of 5 flights
Error bars - sdev
Error bars - sdev*2
Dots for 2x and 0.5x the mean
Dots for 10x and 0.1x the mean
Outside lines for min / max
Altitude (km)
Total aerosol concentration (cm-3)


Jan_2009-2012_CN.ave
Mean of 4 flights
Feb_2006-_CN.ave
Mean of 1 flights
Mar_2008-2012_CN.ave
Mean of 6 flights
Apr_2007-_CN.ave
Mean of 2 flights
May_2002-2012_CN.ave
Mean of 5 flights
Jun_2008-2011_CN.ave
Mean of 9 flights
Jul_2003-2012_CN.ave
Mean of 6 flights
Sep_2012-2012_CN.ave
Mean of 2 flights
Oct_2005-2008_CN.ave
Mean of 4 flights
Nov_2007-2012_CN.ave
Mean of 7 flights
Dec_2006-_CN.ave
Mean of 1 flights
Error bars - sdev
Outside lines for min / max
Altitude (km)
Total aerosol concentration (cm-3)
2002 - 2012
Mean of 1 flights
Error bars - sdev
Error bars - sdev*2
Dots for 2x and 0.5x the mean
Dots for 10x and 0.1x the mean
Outside lines for min / max
Altitude (km)
Total aerosol concentration (cm-3)