We thank you for taking the time to review the manuscript and for your helpful comments. We have revised the manuscript in response to your comments. We believe that the manuscript has been greatly improved thanks to your suggestions.

AGeneral comment

This is an interesting paper which develops an optimal estimation aerosol retrieval using combined CALIOP and MODIS observations. The retrieval attempts to retrieve effective aerosol size and optical properties for each of four aerosol types. This goes far beyond the standard retrievals of CALIOP and MODIS, or any other satellite sensor, so it is good that comparisons with Aeronet retrievals are included to evaluate the performance of the retrieval. The algorithm is described well and the authors do a good job of examining some of the uncertainties (particle model assumptions) but more details on uncertainties (Section 4) and a few other topics would be helpful.

Specific comments

CALIOP Level 1B data is pre-processed by calculating running means using horizontal averaging over 10 km. Are retrievals performed only on 10-km averages that do not contain clouds, or are cloudy profiles removed before averaging to 10 km?

The running means were performed after the cloud layer data was removed by the VFM. We have revised Sect. 2.1.

The authors mention several times in Section 3 that the DVCs and DMRS are 'optimized to all CALIOP and MODIS measurements'. It is not clear to me what this means. Are the retrieved parameters adjusted to minimize the merit function for each MODIS-CALIOP data pair, or is there some sort of global optimization which is performed?

"each MODIS-CALIOP data pair" is correct. We corrected them in the manuscript.

What is the altitude range of the CALIOP-MODIS retrieval? It appears to be 0-10 km from Figure 8b.

The upper limit of the altitude is about 30 km. This is described in Sect. 2.1 of the revised manuscript.

I'm not sure how to read Table 3. What is meant by "relative value"? Relative to what? In the first row, is the mean difference between retrieved and simulated values an AOD of -0.15 or is it 15% of something?

The equations for the calculations of the differences were added to Tables 3 and 4, and the relative values were expressed as percent values in the revised manuscript.

Lines 284-285: How does it follow that the AOD of WS is greater than the AOD of LA because the SSA of LA is lower? The scattering extinction coefficient is reduced when SSA decreases, but the total extinction is unchanged. Please explain.

We added the following description in Sect. 3.1.3.
"In the AERONET product at worldwide locations, ω_0 ranges from 0.8 to 1.0 (Dubovik et al., 2002). ω_0 is about 0.96 for WS and about 0.44 for LA (Table 1), and ω_0 for an external mixture of WS and LA is calculated by $\omega_0 = (\tau_{a,WS}\omega_{0,WS} + \tau_{a,LA}\omega_{0,LA})/(\tau_{a,WS} + \tau_{a,LA})$. Thus, τ_a of WS must be greater than that of LA."

To understand the realism of the retrieval simulations more detail should be added on how the satellite Level 1 data was simulated in Section 4. From Line 256, it appears that a single number is used for random error in lidar backscatter (15%). But the relative random error of CALIOP attenuated backscatter profiles varies with altitude and with the albedo of the underlying surface and can be much worse than 15%, especially at higher altitudes. Was noise from the solar background simulated and added as a random variable to each sample in the vertical profile? Were retrieval errors due to systematic MODIS calibration errors or estimates of surface albedo considered?

The altitude dependency, surface contribution, and solar background noises should be included in the simulations. However, the information available from the published papers is limited, particularly, for the CALIPSO version 4 data set. It is difficult to simulate these realistic noises. The measurement accuracy of the attenuated backscatter coefficient at 532 nm in the version 3 product is evaluated by the comparison with the airborne HSRL. The mean difference is 2.9 % and the standard deviation is 20 % (Rogers et al., 2011). The bias of the version 4 product is smaller that the version 3 (Getzewich et al., 2018). Furthermore, we smoothed the data by the running mean, and the CALIOP-MODIS retrieval optimizes the state vector to only the data discriminated as aerosols by the VFM. Therefore, we think the random error of 15 % in the backscatter coefficient is an appropriate value.
The systematic errors of MODIS calibration and surface albedo were not considered. However, the random errors are important in this study. We use two bands of visible and near-infrared wavelengths. The volumes of each fine and coarse modes are estimated separately from the two bands. And the median radii of the fine and coarse modes depend on the spectral dependency of two bands. The random errors affect the spectral dependency of two bands, and make it difficult to retrieve the volumes and median radii of the fine and coarse modes.
We are now processing the observation data in the period from 2007 to 2021. We will conduct the further validation study using the ground-based remote sensing networks of AERONET, SKYNET, and AD-Net to estimate the practical uncertainties of the retrievals. We think the retrieval simulations in this study is

useful to interpret the retrieval errors in the validation study.

It is odd that retrieval uncertainties are larger over ocean than over land, while CALIOP and MODIS retrievals are both better over ocean than land. Is this really because SS is retrieved over ocean but not land, as the authors say, or could it be because AOD over ocean tends to be much smaller than over land? Or is it due to uncertainties in the optical model used for SS? It would be good to discuss reasons for this behavior in more detail. Marine aerosol is not just 'sea salt' and often contains internally mixed biogenic sulfate or biogenic organic compounds. This might impact the refractive index of the particle model used.

The simulated AOD is from 0.05 to 1.0 for both the land and ocean cases. The same optical model of SS is used in the simulations and retrievals. However, the particle radius of SS is different in the simulation and retrieval because the particle radius of SS is determined by the ocean surface wind speed and the random errors are added to the ocean surface wind speed in the simulations. The particle radius of SS is not optimized in the retrieval. Therefore, the difference of the particle radius of SS affects the AOD of SS. Since both WS and SS are less light-absorbing particles, the AOD of WS is overestimated (underestimated) when the AOD of SS is underestimated (overestimated). This different sign of the retrieved AOD can be seen in Table 3. We added these descriptions in Sect. 4.2.
WS would be estimated together with SS, if the biogenic sulfate or organic compounds have similar optical properties to WS. Actually, WS is estimated over the ocean, and the distribution of WS over the ocean is similar to the distribution of SS (Fig. 10a). WS over the ocean may be the biogenic sulfate and organic compounds, and fine particle of sea salt.

The authors comment that extinction coefficients are unnaturally large at 70N and 70S-80S. I do not see evidence of this in figure 8a or 8b and am wondering what the authors are referring to. There appear to be very few retrievals at 70S-80S. The text says the large EC are due to cloud contamination, but could it be due to ice cover and thus high surface albedo? Are retrievals attempted over ice or only over ice-free ocean?

The color bar of Fig. 8 was modified to emphasize the spatial variations. The slightly large ECs are seen at the altitudes from 0 to 9 km and latitudes from 70S to 80S, and are indicated by blue color in Fig. 8b of the revised manuscript. A peak of EC is seen at altitudes from 0 to 1 km and latitudes at 70N, and are indicated by the colors from green to red. The retrievals have been attempted to the observations over the ice surface. As you say, it is possible that the high surface albedo of ice has affected the retrieval of the ECs. We mentioned the influences of high surface albedo of ice in the revised manuscript.

Minor comments

It was not clear to me what is meant by 'dry volume concentration' (line 125). What are the units?

The unit is um^3/um^3. We added the following explanation in Sect. 3.1.1.
"V_{dry} is defined as the volume of aerosols at a relative humidity of 0 % per unit atmospheric volume, and $r_{m,dry}$ is defined as the median radius of aerosols at a relative humidity of 0 %."

Equations 10 and 13 explain constraints applied to the solution, using somewhat different approaches to notation. I find the approach used in Eqn 13 to be more clear than Eqn 10.

Eqs. 10 and 11 are changed to the same approach as Eq. 13.

Lines 422-423: rather than "SSA of the land ..", I think "SSA over land .." is meant, and the same for "of the ocean" and for AF

We corrected them.

The authors introduce a large number of non-standard 2- and 3-letter abbreviations for various parameters (ABC, LR, DMR, …) , and then later introduce mathematical symbols for some of these parameters when used in equations. It would be simpler to define the math symbols and use them throughout the paper. I found DMR and DVC especially awkward and had a much easier time reading re and vdry .

We changed the abbreviations to the mathematical symbols.

Depolarization ratio (DR) and linear depolarization ratio (LDR) are both used. Aren't these the same parameter?

We have used DR for the CALIOP measurements, and LDR for particle optical property, but these are confusing to readers. In the revised manuscript, the total depolarization ratio (δ) was used for the CALIOP measurements, and the particle depolarization ratio (δ_a) was used for the particle optical property.