Reply to General comments:

1. The paper by Fu and Hasekamp provides very useful insight in the different performance of retrievals applying a parameterization of the 2-mode (fine and coarse) size distribution on the one hand (parametric mode), and retrievals utilizing more than two modes of size distribution parameters (multi-mode retrievals). I can recommend the paper for publication because it addresses a fundamental choice made by a lot of aerosol retrievals, which may have significant consequences not only for the accuracy of the retrievals (as the authors show) but potentially also for its computational cost. The paper may therefore provide some guidance for selecting an appropriate retrieval method in the context of accuracy needs, computational performance and the available information content in the measurement. Fu and Hasekamp are basing their results of the evaluation of the performance of the both mode-fitting approaches on the same SRON based RT and inversion algorithm using both synthetic and PARASOL measurement.

The latter are then evaluated with respect to Aeronet data. While the presented approach is solid in terms of performance differences for the SRON Algorithm, what is missing is a discussion to which extend the results are significant even for other full inversion approaches and RT models. Since there are currently not many which can perform at a global scale (as the authors state themselves) it would have been interesting to understand if these few algorithms would converge in their performance when using the same type of mode fitting. While I would guess that the authors ultimately had such a comparison in mind, a discussion of the results in this context - and maybe providing some outlook on how to apply this kind of sensitivity test also in the context of other retrieval schemes - is currently missing.

Response:

The performance of a retrieval algorithm depends on a number of things in addition to the state vector definition for aerosols (studied here): The inversion approach (cost function, regularization strength, multi- versus single pixel), the accuracy of the forward model, the surface reflection model. Now the SRON algorithm is extended to an arbitrary number of 'fixed' modes, it can be more easily compared to e.g. the GRASP algorithm which is also based on fixed modes. For such comparisons, it is important to study, where possible, the algorithm differences in a systematic manner. This would include using the same surface reflection model and ensuring that the RT models in both algorithms agree with benchmark results.

We added the following paragraph to the 'Discussion and Conclusion' section:

"When comparing retrievals between different algorithms, it is important to realize that the performance of a given algorithm depends on a number of factors, the definition of the aerosol state vector being one of them. Other factors are the inversion approach (cost function, regularization strength, multiple versus single pixel), the accuracy of the forward model, and the surface reflection model. It is important to study the above mentioned aspects with an individual algorithm. However, now that the SRON algorithm has been extended to include an arbitrary number of fixed modes, it has become easier to compare to other algorithms using a similar state vector definition (Dubovik et al., 2011; Xu et al., 2017). This would be an important topic for future research. "

Reply to specific comments:

1. While one would expect that retrievals over water surfaces would reduce the parameter space and therefore may make the evaluation of the performance difference for different mode fitting scheme more robust, the study focuses on land surfaces only without further qualification. In this context I am also missing a discussion of the combination of wave-length, surface properties and scattering geometry, on the synthetic results, since some of the combinations may not be realistic and may therefore complicate the interpretation of the comparison of the synthetic results to the performance using PARASOL measurements. Ideally the synthetic retrievals could limited to the observation geometries and surface properties combination available at the Aeronet stations. This greatly would improve the interpretation of the PARASOL retrieval results in the context of the synthetic retrievals.

Response:

- To make explicitly clear that the paper relates to aerosol retrievals over land, we changed the title to:

" Retrieval of aerosol micro-physical and optical properties over land using a multi-mode approach ".

-We believe that the surface properties in our synthetic measurements are representative for those over AERONET sites. However, as noted by the reviewer, the viewing geometries by which a certain ground pixel is seen by POLDER-3 varies drastically over season and over the orbit/swath. The geometry used in the paper is one of the 'best' geometries as it assumes measurements in the principal plane. To investigate the effect of this assumption we also performed the same study for a more 'difficult' geometry with high sun and a relative azimuth angle of 60/-120 degrees, i.e. far away from the principal plane. The geometry used in the paper (GEOM-1) and the more 'difficult' one (GEOM-2) are summarized in Table 1 of this response. Figures 17-23 show the result of the synthetic study for GEOM-2 for AOT, SSA, real refractive index (fine and coarse), imaginary refractive index (fine and coarse), and aerosol layer height. These results can be compared to the corresponding results for GEOM-1 in the paper. Although the overall errors are larger for GEOM-2 on most parameters, the general conclusions of the synthetic study still hold for GEOM-2. Therefore, we are convinced that they are representative. We added the following phrase to the paper in Section 'Discussions and conclusions':

" It should be noted that the geometry used for the synthetic study in this paper is quite favorable as it assumes measurements in the principal plane. We also did the same synthetic study for a much less favorable geometry (SZA=20°, relative azimuth angle= $60^{\circ}/-120^{\circ}$). Although for the latter geometry, the performance is somewhat worse, the main conclusions from the synthetic study still hold for this geometry. "

_	GEOM-1	GEOM-2
Solar Zenith Angle (SZA) (degree)	41.0	20.0
Satellite Viewing Zenith Angle (VZA) (degree)	$60.0, 50.0, 40.0, \\ 30.0, 20.0, 10.0,$	
	0.0, -10.0, -20.0, -30.0,	
	-40.0, -50.0, -60.0	
Satellite Azimuth Angle	180.0 (if VZA $<$ 0)	$120.0 ~({ m if}~{ m VZA}{<}0)$
(SAA) (degree)	$0.0 ~(\mathrm{if}~\mathrm{VZA}{>}0)$	$60.0 ~({ m if}~{ m VZA}{>}0)$
Scatter Angle (ScatA) (degree)	79.0, 89.0, 99.0,	108.8, 118.2, 127.6,
	109.0, 119.0, 129.0,	136.7, 145.5, 153.6,
	139.0,	160.0,
	149.0,159.0,169.0,	162.8, 160.3, 154.1,
	179.0, 171.0, 161.0	146.1,137.3,128.2

Table 1: Different observation geometries in synthetic retrievals. (GEOM-1 corresponds to synthetic retrievals in the paper; GEOM-2 is for a test of effect of different geometry.)



Figure 17: Synthetic retrievals using GEOM-2 for AOT.



Figure 18: Synthetic retrievals using GEOM-2 for SSA.



Figure 19: Synthetic retrievals using GEOM-2 for the real part of refractive index (at 550 nm) of the fine modes (m_r^f) .



Figure 20: Synthetic retrievals using GEOM-2 for the real part of refractive index (at 550 nm) of the coarse modes (m_r^c) .



Figure 21: Synthetic retrievals using GEOM-2 for the imaginary part of refractive index (at 550 nm) of the fine modes (m_i^f) .



Figure 22: Synthetic retrievals using GEOM-2 for the imaginary part of refractive index (at 550 nm) of the coarse modes (m_i^c) .



Figure 23: Synthetic retrievals using GEOM-2 for the central height (z) of the aerosol layer.

2. The difference of the retrieval performance between consistent and inconsistent synthetic retrievals are potentially very interesting to understand the frequent problems when using, for example, generic pre-launch TOA test-data sets for end-to-end system performance studies and developments. However a more detailed interpretation or analysis of the results appears to be missing in the paper. The results for AOT at least seem to indicate that parametric 2 mode retrievals perform better in inconsistent cases than multi-mode retrievals. Can this be understood or explained?

Response:

Although the 2-mode parametric retrieval indeed performs somewhat better for most parameters on the 10 mode synthetic measurements than vice versa, overall we believe that the performance of different retrievals on inconsistent synthetic measurements is sufficiently good compared to the level-2 requirements (in the revised version a requirement table has been added following the comment of reviewer 1). The only exception is the fine mode refractive index for which the poor performance of the 10 mode retrieval on the 2 mode synthetic measurement is not understood (as mentioned in the paper).

3. For the PARASOL retrievals in section 5 it is stated that multi-mode retrivals with more than 4 modes perform well (while at the same time mode-5 seems to have the largest bias), whereas the conclusion from the synthetic retrievals was that multi-mode retrievals perform well for n>5. Could there be a reason for this (although small) discrepancy. Overall the results are presented as if mode-5 is a kind of physical significant lower limit for multi-mode retrievals (if yes, why?), while the results seem to more indicate a general trend for decreasing RMS with higher mode numbers.

Response:

You are right. Only for synthetic retrievals we see for most parameters a decrease in error for increasing number of modes till 5 modes and after that only a small decrease. For the real retrievals indeed this conclusion does not hold and we have removed statements in the paper that suggest this.

4. In Section 3.1 there is a reference missing to the actual PARASOL data and its version used. Are there references available for the expected intensity and polarisation error for PARASOL?

Response:

We added a statement that we use level-1 Collection 3 data, in Sect. 3.1, " The PARASOL level-1 Collection 3 product data have been used in this study."

Information about the measurement errors of POLDER-3 is hard to find. The results of Fougnie et al. (2007) seem to indicate (although not explicitly claimed) a radiometric error of about 2% for all bands except 443 nm where it would be about 5%. For the DoLP error no explicit reference exists. Knobelspiesse et al. (2012) assume 0.02 in their sensitivity study while Dubovik et al. (2011) use 0.005 but add the statement that the error is likely to be a factor 2-3 larger. Given that the exact magnitude of the POLDER-3 errors is of less importance for our study, we decided not to speculate on the magnitude of the errors.

References

- Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., Ducos, F., Sinyuk, A., and Lopatin, A.: Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations, Atmospheric Measurement Techniques, 4, 975–1018, https://doi.org/10.5194/amt-4-975-2011, 2011.
- Fougnie, B., Bracco, G., Lafrance, B., Ruffel, C., Hagolle, O., and Tinel, C.: PARASOL inflight calibration and performance, Appl. Opt., 46, 5435–5451, https://doi.org/10.1364/ao.46. 005435, 2007.
- Knobelspiesse, K., Cairns, B., Mishchenko, M., Chowdhary, J., Tsigaridis, K., van Diedenhoven, B., Martin, W., Ottaviani, M., and Alexandrov, M.: Analysis of fine-mode aerosol retrieval capabilities by different passive remote sensing instrument designs., Optics express, 20, 21457– 21484, 2012.
- Xu, F., van Harten, G., Diner, D. J., Kalashnikova, O. V., Seidel, F. C., Bruegge, C. J., and Dubovik, O.: Coupled retrieval of aerosol properties and land surface reflection using the Airborne Multiangle SpectroPolarimetric Imager, Journal of Geophysical Research (Atmospheres), 122, 7004–7026, https://doi.org/10.1002/2017jd026776, 2017.