Reply to comments:

1. *General comment. I understand that one of the purposes of this work is to determine expected uncertainty on the retrievals for the polarimeters. But AERONET and HSRL should already have documented expected uncertainty. It certainly would be helpful to indicate on the figures what is the expected uncertainty of the known sensors. We see values for MAE, bias, etc., but do not know how to put these values into context. If we knew AERONET uncertainty for that parameter, for example, context could be established.*

   Response:

   We agree. We added a phrase to the paper with respect to the expected AERONET AOD uncertainty in Sect. 3.5:

   " The uncertainty on AERONET AOD is 0.01 for mid-visible wavelengths and 0.03 for UV wavelengths (Eck et al., 1999) and is dominated by a calibration (systematic) error. "

   For the expected uncertainty in HSRL-2, we refer to the comparison between HSRL-2 and AERONET. We added the phrase in Sect. 4.2.1: " The bias between HSRL-2 and AERONET is within the AERONET uncertainty. The random differences, with standard deviation 0.029 at 380 nm and 0.014 at 532 nm are most likely due to HSRL-2 uncertainties. "

2. *General comment. This is a corollary to (1). AERONET AOD has very small uncertainty, but AERONET retrieved products and these include the SDA products have larger error bars. The goal in comparing polarimeter retrievals to these other retrievals is comparison, not validation. This was not explicitly stated in the paper.*

   Response:

   We agree and explicitly state this to the paper in Sect. 3.5:

   " The multispectral aerosol optical depth (AOD) from the MAP and lidar retrievals is validated with AERONET (AErosol RObotic NETwork) level 1.5 data (Holben et al., 2001) (version 3.0). The data are cloud cleared. The uncertainty on AERONET AOD is 0.01 for mid-visible wavelengths and 0.03 for UV wavelengths (Eck et al., 1999) and is dominated by a calibration (systematic) error. The effective radius for fine and coarse modes are compared with AERONET level 1.5 Almucantar Retrieval Inversion Products (Dubovik et al., 2002). The AOD of fine and coarse modes are compared with AERONET level 1.5 spectral de-convolution algorithm (SDA) data (O'Neill et al., 2003). It should be noted that the inversion- and SDA products are quite uncertain themselves at low AOD so the comparison to these products should not be considered a validation. "

3. *General comment. I see in the description of the different data sets mitigating strategies for inhomogeneity for registering the different angular views. Does this include topographical variation?*

   Response:
   Yes, this does include topographical variation.

4. *Page 9. Last paragraph that begins with "As with the extinction products", I'm a little unclear on what is being said here. "HSRL method" is when HSRL measures extinction. "assumed lidar ratio" is when it does not. The HSRL method is not available in many situations during ACEPOL, so the lidar information is going to come to us like an old-fashioned backscattering lidar with an assumed lidar ratio. It's not clear why the HSRL method is going to be unavailable. Then here it seems to imply that there is going to be a choice between the two methods, not that the HSRL method is unavailable, but that both are available. And then it says that the assumed lidar ratio method is actually BETTER than the HSRL method at low loading. This is because one measures its uncertainty in a relative sense and the other in an absolute sense. The fact that the assumed lidar ratio can be better than the HSRL method is very strange to me. Did I understand this paragraph correctly?*

   Response:
   Thanks for pointing out this point. Actually the statement "For ACEPOL, the extinction, AOD, and lidar ratio from the HSRL methodology are not available for many ground pixels." was not accurate (which was based on the old version HSRL-2 data). We have removed this statement and re-wrote that part to avoid confusion in Sect. 3.4: " For ACEPOL, the extinction products from the HSRL method are reported at 150 m vertical resolution and at temporal resolution of 60 s generally and 10 s. Additionally, the aerosol extinction products at 355 nm and 532 nm are also provided based on the aerosol backscatter and an assumed lidar ratio of 40 sr, and reported at the backscatter resolution.

   Similarly, the AOD is reported from the standard HSRL approach and also the AOD calculated using assumed lidar ratio is provided. "

   Yes, we have two AOD products as two choices. For the low AOD case the AOD from the assumed lidar ratio is better than the HSRL method. This is because for low AOD, both approaches are difficult, but the AOD from the assumed lidar ratio is expected to be smaller (given that the AOD is small) than the systematic uncertainty $\sim$0.05 from the HSRL method for ACEPOL.

5. *Section 3.5. AERONET section. Level 1.5 is cloud cleared, but not quality controlled. Also be aware that fine and coarse as defined by both the almucantar retrievals and the*

SDA methods are going to be different than defining fine and coarse by specific modes as is done in the polarimeter retrieval (Table 1). This may introduce differences in your comparisons. It did with the MODIS Dark Target over ocean retrieval.

Response:

Thanks. We added this point to the paper in Sect. 4.1:

" Also, it should be noted that the "fine" and "coarse" as defined by the Almucantar retrievals are different with defining "fine" and "coarse" by specific modes as shown in Table 1. This may introduce differences in the comparisons. "

6. *Page 11-12. Discussion of comparison of effective radius against AERONET. Perhaps AERONET is wrong here? This is retrieval vs. retrieval, not retrieval vs. truth. And the loading is extremely low. I would think that everybody is running on fumes here. This applies to fine mode, but especially to coarse mode. Nobody has SWIR to really nail coarse mode. And AERONET's definition of fine and coarse modes, and their respective effective radii, are defined differently than the five modes in Table 1.*

Response:

We agree and included this aspect to the paper in Sect. 4.1:

" It is important to note that for the low AOD values encountered during ACEPOL, the AERONET retrieved fine and coarse mode AOD and effective radius are very uncertain themselves. Therefore, this comparison should not be interpreted as "retrieval versus truth" but rather as "retrieval versus retrieval". "

7. *Figure 3. If I'm interpreting these plots correctly... The MAP retrievals can be very different from AERONET. For example, RSP has differences of -0.04 where the (AERONET + RSP)/2 = 0.025. This means that RSP retrieved tau_c of 0.005 and AERONET 0.045. In absolute terms that's not a lot, but in terms of relative contributions of the coarse mode to the total AOD it is a lot. Is it within expected error of the AERONET retrieval? It would be very helpful to have some context for the magnitude of the differences.*

Response:

We added some context for the magnitude to the paper in Sect. 4.1:

" The comparison shows a MAE of 0.028, 0.029, and 0.012 for SPEX, RSP, and airMSPI, respectively for τ^{f} and 0.026, 0.028, 0.017 for τ^{c}. The bias is 0.028, 0.019 and 0.004 for τ^{f} and 0.025, 0.028, and 0.003 for τ^{c}. So, SPEX and RSP have an overestimation of the fine mode and an underestimation of the coarse mode, compared to AERONET SDA product. Although these biases are large in a relative sense (given the low AOD, especially for the coarse mode), they are within the expected error from the AERONET SDA product. AirMSPI compares better to the AERONET SDA product than SPEX airborne and RSP. ”

8. *Section 4.2.1 These comparisons are all with "assumed lidar ratio". Are these the only days with collocations? If there is a choice between assumed lidar ratio and HSRL method, how does the HSRL method compare?*

   Response:
   Indeed all HSRL2-AERONET collocations are included. We have explicitly mentioned in Sect 3.4, the comparisons in Figure 4 are with the assumed lidar ratio. For these low AOD cases, the comparison against AERONET for HSRL AOD from the HSRL method is worse than from the assumed lidar ratio. The reason has been explained in Sect 3.4.

9. *Section 4.2.2. I grew up in Los Angeles and the Central Valley, so I know this territory well, but not everybody does. Maybe use "east" and "west" without place names, or annotate the image.*

   Response:
   Thanks. We changed them to the paper in Sect. 4.2.2:
   “ From this figure it follows that there were very low AOD values for the eastern part of the scene and somewhat higher values in the western and south-western part of the scene. ”

10. *Final sentence of Section 4.2.2. "The differences from the direct comparison between SPEX and RSP are somewhat larger than those from individual comparisons with HSRL-2 of SPEX and RSP, respectively. This suggests that the differences with HSRL-2 are not caused by common assumptions in the SPEX and RSP retrievals, but are rather caused by errors that are specific to each MAP". I don't follow the logic.*

    Response:
    If the differences with HSRL-2 are caused by common assumptions in the SPEX and RSP retrievals, the differences should be smaller when comparing SPEX and RSP (than comparing MAPs with HSRL-2) because the common assumptions should have little effect when comparing SPEX and RSP. However, the comparison between SPEX and RSP is worse than comparing MAPs with HSRL-2. Thus, we reach the statement “ This suggests that the differences with HSRL-2 are not caused by common assumptions in the SPEX and RSP retrievals, but are rather caused by errors that are specific to each MAP”

11. *Section 4.2.3. page 13. Lines 15-17. "It should be noted that the smoke plume exhibits large spatial variation so part of the MAP-lidar differences can be attributed to the fact that different instruments see a slightly different part of the smoke plume". What about different angles from the same instrument seeing different parts of the smoke, or what if the smoke changes between the fore and aft angles are measured? What happens to the retrieval? It would be really nice to have a quantitative sense of how variable that plume is. Could we see a spatial plot of the smoke retrievals or at least have stdev on the parameters shown in Table 2.*

    Response:
    If the different viewing angles would see different parts of the smoke plume, this would result in a large χ^2 of differences between forward model and measurements. This is not the case for the points in the Figure 7.

    To illustrate the spatial variability of the smoke plume, we have included Figure 7a, which gives a sense of how variable the AOD of smoke plume is. We have also included the standard deviation inside the plume from SPEX and RSP on the parameters shown in the Table 3.

12. *Page 13. Lines 25-26. "Our explanation for this, is that at high AOD the measured radiance and DoLP are less affected by the co-registration errors between viewing angles than for low AOD". How could this be? The evolving, hetereogeneous smoke plume has to be more difficult to co-register between angles than the unmoving ground.*

    Response:
    The land surface can be very patchy, especially near Fresno and Bakersfield. This leads to higher spatial variability in the radiance than the spatial variation of the smoke plume at 100 meter scale. The difference in co-location between the MAPs and HSRL-2 sampling may however be 1 km.

13. *Page 13. Lines 33-34. On the other hand, I think this is a really good explanation: "A possible explanation for the difference could be the simplified description of non- spherical particles in our retrieval approach. "*

    Response:
    Thanks.

14. *Figure 7d-f. Are lidar ratios here retrieved via HSRL method, or assumed? If assumed,*

does these figures make any sense. If retrieved, then why not use retrieved throughout the paper? Or show that they are worse than assumed. This whole retrieved vs. assumed lidar ratio choice never sat well with me throughout the manuscript.

Response:
Shown in the plots are those retrieved from the HSRL method. This is consistent with our comparison for the smoke case for high AOD from the HSRL method. The only place we use the assumed lidar ratio is for the comparison of AOD in the low AOD case. For these cases we do not compare the lidar ratio because indeed that would not make sense.

15. *Table 2. Maybe show stdev along with mean? Or show spatial distribution if any of these properties are varying downwind?*

Response:
We have done so in the revised manuscript. Please see Figure 7a and Table 3.

16. *Page 14. Line 13. "the latter value is closer to the ALH derived from HSRL-2 (2.64 km)". Sure slightly closer, but still 1 km off. Not that much different from SPEX.*

Response:
We agree. We re-wrote the phrase to the paper in Sect. 4.2.4:
" For the Aerosol Layer Height (ALH), SPEX retrieves a higher value (4.417 km) than RSP (1.148 km), where the latter value is somewhat closer to the ALH derived from HSRL-2 (2.64 km). "

17. *Page 14. Line 14. The explanation of ALH being difficult to retrieve without UV might be elaborated on a little here.*

Response:
We added a phrase to the paper in Sect. 4.2.4:
" Here, it should be noted that for SPEX the shortest wavelength that is used in the retrieval is 450 nm, so we do not expect an accurate ALH retrieval because the retrieval of ALH from polarization requires a strong signal from Rayleigh scattering (Wu et al., 2016). "

18. *Finally. . . don't you want to state a conclusion? What is the overarching thing you have learned? If this were my paper I would conclude that the 3 polarimeters are producing*

comparable results when forced through the same algorithm. The exception being aerosol layer height and perhaps some coarse mode parameters, which suffer from not having the bands that these parameters are sensitive to: shortwave (410 nm) and SWIR, respectively. So when there is no sensitivity, the retrieval becomes a random number generator. But for parameters that the instruments are sensitive to, there is little difference between instruments. It is still TBD whether algorithmic differences are going to matter. But it is not my paper. The authors can choose to write a conclusion of their choice. Or not.

Response:

Thanks. We included these summaries to the paper in the conclusion part:

" In this study, 3 polarimeters produced comparable results when using the same algorithm. The exception were the ALH and some coarse mode parameters, which were mainly caused by not having the bands that these parameters were sensitive to: shortwave (410 nm) and SWIR, respectively. For parameters that the instruments were sensitive to, good agreements were found among instruments. Our results corroborate the findings of earlier studies that different combinations of spectral and angular measurements yield a very similar retrieval capability for aerosol properties (Hasekamp and Landgraf, 2007; Wu et al., 2015; Hasekamp et al., 2019) "

References

Dubovik, O., Holben, B. N., Lapyonok, T., Sinyuk, A., Mishchenko, M. I., Yang, P., and Slutsker, I.: Non-spherical aerosol retrieval method employing light scattering by spheroids, Geophys. Res. Lett., 29, 1415+, https://doi.org/10.1029/2001gl014506, 2002.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, Journal of Geophysical Research: Atmospheres, 104, 31 333–31 349, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JD900923`, 1999.

Hasekamp, O. P. and Landgraf, J.: Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements, Appl. Opt., 46, 3332–3344, https://doi.org/10.1364/ao.46.003332, 2007.

Hasekamp, O. P., Fu, G., Rusli, S. P., Wu, L., Di Noia, A., Brugh, J. a. d., Landgraf, J., Martijn Smit, J., Rietjens, J., and van Amerongen, A.: Aerosol measurements by SPEX-one on the NASA PACE mission: expected retrieval capabilities, Journal of Quantitative Spectroscopy and Radiative Transfer, 227, 170–184, URL `http://www.sciencedirect.com/science/article/pii/S0022407318308653`, 2019.

Holben, B. N., TanrÃľ, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. V., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G.: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, J. Geophys. Res., 106, 12–12 097, https://doi.org/10.1029/2001jd900014, 2001.

O'Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N., and Thulasiraman, S.: Spectral discrimination of coarse and fine mode optical depth, Journal of Geophysical Research: Atmospheres, 108, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002975`, 2003.

Wu, L., Hasekamp, O., van Diedenhoven, B., and Cairns, B.: Aerosol retrieval from multiangle, multispectral photopolarimetric measurements: importance of spectral range and angular resolution, Atmospheric Measurement Techniques, 8, 2625–2638, https://doi.org/10.5194/amt-8-2625-2015, 2015.

Wu, L., Hasekamp, O., van Diedenhoven, B., Cairns, B., Yorks, J. E., and Chowdhary, J.: Passive remote sensing of aerosol layer height using near-UV multiangle polarization measurements, Geophys. Res. Lett., 43, 8783–8790, https://doi.org/10.1002/2016gl069848, 2016.