We are grateful to reviewers for the efforts they spent on reviewing the paper and providing comments, which allowed considerably improve the paper.

Our detailed responses to the remained questions and comments are provided below.

Reviewer1

I believe the authors have improved this manuscript with their extensive revisions. The language was substantially easier to read, with very few grammatical errors remaining, and the scientific reasoning is explained in more detail. I was largely satisfied with the original submission, but agree with some of the criticisms made by the other reviewers. The authors haven't completely addressed all of those concerns, but I think the issues that remain are a matter of professional disagreement rather than scientific failure. Therefore, I think the paper is suitable for publication after some further minor revision.

I am sympathetic to many of the points made by Reviewer 3. Reviewer 4 makes similar points but their tone detracts from the critique. For my thoughts on their comments,

- The paper is closer to a technical report than a paper than is ideal, but that style is acceptable for AMT.
- I agree that there are far too many very similar plots. The paper would be easier to read if only the most important plots were presented in the main body of the paper to tell a concise and coherent story, with the remainder placed in appendices or supplementary material.
  - o A greater weakness is the repetition of plot style rather their number. The Tukey plot included in the response to Reviewer 3 would be an excellent replacement for one of the scatterplot sets, as the use of various styles of presentation makes data more accessible and comprehensible. The authors already improved this through the addition of data tables, but could do better if they wanted.
  - o However, this paper is far from the most plot-dense I have reviewed for this journal and most of the plots are needed to address the concerns of the other reviewers. As such, I think the author's decisions are justifiable. They could present a better paper by streamlining the figures, but I don't think they are obliged to do so.
- I agree that the use of log AOD would be a better axis for the validation plots but will not require it because it remains an area of robust disagreement in the aerosol community.
- I agree that data assimilation would be an interesting approach to this particular problem.
  - o However, I do not think it is appropriate to reject a paper because we would have done the work differently. SYREMIS has been under development for around a decade, and there are sound reasons to do a retrieval rather than assimilation. Though some material from the second part of the response to Reviewer 3 is included within the revised manuscript, I think this need to be revised to more rigorously defend the strengths and weaknesses of doing a retrieval (and mention that alternative approaches are available).
- I also find the discussion of measurement uncertainty disappointing. While I agree with the papers offered by Reviewer 3, I think it is appropriate that the manuscript does not cite them as the authors disagree with those paper's terminology (as discussed below).

Response:

Many thanks to Reviewer 1, who also provided feedback on the comments from other reviewers.

We would like to provide here some comments about the number of figures and their representation. The number of figures corresponds to the number of inter-comparison tests performed to evaluate the SYREMIS/GRASP approach. These tests reflect the advantages of the SYREMIS/GRASP approach in comparison to the single instrument retrieval, especially taking into account that we wanted to demonstrate the synergetic retrieval possibility not for one retrieval parameter but for the extended set of them, including their spectral dependence. Reducing this number would require providing more supplementary material, for example, tables. We agree that the figures or tables can be presented differently, depending on the preference. Nevertheless, this does not change the conclusions about the advantages of the SYREMIS/GRASP approach we observe.

Finally some new thoughts I had while reading,
- I dislike the use of "standard deviation" in the revised manuscript as it does not distinguish between three distinct concepts: natural variability in the observed field, measurement variability due to unavoidable measurement errors, and uncertainty resulting from the retrieval method and model.

Response:
We appreciate the nice intent of the review to make sure that all technical details are made clear within our article. However, we would like to emphasize that our article doesn't have the objective to describe methodology and all details of used numerical inversion concepts. The paper, as clearly stated in the text, focused on the application of an earlier developed inversion methodology to a practical application. Indeed, the GRASP inversion concept is described in detail in the paper Dubovik et al. (2021). That paper discusses all definitions and provides detailed comparative discussions with other known, and especially mainstream inversion concepts. At the same time, we would like to note that in order to make the paper sufficiently complete and clear for the reader, we did include reasonably extensive summaries of the key aspects of the concept. However, it is expected that the basic principles, as well as more technical details of the used inversion concept, the reader should find in the earlier methodological papers.

The reviewer asks many specific questions. The answers to all these questions should be sought in Dubovik et al.(2021) and earlier papers referenced in that paper. At the same time, for the convenience of the reviewer, we will provide explicit answers to specific questions.

o Despite reading it twice, I am unclear on if "standard deviation" means the variance in the observations that go into the state vector, the variance all observations from that instrument, or some theoretical calculation. For example, line 258 states "standard deviation of uncertainties in each dataset". For the strict definition of uncertainty, "a parameter that describes the range of values that can be ascribed to a measurement", you are describing the spread of the spreads. But 379 mentions "standard deviation of the measurement fitting", which sounds like the variance of the retrieval residual, which is more a measure of retrieval performance than uncertainty.

Response to above specific comment:
We believe that here the reviewer confuses some definitions. The fact is that according the definition, the "variance" is equal to the square of the "standard deviation". Therefore, if the values of standard deviations are determined, the variance values are determined too.

Regarding the phrase "the standard deviation of the measurement fitting, which defines weighting matrices", we agree that it is somewhat confusing and we corrected it in the revised version by
"defining standard deviation of the expected uncertainties, which determine weighting matrices".

o Though you indicate the details of how these variances are calculated can be found in old Oleg papers, it would be useful to have the details here. Particularly distinguishing between natural variability (which is covered by your "a priori" components), measurement noise (which is what at I think you want standard deviation to cover), and forward model uncertainty (which appears to be undiscussed)
o I wouldn't require the authors to change this, but I ask that they reconsider to reduce confusion. Aerosol scientists are already very bad at statistical terminology.

Response to the above specific comment:

Here we would like to emphasize that in the GRASP (Multi-Term LSM) concept, all uncertainties in input data sets (both measurements and a priori) are described in the same way: i.e. input data presented as normally or log-normally distributed vectors described by the means and corresponding covariance matrices. It is assumed that the forward model of either the measurements or a priori data is perfect. Certainly, this is an approximation, but this is a fundamentally common approximation used everywhere in all mainstream inversion approaches, including Maximum Likelihood and Least Squares Methods. This aspect is clearly stated and discussed by Dubovik et al. (2021).

- I wish to register my disagreement with the use of " a priori" to describe gradients. In atmospheric science, the widespread adoption of " a priori" is in the sense of an optimal estimation retrieval. typically describing observations from a different instrument. While that can include spatio-temporal correlations, the methods here more closely resemble the use of Hessians to smooth solutions, a central feature in regularization, e.g. doi: 10.1364/ao.41.003685 or 10.1364/ao.41.003685, which is a competing retrieval method. Describing equation 3A (and downstream concepts) as "smoothing" or "variability" constraints/knowledge would avoid confusing readers and more clearly distinguish this work from optimal estimation (which I believe these authors wish to do as they make a reasonable argument as to the weaknesses of that method).

Response to the above specific comment:

Here we would like to stress again that the numerical inversion of GRASP is based on the "Multi-Term LSM" approach used in detail by Dubovik et al. (2021). The key idea of this approach is the use of multiple a priori constraints. In this respect, Dubovik et al. (2021) claim that the "Multi-Term LSM" generalizes the mainstream approaches, including C. Rodgers "Optimal Estimation", Phillipe-Tikhonov-Twomey "smoothness constraining"/"regularization", etc. approaches. This means that either of these methods can be easily formulated as specific cases of "Multi-Term LSM", while the inversion configuration proposed by "Multi-Term LSM", as those used in this paper, can't be naturally derived from neither C. Rodgers "Optimal Estimation" nor Phillipe-Tikhonov-Twomey formulation. Therefore, the formulations used in our current paper simply can't be reduced to the mentioned mainstream approaches. Once again, this is discussed in Dubovik et al., (2021).

- Line 339: Do you mean that these constraints were only used for retrievals of 2x2 size or that the correlations are only included between adjacent pixels within your larger retrieval area?

Response: Yes, they are. To clarify this, we modified the phrase as follows: "these constraints played a rather minor role, as they were applied only within a 2-by-2-pixel spatial segment in the spatial-temporal multi-pixel block."

- Line 369: While the description of these experiments is improved, you still do not provide any understanding of which setups were considered. You explain why different channels might have different weights but not why you considered only these increments. If the weights had been based in empirical results, one would expect every channel to be slightly different (due to noise). The use of round numbers implies a decision was made, but that is not explained. A formal way of saying "We tried 0.001 and it worked first time so why try harder" would suffice.

Response: The optimal combination of the standard deviations for different instruments was obtained through a few tenths of experiments. Describing all these tests will make the paper too technical. To avoid this, we added a general phrase: "Experimenting with different values of the standard deviation, it was found …"

- Table 4: Why are the temporal constraints provided qualitatively? Surely there is some range of values that could be given, and you could explain the process by which that value is chosen. The current text provides no guidance so we can only assume you do whatever makes you look best, which isn't very scientific.

Response: As we explained in the paper, the values "are not unique, since quite similar retrieval performance of aerosol and surface properties can be obtained within a certain range of the constraints." We would rather avoid providing a specific ranges in this paper since this requires much more extensive studies.

- Figure 4+: Why is a colour bar never provided? It becomes very hard to interpret the images without a scale, particularly Fig. 7 where N~100 implies that every point should denote a single collocation and, therefore, every point would have the same colour since no two points overlap.

Response:

A colour bar was not provided because we think that it doesn't represent the value that can be used in analysis. As it was described in the caption to Fig.4. "In the scatter plot, the colour of each datapoint indicates the two-dimensional probability density function (2D PDF) for AOD values (increasing PDF value corresponds to colour gradient from blue to red)". To show typical values of 2D PSD for AOD, AE, and SSA, Figures 4, 6, and 7 were updated with a colour bar representing PDF.

- Line 516: "reasonable" would be a better description than "high" as the optimal GCOS fraction is 68% (as opposed to 100%).

Response: Overall, AOD retrieval from space-borne measurements with fulfillment of GCOS requirements higher than 50% is usually considered a high-quality retrieval. A few algorithms can reach such requirements (Popp et al., 2016; Cheng et al., 2020, 2024; Hasekamp et al., 2024, Litvinov et al., 2024).

- Line 731: I'm not convinced by this explanation, needing something like Fig. 2 from doi: 10.1029/2002JD002803 to illustrate that your data sample more of angular space than MODIS. However, I don't think you have to do that here is a later paper on BRDF is produced.
Response: This conclusion is based on a general understanding of Sentinel-3A/OLCI, Sentinel-3B/OLCI, and S5P/TROPOMI measurements, as well as the number of measurements used in the synergy. In particular, the same pixel observed from the Sentinel-3A and Sentinel-3B belongs to very different parts of the orbits: the western part of Sentinel-3A/OLCI can overlap with the eastern part of Sentinel-3B/OLCI. As a result, within a several minutes, the same pixel on the surface can be observed at very different geometries in forward and backward directions. On top of this, TROPOMI afternoon observation geometry is very different from the morning observations of the Sentinel-3A and Sentinel-3B. GRASP multi-pixel approach with temporal smoothness constraints accumulates such measurements within 1 month. Therefore, we expect better angular sampling from the synergy. Indeed, a detailed analysis will be performed in a separate paper, where this will be demonstrated with statistics of the measured geometry.

- Appendix A: I am unclear on what the star is supposed to indicate.
Response: The star indicates measurements or measurements a priori.

- Typographic corrections
o Line 296: The successful application of
Response: Fixed.
o Line 304: performed on data accumulated in spatial-temporal blocks
Response: Fixed.
o Line 312: by applying temporal a priori
Response: Fixed.
o Line 336: as was done
Response: Fixed.
o Line 578: SYREMIS LEO+LEO
Response: Fixed.
o Line 590: comparable to those
Response: Fixed.
o Line 605: It enables transfer
Response: Fixed.
o App. A: Many equations have empty square symbols where post/sub-scripts have been omitted from the Word interface.
Response: Fixed.
o Eq. 3A: I believe the g(a) functions should have subscripts of t,x,y rather than t for all three lines.
Response: Fixed.
o Line 827: systems (Eqs. 2A and 3A) simultaneously
Response: Fixed.
o Eq. 4A: I think the last a on the second line should be bold as the quantity is still a vector.

Response: Fixed.
o Line 835: comma before the first C
Response: Fixed.
o Line 875: retrieval are is not only
Response: Fixed.
o Eq. 11A: The last line has two symbols that did not render.
Response: Fixed.
o Line 897: I don’t think “continues” is the word meant here but it isn’t obvious what would be.
Response: Fixed.