Response to Reviewer

We sincerely thank the reviewer for the evaluation of our work and the helpful suggestions to improve the clarity and rigor of the manuscript. Below, we provide detailed point-by-point responses. Comments from reviewers are shown in blue, **our responses in black**, and the corresponding changes made in the manuscript are highlighted in orange.

Comment 1

Line 5: I would add here that the method requires prior assumptions about effective radius, single scattering albedo, and asymmetry factor. These are non-trivial constraints and the current title and abstract imply that just spectral AOD is needed, which is misleading. I would even consider adding "and optical properties" or something in the title.

Response 1

The title is changed based on this comment: "Retrieval of aerosol composition from spectral aerosol optical depth and optical properties using a machine learning approach"

Comment 2

Line 11: "In the total retrieval uncertainty, the forward model contributes less than 10%, confirming its robustness." I am still uneasy about this statement because it depends strongly on the relative uncertainty of AOD (which depends on the absolute AOD) and on the strengths of the prior and model parameter constraints. I suggest deleting it.

Response 2

We agree and delete it.

Comment 3

Line 81: "compositions" here should be singular "composition".

Response 3

Line 81: "compositions" has been changed to "composition".

Comment 4

Section 2: I think there should be a section for MERRA2 here, since this is an input to the retrieval. Part of my confusion on the previous version of the manuscript was on the role of MERRA2. I think having it up front in this Data section would make things clearer and the method more reproducible. Additionally, as there are so many file types and variable names, the specific file types and variables used should be written (i.e. the five components, how they are converted to number concentrations from optical depths, and then the simple fraction definition).

Response 4

Thanks, following your comment, we have substantially revised the “Data and Methods” sections to introduce MERRA2 as an input dataset to the retrieval framework.

Specifically, we have added a dedicated subsection in Sect. 2 describing the MERRA-2 aerosol reanalysis, including the exact product used (M2T1NXAER / tavg1_2d_aer_Nx), the temporal resolution, and the aerosol components considered (sea salt, sulfate, dust, black carbon, and organic carbon). The relevant variables extracted from the MERRA-2 files are now explicitly listed to improve transparency and reproducibility.

In addition, a new subsection has been added to the “Methods” section describing how component-resolved AOD from MERRA-2 is converted into aerosol number concentration. This includes explicit formulations relating AOD to particle number concentration through representative extinction cross sections and effective radii, followed by normalization to obtain relative number concentration fractions. We emphasize in the revised text that these fractions are used as auxiliary prior information rather than as exact microphysical constraints. The revised section is as followed:

Line 108-117:

2.2 MERRA-2 aerosol reanalysis data

The Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) is the latest global atmospheric reanalysis by NASA’s Global Modeling and Assimilation Office (GMAO) using the GEOS atmospheric model (version 5.12.4). MERRA-2 provides a physically consistent, long-term record of meteorological and aerosol variables from 1980 to the present (Gelaro et al., 2017).

In this study, we use the MERRA-2 aerosol product M2T1NXAER (also referred to as tavg1_2d_aer_Nx), which is an hourly, time-averaged, two-dimensional dataset. This product includes assimilated aerosol diagnostics such as column-integrated AOD at 550 nm. The aerosol species in MERRA-2 include black carbon, dust, sea salt, sulfate, and organic carbon. The aprior information used in this study is the relative number concentration fractions of individual aerosol components. The derivation of these fractions from MERRA-2 aerosol optical depth data is detailed in Sect.3.7.

Line 292-312:

3.7 MERRA-2 a priori information

Aerosol optical depth (AOD) can be expressed as the product of aerosol number concentration and the particle extinction cross section. For a given aerosol component, the column AOD at wavelength λ can be written as

$$AOD(\lambda) = \int N(z) \cdot \sigma\_ext(\lambda)\, dz,$$

where $N(z)$ is the particle number concentration at height z, and $\sigma\_ext(\lambda)$ is the extinction cross section of an individual particle.

The column-integrated number concentration N_col is defined as the vertical integral of the particle number concentration,

$$N\_col = \int N(z)\, dz.$$

Substituting this definition into the expression above yields

$$AOD(\lambda) = N\_col \cdot \sigma\_ext(\lambda).$$

The extinction cross section can be further expressed as

$$\sigma_{ext}(\lambda) = Q_{ext}(\lambda) \cdot \pi \cdot r_{eff}^2,$$

where Q_{ext} is the dimensionless extinction efficiency and r_{eff} is the effective particle radius.

Combining these expressions gives a first-order estimate of the aerosol column number concentration,

$$N_{col} = AOD(\lambda) / [Q_{ext}(\lambda) \cdot \pi \cdot r_{eff}^2].$$

For each aerosol component in MERRA-2, representative values of the effective radius r_{eff} are adopted following the aerosol size assumptions used in the GEOS model. Accumulation-mode aerosols, such as sulfate and organic carbon, typically have effective radii on the order of $r_{eff} \approx 0.3$–0.4 μm, while other aerosol components exhibit component-dependent sizes. The corresponding extinction efficiencies $Q_{ext}(\lambda)$ are obtained from MOPSMAP simulations using consistent refractive indices and particle size assumptions.

Finally, the estimated column number concentrations are normalized across aerosol components to obtain relative number concentration ratios. These ratios are used as prior information in the retrieval.

Comment 5

Line 105: which wavelength(s) are SSA and AF used at? From table 1 I think just 440 nm, this should be specified here as well.

Response 5

Yes, thanks for your suggestion. SSA and AF are at 440 nm. We add this in the text.

Comment 6

Line 138: I think this sentence as written is too optimistic and would say that these are “sometimes” available from “ground-based” remote sensing observations. They are rarely available robustly from satellite retrievals, especially for the bulk of scenes where AOD is low. And from ground-based remote sensing like AERONET, they are also much more limited than direct-Sun data due to a less frequent measurement cadence and various other scene requirements (e.g. azimuthal symmetry which limits the applicability in cases of plumes where the aerosol is spatially hetereogeneous).

Response 6

We agree with you and have made the following revisions as you suggested:

Line 126: Specifically, the ML model takes as input the aerosol component fractions (i.e., number concentrations of sea salt, sulfate, black carbon, dust, and insoluble aerosols), together with auxiliary parameters such as single scattering albedo (SSA), asymmetry factor (AF), and effective radius (Reff). These parameters are not routinely available from all remote sensing observations: they are only sometimes retrievable from ground-based measurements under specific viewing and scene conditions, and are rarely robustly retrieved from satellite observations, particularly in low-AOD scenes or in cases where the aerosol field is spatially heterogeneous.

Comment 7

Line 191: I still don't see why using MOPSMAP directly should result in worse results than the NN emulator. To me it suggests a code bug or an issue with e.g. minimization settings (maybe something numerical which becomes less of an issue in the NN normalization process). The underlying reality is the same and if the code is written well then it should not care whether it is being fed MOPSMAP or an emulator. So I am still suspicious that the NN emulator is needed at all, though it is of course useful as a speedup.

Response 7:

Thank you for raising this important point. We added the additional experiments and analyses presented in Appendix A, which are intended precisely to clarify this issue. Our results demonstrate that the difference between using MOPSMAP directly and using the NN-based forward model does not arise from numerical instability, code implementation, or minimization settings, but from the dimensionality of the inversion state vector and the intrinsic information content of the observations.

When MOPSMAP is used directly as the forward model, the inversion naturally requires a high-dimensional microphysical state vector (15 parameters in our setup, including mode radii, size distribution widths, and number concentrations). The results in Appendix A show that, even under idealized conditions with synthetic observations generated by the same forward model, the degrees of freedom for signal (DOFS) is only about 1–2 for this full state vector. Most microphysical parameters are therefore unconstrained by the available optical observations, and the inversion is fundamentally under-determined. This behavior is reflected in the averaging kernel analysis, which shows that the majority of state vector elements are dominated by the prior rather than by the measurements.

Importantly, we also used the fixed aerosol size experiment in Appendix A to see that, if the aerosol size distribution is assumed to be known, the dimensionality of the state vector can be substantially reduced and the retrievable information on aerosol composition increases or not. This naturally raises the question of whether further information could be gained if additional optical properties, such as SSA and AF treated as inputs. In a fully physical forward model like MOPSMAP, however, these quantities are unknown outputs of the model rather than known inputs. Therefore, they cannot be fixed independently in the inversion. This limitation prevents the physical model from exploiting externally available information on SSA or AF.

By contrast, the machine-learning–based forward model adopted in the main text is constructed as a data-to-data mapping, allowing SSA, AF, effective radius, and other optical parameters to be incorporated explicitly as inputs. In this way, microphysical complexity remains embedded in the learned mapping, while the inversion state vector is aligned with the actual information content of the observations. This explains why the ML-based forward model enables a stable and informative retrieval, whereas a purely physical forward model becomes ineffective when applied to a high-dimensional microphysical state space.

Comment 8

Equation 2: I would move lines 200-202 here instead of where they currently are. I was initially confused because Table 1 gives not just AOD but also Reff, SSA, and AF as outputs but in the sense of this section, they are inputs. But then line 200 explains where the 9 inputs come from and what the outputs are.

Response 8:

Thanks, we agreed with you. We've revised here as you suggested.

Comment 9

Line 205: the citation for Adam optimization should be given. Yes it's common, but the ReLU citation is given and that is also common...

Response 9:

Thank you. We've revised here as suggested.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, https://arxiv.org/abs/1412.6980, 2017.

Comment 10

Equation 12 and sections 3.5 and 4.3: it looks like uncertainties in the model parameters (theta, i.e. SSA, AF, Reff, RH) are not accounted for in the cost function, or in Sy. This is essentially making the assumption that the constraint on these (from AERONET or MERRA2) is perfect, which is obviously unrealistic. For example, even in good conditions the AERONET SSA uncertainty is about 0.03. Given most of the time SSA various from about 0.8 to 1, this is not a small fraction of the range. Looking at e.g. the early Dubovik 2002 aerosol type paper (https://doi.org/10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2)one can see that an envelope of +/- 0.03 can encompass e.g. smoke, dust, urban aerosol, and marine dependent on where it falls. And that asymmetry factor shows small distinction between types at 440 nm so may have limited value. This means that the error budget is incomplete and, as a result, analyses based on the propagated uncertainties will be overconfident (the results seem better than they really are). This should at minimum be clearly acknowledged in the manuscript. Again, it is misleading about the fidelity of the results otherwise. Ideally, these uncertainties should be accounted for. One method to do so would be to perturb the four parameters in theta to simulate realistic errors (which might be different from MERRA2) and incorporate this either into Sy (via seeing how it affects calculated AOD in the MOPSMAP emulator). Incorporating model parameter error like this into Sy is common within the OEM. Or, this could be added as an additional uncertainty term within Equation 16.

Response 10

Thank you very much for pointing out that. The uncertainties in auxiliary parameters (SSA, AF, effective radius, and RH) were not explicitly accounted for in the original error budget, which could lead to overconfident retrieval results. Following this comment, we have added a dedicated subsection (Sect. 3.6, "Sensitivity analysis to auxiliary parameters") to explicitly quantify the impact of these uncertainties on the retrieved aerosol composition.

In the revised manuscript, we adopt representative uncertainty ranges for each auxiliary parameter (±0.03 for SSA, ±0.05 for AF, ±20% for Reff, and ±5% for RH), consistent with reported uncertainties from AERONET and previous studies (e.g., Dubovik et al., 2002). We then propagate these uncertainties through the trained machine-learning forward model using an ensemble-based retrieval experiment, in which all auxiliary parameters are perturbed simultaneously according to Gaussian distributions and the optimal estimation retrieval is repeated 300 times.

The resulting distributions of retrieval deviations (new Fig. 5) demonstrate that auxiliary-parameter uncertainty induces only small perturbations in the retrieved aerosol composition, typically within ±5%. More specifically, different aerosol components exhibit distinct responses to perturbations in the auxiliary parameters. Sea salt and dust tend to show slightly positive deviations, whereas sulfate exhibits a weak negative deviation. In contrast, black carbon and insoluble aerosol display no systematic offset, with their deviation distributions centered near zero. These effects do not alter the identification of the dominant aerosol component.

The following part have been added as followed:

3.6 Sensitivity analysis to auxiliary parameters

To assess the impact of uncertainties in auxiliary optical parameters on the aerosol composition retrieval, we conducted a two-step sensitivity analysis using the trained machine-learning (ML) forward model. Aerosol composition fractions are retrieved from multi-wavelength AOD observations, while single scattering albedo (SSA), asymmetry factor (AF), effective radius (Reff), and relative humidity (RH) are treated as auxiliary inputs. In the retrieval framework, these auxiliary parameters are assumed to be fixed but subject to uncertainty. In the following analysis, we quantify how uncertainties in these auxiliary parameters propagate into the simulated AOD spectrum and subsequently affect the retrieved aerosol composition.

We adopt representative uncertainty ranges of ±0.03 for SSA, ±0.05 for AF, ±20% for Reff, and ±5% for RH. These values are consistent with typical AERONET uncertainty estimates for SSA and with commonly reported uncertainties for aerosol optical properties (e.g., Dubovik et al., 2002).

The impact of auxiliary-parameter uncertainty is quantified using an ensemble-based retrieval experiment designed to propagate realistic parameter errors. In this experiment, all four auxiliary parameters (SSA, AF, Reff, and RH) are perturbed simultaneously according to Gaussian distributions defined by their assumed uncertainty ranges, and the optimal estimation retrieval is repeated for each perturbed realization. A total of 300 realizations are performed to characterize the resulting variability in the retrieved aerosol composition. The ensemble of retrieved aerosol compositions is then summarized to assess the sensitivity of the retrieval to auxiliary-parameter uncertainty.

4.4 Results of sensitivity analysis to auxiliary parameters

The distributions in Fig.5 show the deviations of the retrieved aerosol component fractions relative to the baseline retrieval obtained with fixed auxiliary parameters. For all aerosol types, the perturbation-induced deviations are narrowly distributed within ±5% and centered close to zero, indicating that uncertainty in the auxiliary parameters does not introduce systematic biases in the retrieval.

The dominant aerosol components (e.g., sulfate and dust) exhibit approximately symmetric distributions with typical spreads on the order of a few percent. In contrast, minor components such as black carbon and insoluble aerosol show strongly peaked distributions near zero, reflecting both their small absolute contributions and the limited sensitivity of the spectral observations to these components. More specifically, different aerosol components exhibit distinct responses to perturbations in the auxiliary parameters. Sea salt and dust tend to show slightly positive deviations, whereas sulfate exhibits a weak negative deviation. In contrast, black carbon and insoluble aerosol display no systematic offset, with their deviation distributions centered near zero.

It indicates that if uncertainties in the auxiliary parameters (SSA, asymmetry factor, effective radius, and relative humidity) are not explicitly accounted for in the inversion, the retrieval may underestimate sea salt and dust while overestimating sulfate. However, the magnitude of these deviations remains small, and the overall aerosol composition, particularly the dominant component, is not qualitatively affected.

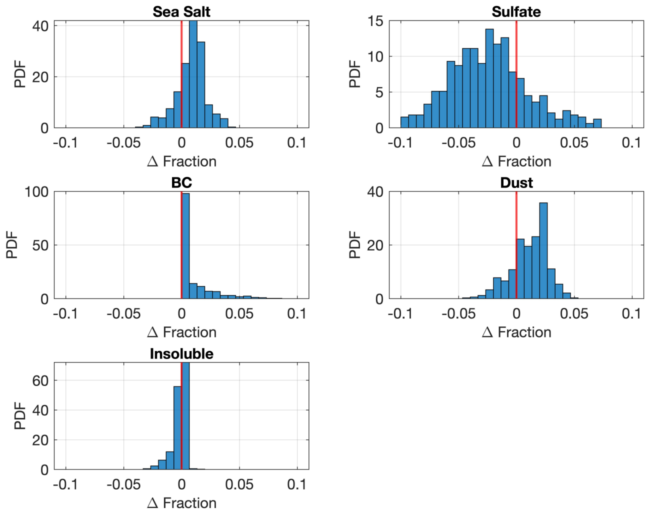

Figure 5. Distributions of relative changes in the retrieved aerosol component fractions from the sensitivity tests described in Section Sensitivity analysis to auxiliary parameters. Four auxiliary parameters (single scattering albedo at 440 nm, asymmetry factor at 440 nm, effective radius, and relative humidity) are perturbed, and deviations from the baseline retrieval are shown for each aerosol fraction. Red vertical lines indicate zero deviation.

Comment 11

Line 248: what is the justification for these prior uncertainties? It seems arbitrary. If the numbers are changed then the uncertainties and averaging kernel will change as well. If the end use is to take e.g. MERRA2 component fractions as input, then it should be determined by analysis of MERRA2 component fraction uncertainty (which I would imagine is a function of location, among other things).

Response 11:

We thank you for this important question regarding the justification of the prior uncertainties. We agree that the choice of prior uncertainty affects the posterior uncertainty and averaging kernel and therefore requires clear justification.

In this study, two different treatments of prior uncertainty are adopted for two different purposes. For the synthetic (virtual-spectrum) experiments, a fixed prior uncertainty (0.1 in relative fraction) is prescribed. The purpose of these experiments is methodological: to diagnose the intrinsic information content of the observations and to examine the behavior of the inversion framework under controlled conditions.

For the retrievals using actual observations, the prior uncertainty is not treated as a fixed or arbitrary value. Instead, it is derived empirically from MERRA-2 data and is location dependent. Specifically, for each observation site, we compute the standard deviation of the MERRA-2 aerosol component fractions within a latitude–longitude box of ±1° centered on the observation location. This spatial variability is then used as a proxy for the prior uncertainty, reflecting regional heterogeneity in aerosol composition and addressing the reviewer's concern that prior uncertainty should depend on location.

The a priori covariance matrix Sa is specified as a diagonal matrix. For the synthetic (virtual-spectrum) experiments, a uniform variance of 0.01 (i.e., a standard deviation of 0.1) is prescribed for each aerosol component. For retrievals using actual observations, the diagonal elements of Sa are derived empirically from the spatial variability of MERRA-2 aerosol component fractions within a ±1° latitude-longitude box centered on the observation location, and are used as a proxy for location-dependent prior uncertainty.

Comment 12

Line 252: Note the AERONET team report direct-Sun AOD uncertainty of 0.02 at 440 nm and shorter wavelengths. 0.01 as used here is for the longer visible wavelengths. So this should be updated.

Response 12

Corrected as recommended.

For the measurement error covariance matrix Sy, we distinguish between visible and shortwave infrared (SWIR) wavelengths. For visible bands, we adopt wavelength-dependent AOD uncertainties following AERONET direct-Sun uncertainty estimates, with a standard deviation of 0.02 at 440 nm and shorter wavelengths, and 0.01 at longer visible wavelengths.

Comment 13.

Section 4.5: if I understand correctly, in this section Reff, SSA, AF, and RH are switched around to be retrieved and not assumed. But none of these results are shown, and this is inconsistent with the rest of the retrieval development and analysis in the paper. It is also not clear why the authors chose this as opposed to just using auxiliary inputs like they did elsewhere in the paper. Line 381 sounds honestly like an excuse like the authors did not want to download the MERRA2 data needed for the case study. I also previously had concerns about the reasonableness of using monthly data for this purpose (it ignores sampling issues and real sub-monthly variability). In my view a simple monthly plot and then showing dust AOD is not convincing enough. So I do not think this section as presented is very useful in the context of the paper. My recommendation would be to keep the same retrieval methodology (i.e. auxiliary Reff, SSA, AF, and RH) as elsewhere in the paper, and to do the analysis using daily instead of monthly data. Then the data could be compared on a daily basis with the GEOS fields (sampled around the early-pm satellite overpass time) and a commonly-sampled dust AOD (and AOD for the other aerosol types) could also be presented here. This would be much more direct and meaningful demonstration of the method.

Response 13

As you correctly point out, the inversion strategy within the satellite differs from the primary methodology, treating Reff, SSA, AF, and RH as inversion variables rather than auxiliary inputs. Moreover, as you suggest, meaningful satellite validation necessitates ensuring consistency in spatio-temporal sampling.

We concur with your recommendation that inversion should employ identical methodologies during validation to guarantee consistency and clarity of results. We further agree that existing monthly averaging overlooks submonthly variability and sampling issues, and as the reviewer suggests, meaningful satellite validation must ensure consistency.

Therefore, without compromising the overall methodology or conclusions in our manuscript, we have decided to remove this satellite validation part from the main context. The revised version will retain the original auxiliary parameter inputs (Reff, SSA, AF, and RH) and focus entirely on ground applications, thereby enabling more precise constraints and validation.

We appreciate your suggestions and will undertake more precise and detailed satellite validation in future work. Building upon the current methodology, we will update and validate its feasibility and uncertainty in satellite inversion applications.

Comment 14. Appendix A: is this only supposed to be one paragraph? I am not sure why it is an Appendix, and as mentioned earlier think it would be better as a subsection of Section 2.

Response 14

We have removed all results concerning satellites, so this part has been deleted accordingly.