Review1

Review for “An AI Based Algorithm for Retrieving Aerosol Optical Depth and Single Scattering Albedo Using All-Sky Imager Observations" by Ni et al.

This study presents a machine learning (XGBoost) approach to retrieve Aerosol Optical Depth (AOD) and Single Scattering Albedo (SSA) from all-sky imager (ASI) data, using collocated AERONET measurements as training targets. The method is tested at two contrasting sites (Beijing, China and Southern Great Plains, USA) and demonstrates promising skill, particularly for AOD. The work addresses an important gap—low-cost, high-temporal-resolution aerosol monitoring—and offers a clear advancement over traditional radiative transfer-based methods by avoiding iterative inversion.

Overall, the manuscript is well-structured, the methodology is sound, and the results are presented clearly. However, several issues—ranging from methodological transparency to the interpretation of SSA performance—should be addressed before publication.

Major comments:

1. While the method has been applied to two independent sites, Beijing and SGP, I noticed different versions of AERONET data were used. Specifically, the authors state that Level 2.0 SSA retrievals at Beijing are too sparse due to strict quality control, so they use Level 1.5 instead. But Level 1.5 SSA certainly has larger uncertainty, the authors should explain how this uncertainty impact the retrieved SSA results, and if possible, provide some quantified estimation of the retrieved SSA uncertainty.

   Response: Thank you for pointing this out. We have clarified that AERONET Version 3 Level 2.0 SSA inversion products are not available for the period overlapping with the Beijing_PKU ASI observations, rather than being merely sparse or unavailable for the site in general. This is mainly because the post-deployment calibration required for fully quality-assured Level 2.0 inversion products could not be completed in time, due to pandemic-related restrictions on international shipment and calibration of the instrument. Therefore, Beijing SSA retrievals were trained and validated against Version 3 Level 1.5 cloud-screened inversion products. We agree that Level 1.5 SSA has larger uncertainty than Level 2.0.

   Furthermore, we performed a Monte Carlo sensitivity experiment by adding Gaussian random perturbations to the AOD and SSA training labels. The results show that SSA retrievals are more sensitive to the assumed label uncertainty than AOD retrievals. For SSA, the propagated prediction uncertainty increased from 0 to 0.0130 and 0.0214 under the low- and high-uncertainty scenarios, respectively. Correspondingly, SSA RMSE increased from 0.0318 to 0.0331 and 0.0384, while the correlation decreased from 0.608 to 0.580 and 0.489. In contrast, AOD RMSE and correlation remained nearly unchanged, indicating greater robustness of AOD retrievals to the assumed label uncertainty.

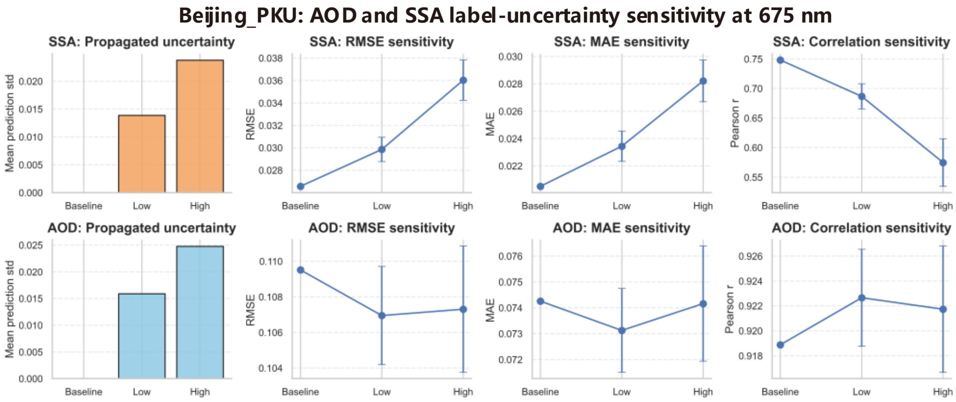

Figure S2. Sensitivity of Beijing_PKU AOD and SSA retrievals to assumed AERONET label uncertainty at 675 nm. Gaussian perturbations were added to the AOD and SSA training labels under three scenarios: Baseline (σSSA = 0, σAOD = 0), Low uncertainty (σSSA = 0.03, σAOD = 0.02), and High uncertainty (σSSA = 0.06, σAOD = 0.04). The model was retrained 100 times for each scenario. The first column shows the mean standard deviation of the predicted values, representing the propagated retrieval uncertainty. The remaining columns show the mean RMSE, MAE, and Pearson correlation coefficient across the Monte Carlo runs, with error bars indicating ±1 standard deviation. The perturbation amplitudes are used as assumed uncertainty scenarios rather than official AERONET Level 1.5 or Level 2.0 uncertainty values.

To reduce the potential uncertainty of the Level 1.5 SSA records, this revised version has further applied additional quality-control criteria to the Beijing_PKU inversion data. Specifically, only SSA retrievals with solar zenith angle greater than 50° and sky residual less than 5% were retained. These criteria remove less stable sky-radiance inversion cases associated with unfavorable observation geometry or relatively large fitting residuals, for which SSA retrievals are expected to be more uncertain. We have revised Section 2.2 to make this quality-control procedure explicit. We also clarify that the Beijing SSA validation should be interpreted as agreement with additionally quality-screened Level 1.5 AERONET inversion products, rather than with fully quality-assured Level 2.0 products.

2. The practice to simultaneous retrieve AOD and SSA from ASI is indeed novel. Butthe SSA performance—especially at Beijing (r as low as 0.33 at 1020 nm, Fig. 4 and Table 1)—is relatively weak. I wonder how this performance compares with existing SSA products, perhaps from satellites or other remote sensing techniques? Such comparison is important in understanding the capability and limitions of ASI in retrieving aerosol properties.

Response: We agree that comparison with existing SSA products is important for interpreting the capability and limitations of the proposed ASI approach. So we have added a comparison with representative ground-based and satellite-based SSA retrieval products in Section 3.1. The revised manuscript compares the reported correlation coefficients and RMSE values with AERONET, PARASOL/POLDER-GRASP, GEMS, and Gaofen-5 DPC products. We also clarify that these comparisons should not be interpreted as evidence that the ASI retrieval outperforms established products, because the products differ in wavelength, viewing geometry, aerosol loading, spatial representativeness, and reference-data quality. Instead, the results suggest that ASI observations may provide complementary, low-cost, high-temporal-resolution local SSA information.

Minor Comments

1.Inconsistent figure labeling in caption of Figure 1

The caption for Figure 1 lists (a)-(d) but the text refers to (a-c) and (d-f) in the following paragraph. Please check and harmonize the labels. Also, the source attribution for (c,d) is given, but not for (a,b) – add "source: authors" for clarity.

Response: We have corrected the panel labels in the text and caption of Figure 1. We have also added source attribution for panels (a) and (b), which are images from our team's own all-sky imager observations at the Beijing_PKU site. The source attribution for panels (c) and (d), which are from the ARM user facility, has been retained.

2.Lack of detail on the cloud and lens flare masking algorithm

Section 2.3 mentions "iterative threshold segmentation" (citing Anon, 1978; Gonzalez and Faisal, 2019) but does not specify the actual threshold criterion (e.g., Otsu? fixed percentile?). Given that cloud masking quality directly affects sky radiance extraction, please provide the specific algorithm or pseudo-code, or at least cite a more recent and relevant ASI cloud masking method.

Response: We have added a detailed description of the cloud and lens-flare masking procedure in Section 2.3. The revised text specifies the solar-region detection procedure, mask radius, initial threshold, iterative threshold-update rule, and convergence criterion. We have also replaced the former "Anon, 1978" citation with the complete reference to Ridler and Calvard (1978).

The revised text in section 2.3: *"Clean-Sky Image Processing: Upon acquiring time-stamped raw sky images, we first mask areas that interfere with the extraction of cloud-free sky radiance—namely buildings, clouds, and glare spots caused by the sun and lens artifacts. The specific steps are:*

(a) Masking buildings at the image periphery.

(b) Identifying and masking the brightest circular region (sun and its immediate surroundings): the solar center was identified from the grayscale image using an integral-image-based search. A square window with a radius of 45 pixels and a step size of 5 pixels was used to locate the brightest region, and a circular mask with a radius of 220 pixels was applied around the detected solar center. For the TSI images, an additional fixed circular/elliptical mask was applied to remove the instrument support structure.

(c) Cloud and lens flare masking: all-sky images were first converted to grayscale for masking. Cloud-contaminated and other bright regions were removed using an iterative thresholding scheme (Ridler and Calvard, 1978)(Anon, 1978)(Anon, 1978)(Anon, 1978). The initial threshold was set to 150. At each iteration, pixels with non-zero intensities were separated into two groups according to the current threshold: pixels with intensity values greater than the threshold were assigned to the background group, whereas pixels less than or equal to the threshold were assigned to the foreground group. The mean intensities of the two groups were then calculated, and the threshold was updated as the average of the two means. The iteration was stopped when the threshold change was less than or equal to 1. Pixels classified as background were set to zero in the RGB image, producing a cloud-masked image. Iterative threshold segmentation is applied until the threshold stabilizes, after which a black mask is used to occlude high-brightness regions associated with clouds and flares (Anon, 1978; Gonzalez and Faisal, 2019; Zhang and Xiao, 2014)."

3. The diurnal cycle analysis (Section 3.3) would benefit from a quantification of the 5-minute smoothing effect.

The authors apply a 5-minute smoothing to the ASI retrievals in Figure 6 but then report statistics on unsmoothed data. This is acceptable, but the magnitude of high-frequency noise should be characterized (e.g., standard deviation of the difference between smoothed and unsmoothed). Without this, readers cannot assess whether the fine-scale variability in Figure 6a is signal or noise.

Response: We have quantified the effect of the 5 min smoothing by calculating the difference between the raw ASI retrievals and the 5 min moving average. For the case shown in Fig. 9, the mean raw-minus-smoothed difference is 0.0008, with a standard deviation of 0.013, mean absolute difference of 0.0076, and 95th percentile absolute difference of 0.027. We have added these values to Section 3.5 and included a supplementary Figure S1 showing the raw series, smoothed series, residual time series, and residual distribution. We also clarified that all validation statistics were calculated using unsmoothed retrievals.

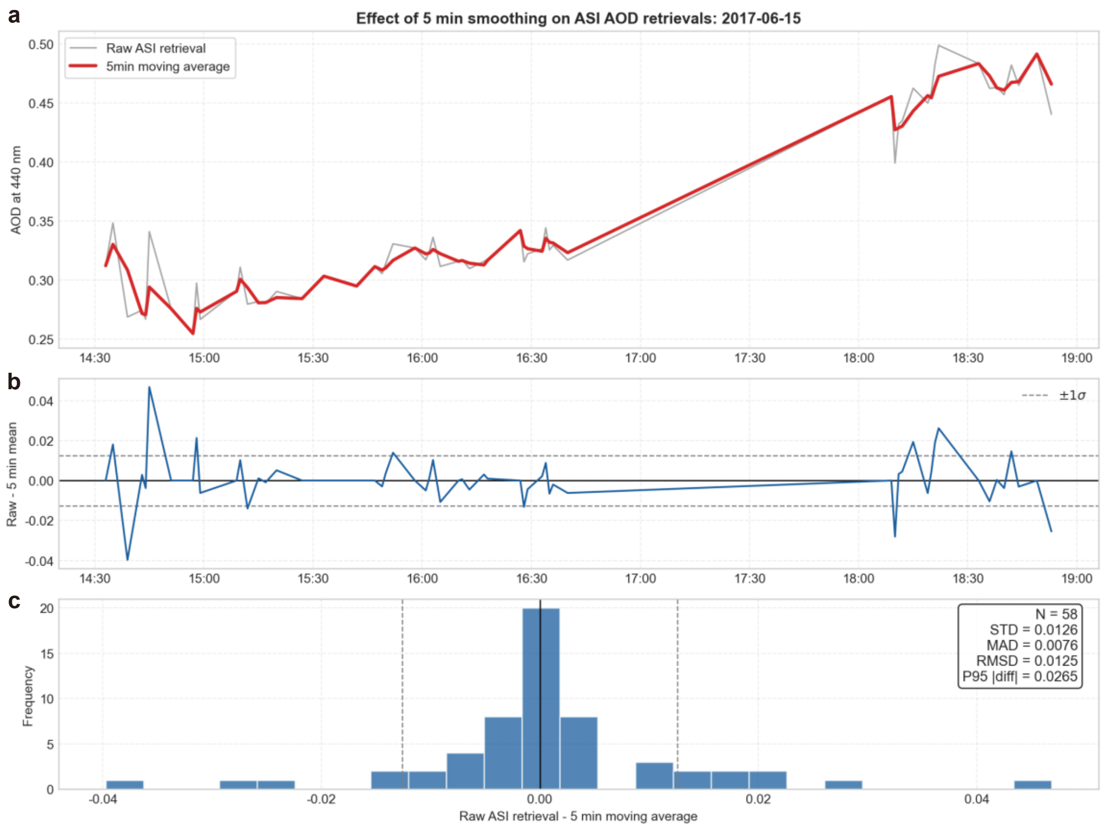

Figure S1. Quantification of the 5 min smoothing effect for ASI-retrieved AOD at 440 nm on 15 June 2017. Panel (a) compares the raw ASI retrievals with the 5 min moving average, the middle panel shows the difference between the raw and smoothed retrievals, and the lower panel shows the distribution of the difference. The standard deviation of the raw-minus-smoothed difference is 0.0126, with a mean absolute difference of 0.0076 and a 95th percentile absolute difference of 0.0265.

4. Typographical and formatting issues

- Line 30: "Almuantar" should be "almucantar"
- Line 135: "XGBoost" appears twice in close succession – rephrase for readability.
- Table 1: The column headers "440nm, 675nm..." lack spaces (use "440 nm" for consistency with text).
- Line 195: "alucantar" → "almucantar"
- Reference "Anon, 1978" – please replace with the actual author names (it is a well-known paper by Ridler and Calvard or similar).

Response: Thank you for these careful checks. We carefully checked all occurrences of “almucantar” in the revised manuscript and did not find the spellings “Almuantar” or “alucantar” in the current text. To ensure consistency, we standardized the term as “almucantar” throughout the manuscript. We carefully checked the relevant sentence and rephrased the nearby model-description text to avoid unnecessary repetition of “XGBoost” and improve readability. We also revised the wavelength labels in Table 1 from “440nm”, “675nm”, etc. to “440 nm”, “675 nm”, etc., and replaced the incomplete “Anon, 1978” citation with the full reference to Ridler and Calvard (1978).