We would like to express great appreciation to the reviewer for their insights and detailed review that made this paper stronger and clearer. We have taken them all into serious consideration. Our responses (in blue) for each comment (in black) and updates to the paper (in *italics*) are provided below.

Report #1 Anonymous referee #3

The manuscript compares aerosol layer height retrievals from three different passive sensors (GEMS, TROPOMI, EPIC), using the active sensor CALIOP as a reference. In addition, the aerosol optical depth and UV aerosol index are compared between the passive sensors and AERONET observations over southeast Asia. Altogether, the text is clearly written, and the comparison methodology is scientifically sound and mostly well documented. There are few details that should be elaborated upon. For example, the text does an excellent job summarizing the different definitions for aerosol layer height used by the instruments. To ensure an apples-to-apples comparison, the text says that some of the aerosol layer heights are converted from one definition to another so that they all use the same definition (lines 258-259). More details should be added to explain how the aerosol layer height values are converted. Also, CALIOP data are used as a reference in the methodology, but some critical details about the CALIOP products used are missing. For example, which data version number, what quality screening was used, and what are the limitations of using the CALIOP retrievals as a reference? These details along with a summary of how the CALIOP extinction retrievals are calculated should be added as a subsection of the Section 2 (Data and methodology). Likewise, a summary of the AERONET data should be added to that section, including details on version number, which ground stations were evaluated, and quality screening information.

The analysis strategy is sound. The manuscript provides an objective summary of the intercomparisons and the conclusions are supported by the data presented. This manuscript is suitable for Atmospheric Measurement Techniques. Between the comments above and the specific comments below, I recommend minor revisions are needed prior to publication.

Thank you for your feedback. We agree that adding details will strengthen our paper. First, for the conversion of aerosol layer height definitions, we used a look-up table approach. As shown in Figure 2, we highlight the differences in aerosol layer height definitions based on height, which lead to the development of our conversion look-up table. Throughout this paper, we conducted two conversions: first, we aligned all passive sensor aerosol layer height products with the CALIOP AOCH definition for comparison, as illustrated in Figure 1b. Second, we converted GEMS AEH to the EPIC and TROPOMI AOCH definition, based on Figure 1c. These details are added at the end of Section 2.2 as follows:

In our further comparison of ALH, we count for these inherent differences by converting one definition to another to ensure consistency. Varying AOCH from 0-10 km, we created a look-up table of AEH and, EPIC/TROPOMI AOCH, and CALIOP AOCH corresponding to the same aerosol extinction profile according to their different definitions. Throughout this paper, we conducted two conversions to ensure consistency: first, we converted all passive sensor ALH products with the CALIOP AOCH definition for comparison with CALIOP data (Fig. 1b). Second, we converted GEMS AEH to the EPIC and TROPOMI AOCH definition for comparisons among passive remote sensing products (Fig 1c).

Additionally, we changed the title of Section 2 from ‘Satellite data’ to ‘Remote sensing data’ and added 2.1.4 CALIOP / CALIPSO and 2.1.5 AERONET as follows:

2.1 Remote sensing data

2.1.4 CALIOP / CALIPSO

CALIOP is a lidar system on the CALIPSO platform that provides attenuated backscatter vertical profiles of aerosols and clouds in the atmosphere using a two-wavelength laser operating at 532 nm with linear polarization and at 1064 nm (Winker et al., 2009). While the global coverage of CALIOP is *less than 0.2%, it provides high vertical resolution for retrieving aerosol extinction profiles (Winker et al., 2013). In this paper, we used CALIOP 5 km Level 2 aerosol extinction profile product at 532 nm to derive optical depth weighted heights. Specifically, Level 2 Aerosol Profile, Version 4-21 data product for the year 2021 (CAL_LID_L2_05kmAPro-Standard-V4-21) is used. For the years 2022 to 2023, level 2 aerosol profile, version 4-51 (CAL_LID_L2_05kmAPro-Standard-V4-51) is used due to the data availability. To validate aerosol height retrievals from passive remote sensing with CALIOP observation, the optical depth weighted heights derived from CALIOP 5 km level 2 aerosol extinction profile product at 532 nm following previous studies are used (Chen et al., 2021b; Lu et al., 2023; Lu et al., 2021; Xu et al., 2019).*

2.1.5 AERONET

AERONET is a ground-based remote sensing network to designed to measure and characterize aerosol optical properties through direct sun measurements taken with sun-sky scanning spectral radiometers (Holben et al., 1998). AERONET serves as a critical tool for validating satellite-retrieved aerosol optical properties including AOD. In this study, we used AOD data at 675 nm and 440 nm from AERONET Version 3 Level 1.5 to assess the accuracy of satellite AOD retrievals. AERONET sites located within our study domain of East Asia and Southeast Asia, as illustrated in Figure 2, were selected for this analysis. Additional information of these sites can be found in S1.

Line 57. Comma not needed “present, from”

Comma deleted.

Line 79. The grammar is a little off here. I think it should be “Not only are these techniques based on different physical theories,”

Modified.

Section 2. Background information on the CALIOP data used and AERONET data used should be added to this section. Version numbers, specific products, ground stations (for AERONET), quality screening, etc…)

We added 2.1.4 CALIOP / CALIPSO and 2.1.5 AERONET in the Section 2.

Line 234. Is there a term for σH? In other words, is there a name for this value? This is not critical but might make the description clearer.

σ_H *is the half width parameter defined as:* $\sigma_H = ln\left(3+\sqrt{8}\right)/\eta$

Lines 235-236. “EPIC and TROPOMI defined their retrieved ALH as H in Eq. (1) and called AOCH” Should this sentence say, “…and it is called AOCH”? I think this might make the sentence clearer (or another suitable modification).

Modified to:

EPIC and TROPOMI defined their retrieved ALH as H in Equation (1) and referred to it as AOCH.

Line 243. This is called the “CALIOP 5 km level 2 aerosol profile product”. It contains profiles of aerosol extinction. This and the references in the sentence should be moved to a Section 2 subsection describing the CALIOP data used.

We moved this information to 2.1.4 CALIOP / CALIPSO and 2.1.5 AERONET.

Line 253. It would be more precise to say “reaches approximately 4 km and above” rather than “4 km and beyond”.

Modified.

Lines 258-259. “In our further comparison of ALH,…by converting one definition to another” More details should be added to explain how this is accomplished. My understanding is that the passive sensors report AHL altitudes (with different definitions). How are these altitudes converted?

Answered above in general comments.

Figures 4, 5, 7 captions. Add text to the captions explaining what the gray points indicate.

Grey points indicate points where the data density is less than 0.01.

Lines 391-392. “the ALH values of all passive sensors are converted to AOCH following the CALIOP AOCH definition”. Add the details of how they are converted (here or another suitable location).

Answered above in general comments.

Lines 444-445. “ GEMS AEH is converted to align with the EPIC and TROPOMI AOCH definition”. How is this done?

Answered above in general comments.

Line 457. “This is a combination of inaccurate cloud detection and inherent sensitivity in the retrieval process.” Because the previous sentence involves three different instruments, be specific on which instruments are influenced by inaccurate cloud detection and sensitivity.

This is a combination of inaccurate cloud detection and inherent sensitivity in the retrieval process of TROPOMI and EPIC.

Figure 7 caption. It is not clear which panels are dust and which are smoke. An option could be to say “dust (top row) and smoke (bottom row)”

All panels of Figure 7 are dust and smoke cases combined (all cases). The top row is for GEMS and TROPOMI comparison and the bottom row is for GEMS and EPIC comparison. We changed the caption for clarity.

Intercomparison of AOCH values from GEMS, TROPOMI, and EPIC for all cases (dust and smoke combined) as a function of UVAI. The density scatter plots show the AOCH comparison between GEMS and TROPOMI (a – c), and between GEMS and EPIC (d – f). (b) and (e) represent GEMS data for UVAI < 3, while (e) and (f) represent data for UVAI ≥ 3. GEMS AEH values have been converted to align with the AOCH definitions used by EPIC and TROPOMI.

Line 476. “GEMS and EPIC products for the day are adjusted to relative local solar time (LST) accordingly.” Please rephrase to clarify: the wording makes it sound like the ALH data is adjusted, but I think it is the time of observation that is converted to local time. Is that right? It is not clear. If the actual ALH data is adjusted then an explanation should be added about how this adjustment is made.

How you understood is correct. We rephrased to clarify the time of observation is converted to local time:

Therefore, we define the relative local solar noon time for a given day as the moment when the solar zenith angle at a particular location reaches its minimum value. Using this relative local solar noon as a reference, we adjust the observation times of GEMS and EPIC products with relative local solar time (LST).

Figure 9 and 12. The size of the star is obscuring the ability to see the data underneath. Recommend reducing its size. A description of the star should probably be added to the captions too instead of only the text (e.g., “black star indicates center of study domain”)

We changed the colormap of Figure 9 to the same as that of Figure 12 (according to the comment below) as well as the start color from black to red.

The red star indicates the dust plume area near Beijing.

Figures 9 and 12. Recommend using the same colormap for these images since they are the same quantity. The figure 12 colormap is better given the sequential nature of the data. This an optional recommendation (though I am not sure if the AMT style guide has any rules about this).

We updated the colormap of Figure 9 to match that of Figure 12, as it is more accessible for individuals with color vision deficiencies and adheres to AMT's stylistic guidelines.

Figure 13 caption. Should read, "Same as figure 10" instead of "Same as figure 11"

Yes, thank you for the correction.

Lines 639-640. "…further compounded by inaccurate surface reflectance in TROPOMI and EPIC." This was conjectured in Section 3.1, but not proven. ("The positive y-intercept observed for both EPIC and TROPOMI suggests that the surface reflectance employed in the dust aerosol model may be underestimated…"). The positive y-intercept is consistent with an underestimate of the surface reflection, but other causes where not discussed and ruled out. The statement in lines 639-640 should be modified to say that this is a possible cause that is consistent with the observations in this study.

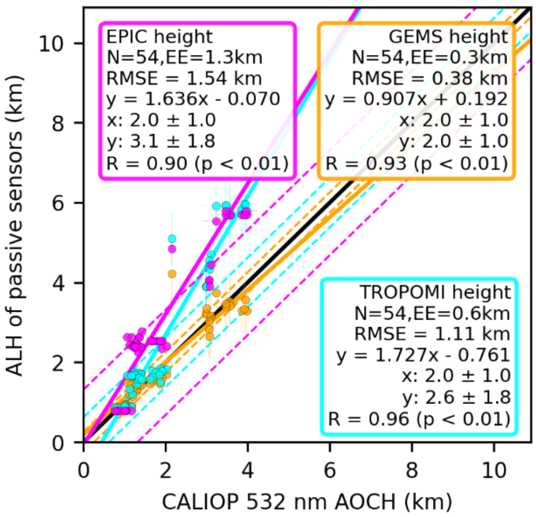
The inaccuracies in the GEMS dust aerosol model contribute to the significant differences in GEMS AOD compared to TROPOMI and EPIC. Additionally, the differences are compounded by other causes including potential inaccuracies in surface reflectance in TROPOMI and EPIC.

Lines 656-661. This paragraph summarizes the intercomparisons of ALH between CALIOP and the three passive instruments. It contrasts the correlations and biases between EPIC, TROPOMI, and GEMS with respect to the co-located CALIOP observation based on the statistics reported in Figs. 10a and 13a. I do not think the passive sensors are being equitably compared because all three passive sensors do not always report ALH values in every profile where CALIOP reports an ALH. For example, in Fig. 10, only GEMS reports ALH south of 35°N. However, those profiles contribute to the statistics for GEM-CALIOP comparisons in Fig. 10a. EPIC and TROPOMI had no ALH reported in these profiles so their comparison statics do not consider this part of the scene. To fairly compare the three passive instruments with respect to CALIOP would be to only compare statistics based on profiles where all three instruments and CALIOP reported ALH. Figures 10a and 13a could probably stay as they are (scatter plot based on the whole scene), however in this section of the conclusions where general statements are being made about the relative accuracy of each instrument compared to another, they should each have ALH values to report.

We agree that it will be helpful to compare ALH values only where all retrieval products (EPIC, TROPOMI, and GEMS) have valid data. However, it is important to highlight the distinct retrieval characteristics of each instrument. While TROPOMI and EPIC AOCH primarily provide retrievals for absorbing aerosols, GEMS retrieves both absorbing and non-absorbing aerosols. This difference reflects the inherent variations in the algorithms of each product and is a critical aspect of the intercomparison.

To address the concern, we have included a supplementary scatter plot where all three instruments and CALIOP report valid ALH data. In this subset, we observe significantly high correlations ($R > 0.9$) between the passive sensors and CALIOP ALH. Specifically, GEMS demonstrates the lowest RMSE (0.38 km), while EPIC and TROPOMI show larger RMSE values of 1.54 km and 1.11 km, respectively, with a tendency to overestimate ALH.

The paragraph is summarizing for individual dust and smoke case studies. Therefore, this additional comparison is now reflected in the 3.3 ALH validation with CALIOP, conclusion, and the supplementary (Fig. S6).



When valid data are available from all retrievals, the passive sensors show a notably high correlation with CALIOP AOCH (R > 0.9), with GEMS having the lowest RMSE (0.38 km), while EPIC and TROPOMI exhibit higher RMSE values of 1.54 km and 1.11 km, respectively (Fig. S6)

Appendix A. List of acronyms should be alphabetized

We alphabetized the list of acronyms.

Report #2 Anonymous referee #4

Dear Authors,

Thank you for this well-written and extremely detailed manuscript describing your comparative analysis of aerosol layer height retrievals from multiple satellite sensors. The potential to characterize the vertical position of aerosol plumes from space using passive sensors is a very exciting one with important applications for air quality monitoring and other areas.

I think your scientific analysis is fundamentally sound. I have some recommendations for clarifying your account, and a few places where rewording is needed to avoid misunderstanding of your results. I believe with these minor edits, this manuscript can be ready for publication.

I have a few more substantial comments, with minor grammar and typographic edits at the bottom.

Line 478: "descends… indicating the deposition process" this is a statement without evidence. Deposition is a process at the surface that removes aerosols from the atmosphere; the process you are thinking of is "gravitational settling" sometimes called "sedimentation." For this case, you could say that this change in retrieved ALH reflects either gravitational settling or atmospheric subsidence, but the actual cause is almost certainly subsidence.

Thank you for your insight and suggestion. We believe that using the term "subsidence" would be the most accurate for our case.

GEMS measurements show that the dust plume is located at 4-5 km on March 27, descends to ~3 km on March 28, and remains at that altitude until March 29, consistent with EPIC and TROPOMI measurements. These daily changes in ALH reflects the atmospheric subsidence of dust aerosols during transport.

Figure 9 (and Figure 11): You have hourly GEMS retrievals, showing a descending plume, and a series of looks from EPIC and TROPOMI, which do not indicate any trend in aerosol layer height. You summarize this by stating that "Likely, multiple GEMS observations reveal clear hourly variations, whereas discerning diurnal changes is challenging with TROPOMI and EPIC." You need to make a clearer statement here—is the difference reflective of different retrieval sampling (TROPOMI and EPIC obviously have large gaps in the observations compared with GEMS), or is it due to differences in the retrieved quantities?

The original statement "Likely, multiple GEMS observations reveal clear hourly variations, whereas discerning diurnal changes is challenging with TROPOMI and EPIC." is written to explain one dust plume case on a March 28, 2021. In this particular case, TROPOMI and EPIC provide fewer observations compared to hourly GEMS observations. We have revised the statement as follows:

For this dust case, hourly GEMS observations reveal clear hourly variations, while characterizing diurnal changes from TROPOMI and EPIC is challenging due to their limited number of observations compared to GEMS.

Introduction Lines 60-70: I am not sure what this summary of Choi et al. 2021 is doing here. The finding of Choi et al. was that only with polarization-sensitive measurements did the DOFS increase to permit largely independent retrieval of aerosol layer height and thickness—still not enough information for actual multi-layer cases. As you describe, none of the sensors you are using have this polarization sensitivity, and so they retrieve only a single layer height parameter using an assumed layer thickness. I think this section could be both shorter and more clear.

We decided to remove this reference as it may not be directly relevant to our study.

Line 295-305: for comparisons where there is significant mean bias between the samples, the linear regression slope and intercept are not very robust; I recommend relying mostly on r2 and RMSE for this analysis.

We removed the analysis from slope and intercept and rewrote it as follows:

For dust cases, both TROPOMI and EPIC AOD show a positive bias compared to AERONET AOD, with values of 0.23 and 0.33 for TROPOMI and EPIC, respectively. In contrast to the dust cases, TROPOMI and EPIC AOD exhibit a negligible bias and smaller RMSE for smoke cases. Although TROPOMI and EPIC AOD do not provide retrievals for values less than 0.2, many AERONET AOD data points exist with values under this threshold, particularly in dust cases. This suggests that the surface reflectance used in the dust aerosol model from EPIC and TROPOMI may be underestimated, resulting in an overestimation in AOD retrieval.

Line 285-295: If I understand correctly, you are saying that AOD for smoke cases correlates equally well with AERONET at 688 and 443 nm, while for dust cases, the correlations are weaker, and the RMSE higher, at 688nm. This could be written more clearly.

Correct. We clarified the writing:

GEMS AOD at 443 nm exhibits a strong positive correlation with AERONET AOD at 440 nm, with correlation coefficients (R) of 0.9 for dust cases and 0.88 for smoke cases (Fig. 3a). At 680 nm, the correlation for smoke cases remains high at R = 0.84, indicating a similar level of agreement with the 443 nm measurements. However, for dust cases at 680 nm, the correlation decreases to R = 0.73, along with a 17 % increase in RMSE, indicating distinct retrieval accuracy of GEMS AOD at 443 and 680 nm for dust.

Abstract, line 30: I think you mean "overestimate ALH"?

It is meant to be AOD, but "km" was incorrectly written. This was corrected in newer version of the manuscript from the first revision.

Line 161 "eastern half" "western region"

Corrected.

Multiple places: I recommend describing EPIC sampling as "near-hourly" rather than "close-hourly"

We changed all "close-hourly" to "near-hourly".

Line 170: "The fact that the lower surface reflectance…motives us" => "Lower surface reflectance in the O2 B band vs the O2 A band suggests that O2 B band can be used to improve ALH retrievals using A band only."

Modified to:

Lower surface reflectance in O2 B band compared to O2 A band over land (Xu et al., 2019) suggests that O2 B band can be used to improve the ALH retrievals using O2 A band only.

Line 187: "this spectral range includes many trace gas absorption bands, and is also sensitive to aerosols and surface properties."

Corrected as follows:

TROPOMI on board the Copernicus Sentinel-5 Precursor satellite was launched in October 2017 to measure solar radiation reflected by Earth from UV to shortwave infrared (SWIR) bands. This spectral range includes many trace gas absorption bands, and is also sensitive to aerosols and surface properties.

Line 202: "Less valid retrievals" => "Fewer valid retrievals"

This part has been removed.

Line 320-321: "The inaccuracy of the GEMS dust aerosol model… has a notable impact"

Corrected grammar.