Response to Referee #1

Thank you for carefully reading the manuscript and providing useful suggestions to improve the paper. The replies to the referee comments are given below. The referee comments are in blue with our responses in black. The sentences in the manuscript are *italic*, with the modifications in the revised manuscript in red.

All the comments are considered, and the text are modified accordingly in the revised version. Some minor revisions are not shown in this document, but can be seen in the track change version of manuscript. References are updated as suggested.

Review of egusphere-2023-1945: "Monitoring biomass burning aerosol transport using CALIOP observations and reanalysis models: a Canadian wildfire event in 2019", by Shang et al.

The authors assess mass transport in the CAMS and MERRA-2 reanalysis data sets using measurements of wildfire smoke acquired by the CALIPSO lidar and by several ground-based lidars. The paper is well organized and the subject matter is appropriate for ACP.

Below I have attached an annotated version of the manuscript that contains several (mostly minor) comments and suggestions. I recommend publication once these have been addressed. Because I am well acquainted with lidar measurements, but less so with the model predictions, my comments focus almost exclusively on the handing and interpretation of the lidar data.

PAGE 5

why not 2D interpolation in both space and time?

We tested this and found that in our case, both methods produced similar results, but the linear temporal interpolation is computationally more expensive. Therefore, we decided to use the nearest time in the interpolation.

PAGE 6

see the "Software and model code" section at https://www.atmospheric-measurement-techniques.net/submission.html#assets. consider adding a "code availability" section to the paper that would include a link to this "custom Lagrangian trajectory model"

We have made an example script for the trajectory computations, and added the "code availability" section in the manuscript:

> *Code availability: A code example of the trajectory computations is available at: https://gist.github.com/anttilipp/29f2cb56d99a054e1aa0fc5bcc1d8622*

I highly recommend including the relevant equations from Shang et al., 2021 in this paper; e.g., combine equations 1 and 2 from the 2021 paper into a single equation in this paper. doing this will make life much easier on your readers.

We agree and have added the equation in the revised manuscript.

PAGE 7

revise for clarity

For the clarity, we modified from:

> *Note that there are about two and half hours difference between these two observations, the sampling smoke layer could be not the exact part of the plume.*

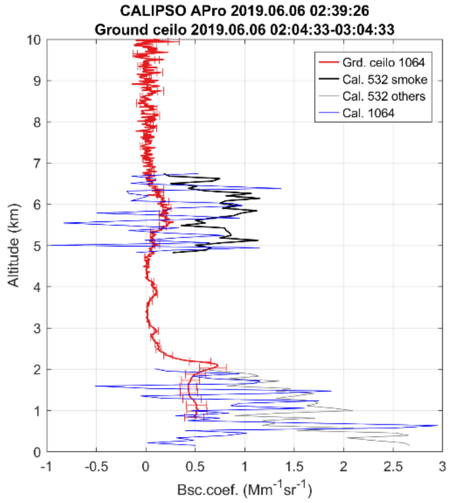
to:

Note that there is about two and half hours difference between these two observations, thus the observed smoke layers could be from different parts of the plume.

PAGE 8

did the authors compare the 1064 nm backscatter coefficients from the ceilometer to the 1064 nm backscatter coefficients reported in the CALIOP APro files? doing that might be useful in any examination of retrieval uncertainties introduce by using a fixed Angstrom coefficient.

The CALIOP 1064 backscatter coefficient (bsc) is much noisier than 532 bsc. especially for the optically thin lofted smoke layer after long-range transportation.



In this case, for the smoke layer between ~4.8-6.8 km, the mean values of CALIOP bsc 1064 is 0.19 ± 0.52 $Mm^{-1}sr^{-1}$, with a median value of 0.10 $Mm^{-1}sr^{-1}$, whereas the mean values of ceilometer bsc 1064 is 0.12 ± 0.07 $Mm^{-1}sr^{-1}$, with a median value of 0.11 $Mm^{-1}sr^{-1}$. The mean values of CALIOP bsc 532 is 0.70 ± 0.26 $Mm^{-1}sr^{-1}$, with a median value of 0.75 $Mm^{-1}sr^{-1}$. When converted CALIOP bsc 532 mean value to 1064 using Angstrom of 2.2, we have a value of 0.15, which is inside the error bar of mean value of CALIOP bsc 1064.
We have added an uncertainty discussion in section 2.4.1.

PAGE 8

please give the measurement time differences for the Granada, Barcelona, and Antikythera cases

The measurements times were given in fig. 2c, we added the text in the revised version for the clarity:
Granada:

The ceilometer-derived backscatter coefficients (1h time-average centered on CALIOP profile) of the smoke layer measured at 1064 nm were used to estimate the layer mass concentrations following the method presented in Sect 2.4.1.

Barcelona:

There is a 4min difference between the two measurements.

Antikythera:

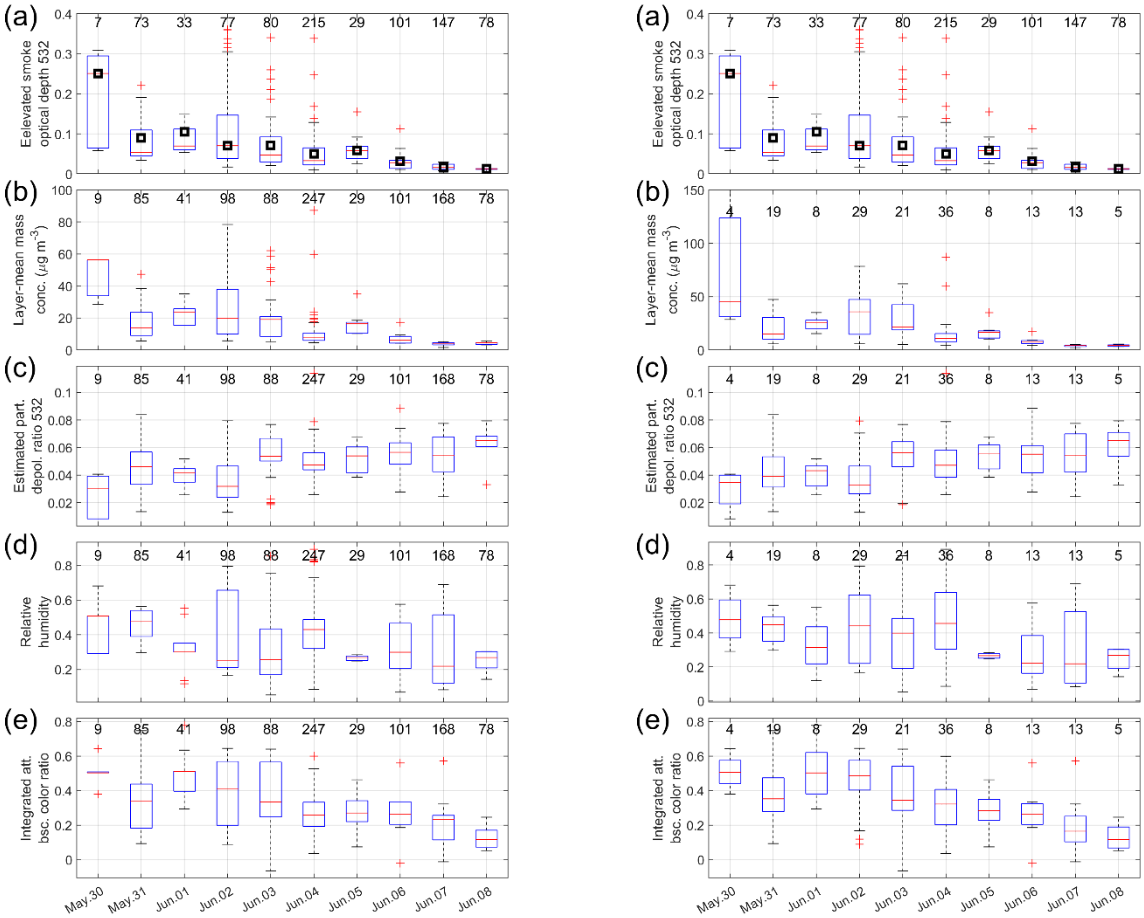
The backscatter coefficients derived from the ground-based lidar (1h time-average) and CALIOP show good consistency (Fig. 2, 4th row, panel c), with about half an hour time difference.

how many of these are unique detections? when searching for layers, the CALIOP detection scheme can average over as much as 80-km, so it's possible for 16 consecutive 5-km data records in the CALIOP level 2 products to report identical information for an aerosol layer detected at 80-km resolution (i.e., the results from one unique detection are replicated 16 times).

Thank you for the comment, this is a good point. We added such information in the revised version:

> *In total, 1336 smoke layers were selected, related to 1194 CALIOP 5km profiles. Considering the different horizontal averaging lengths applied in the CALIOP detection scheme, there are 259 unique detections, among which the layer numbers with horizontal averaging of 5km, 20km, and 80km are 75, 417, and 844, respectively.*

We have compared the statistics from the unique detections with the results in the manuscript, and found that the conclusions stay the same regardless of the sampling method. Here is a comparison using all layers (on the left, fig. 4 in the revised manuscript) and unique detections (on the right) for layers originating from 29 May. The total number of layers is 944, related to 840 5km-profiles, corresponding to 156 unique detections. Overlapping layers are found only in 52 5km-profiles (~6%), thus APro products (instead of ALay) were used to calculate the column AODs (sub-figures (a), the total number is 5km-profiles number). Sub-figures (b)-(e) show minor differences but similar tendencies (the total number on the left figures is the layer number, whereas the total number on right figures is the unique detection number).



see section 2.3 of Thorsen et al., 2013; https://doi.org/10.1002/jgrd.50691. the CALIOP ALay files represent the output of the "nested multigrid averaging scheme" used by the CALIOP layer detection algorithm (see Vaughan et al., 2009; https://doi.org/10.1175/2009JTECHA1228.1). Consequently, regions of differing backscatter intensity within a single vertically and horizontally contiguous layer can be detected at different averaging resolutions. And because the detections at different averaging levels are reported separate and distinct layers, the layer heights and optical properties reported in the ALay files can be notoriously difficult to interpret. So, for readers (and reviewers) to properly interpret the optical depth statistics provided here, the authors need to describe how they have reconciled any apparent vertical overlap of layers reported in the 5-km ALay columns. Failing to account for the apparent vertical overlap will introduce biases into statistical aggregates derived from the ALay parameters.

I note too that 5-km column AODs for smoke are readily obtained from the CALIOP APro products by masking out "not smoke" range bins then integrating. doing this is much more accurate and reliable than adding the layer AODs reported in the ALay products.

Thank you very much for the comment and suggestion. You are right, our original calculation overestimated the columnar AOD due to the overlapping ALay products. We have revised the data processing, and carefully checked the overlapping issue. For overlapping layers, APro products are used instead.
In the whole data set of 1195 5km-profiles, there were overlapping issues on 62 profiles (~5%).
Using the updated data products, we have updated fig. 4, fig 6, fig. 7, fig. 8, as well as the corresponding values in the text. We find that the value differences are small, thus our conclusion remains the same.

The CALIPSO Level 2 5km aerosol layer (ALay) products are used here. Note that the optical depths were summed up (denoted as AOD here) in case there are several smoke layers detected in the same profile (Fig. 4a). In the ALay product, aerosol layers are detected using a "nested multigrid averaging scheme" (Vaughan et al., 2009), which may produce vertically overlapping layers. In such cases, CALIPSO Level 2 5km aerosol profile (APro) products were used instead to calculate the AODs. The same method was also applied later for the column mass calculation to avoid layer overlapping issue.

PAGE 9

overlapping error bars in all cases; in three cases, the CALIOP means lie within the ground-based error bars and vice versa.

Thank you for the suggestion, we have modified the text:

As was discussed in Sect. 3.1 and shown in Fig. 2, mass retrievals from CALIOP ~~agreed well~~ have overlapping error bars in all cases with ground-based lidar/ceilometer retrievals; in three cases, the CALIOP means lie within the ground-based error bars and vice versa.

PAGE 10

also see Thorsen et al., https://doi.org/10.1002/2017GL074521

Furthermore, some aerosol layers could not be fully detected by CALIOP due to the weak signal to noise ratio as discussed in Sect. 3.1 and in Thorsen et al. (2017).

as expected

The column AOD include all aerosol types, and column mass considered here only include BC and OC types. we have modified the text for the clarity:

When compared with CALIOP products, the correlation coefficients are quite similar as the ones for AODs, demonstrating that the smoke aerosols are dominant in the column.

PAGE 12

see https://www-calipso.larc.nasa.gov/; "NASA and CNES agreed to end the CALIPSO science mission on August 1, 2023". so the observational gap has already begun.

There will be an observational gap once CALIPSO ends its life, and during the absence of the upcoming space-borne lidar missions.

to

As CALIPSO science mission ended on August 1, 2023, there currently is an observational gap during the absence of the upcoming space-borne lidar missions.

Response to Referee #2

Thank you for carefully reading the manuscript and providing useful suggestions to improve the paper. The replies to the referee comments are given below. The referee comments are in blue with our responses in black. The sentences in the manuscript are *italic*, with the modifications in the revised manuscript in red.

This is a very interesting study that attempts to investigate an aerosol transport event from different perspectives – perhaps, from almost all possible angles: ground, space, and models.

My main concern is that, in general, the authors should have addressed the possible uncertainties in detail, such as temporal/spatial collocation mismatches, lidar ratio assumptions, and assumptions in the trajectory model. Perhaps it would be helpful to have dedicated sections discussing these uncertainties. I don't mean to suggest that the authors are unaware of these uncertainties. In fact, the later part of the manuscript mentions that future lidar missions/studies would be important. However, that message could be strengthened by a dedicated section on uncertainties.

Thank you for the comments, we have added discussions about the uncertainties in the revised version in section 2.4.1:

> *The conversion factor at 532 nm of 0.16 ± 0.01 $\times 10^{-6}$ m for the fresh and medium-fresh smoke (i.e. less than 2 days), or 0.13 ± 0.01 $\times 10^{-6}$ m for aged smoke, and a particle density of 1.3 g cm^{-3}, were used for the biomass burning particles (Ansmann et al. 2021). Following Ansmann et al. (2021), we assume uncertainties of 10 % and 20 % in the conversion factor and smoke mass density. Using ground-based multi-wavelength lidar measurements, the lidar ratio at 532 nm and the backscatter-related Ångström exponent between 532 and 1064 nm were derived as 71 ± 5 sr and 2.2 ± 0.3, respectively, for the smoke plumes during the same wildfire event as in this study (Shang et al. 2021). For the backscatter coefficient retrievals we used relative uncertainties of 10 %, 15 %, and 25 % for ground-based lidar, ceilometer, and spaceborne lidar, respectively. These values were taken from Shang et al. (2021) and Ansmann et al. (2021). The Ångström value of 2.2 was applied to convert the ceilometer measured backscatter coefficients at 1064 nm to 532 nm (Eq. 2), resulting a relative uncertainty of 24 % on the converted backscatter coefficients. This study employed a lidar ratio at 532 nm of 70 sr, which was the value used for the aerosol subtype of "elevated smoke" in CALIOP version 4 (Kim et al. 2018). The lidar ratios at 532 nm for ground-based lidars are measured with a typical relative uncertainty of about 20 %, which can also be assumed for the 532 nm CALIOP lidar ratio for elevated smoke (the uncertainty is 70 ±16 sr in CALIPSO V4 lidar data). More details of the uncertainties in the CALIPSO products can be found in Young et al. (2013, 2018). Applying the law of error propagation to Eq.1 with the above-mentioned uncertainties, we expected an overall uncertainty in the mass concentration estimates of 32 % for ground-based lidar, and 40 % for ceilometer and CALIOP.*

We have also changed the figure 2 in the revised version, by adding some error bars on sub-figs (c) (e) (f).

The assumption in the trajectory model was that the air parcels, aerosols included, were moved only by the winds (all 3 components). The trajectory model is described in Section 2.3.3. We acknowledge the uncertainties due to trajectory model, wind data and spatial/temporal collocation. However, as we started new trajectories from the fire locations frequently and from multiple altitudes, and as the trajectory and smoke observations match reasonably well also after a relatively long transport we think our trajectories and collocation strategy is working well enough for this study. We have added the following sentence to the manuscript (section 2.4.2):

> *The uncertainties due to trajectory computations, wind data, and temporal and spatial collocation causes uncertainty to the estimates of the dominant air mass pathway. However, we estimate these uncertainties to be small and not to significantly affect the results of our study.*

Additionally, I do not quite understand how you derived these trajectory frequencies. I suggest refining Section 2.4.2.

We have modified the text for the clarity:

> *6480 trajectories were generated considering originated from ~~the same~~ a single day (0 to 24 h), from all initial heights (0 to 7.5 km), and from both wildfire sub-regions (~~18 intial spots~~ 9 initial spots in each, Fig. 1). ~~The trajectory frequencies were used here to determine the dominant air mass pathway.~~ The dominant air mass pathway was determined by the trajectory frequencies, which were calculated via the bivariate bin counts in two steps: The latitude and longitude trajectory frequencies were calculated based on 1° x 1° pixels, whereas the altitude and time trajectory frequencies were calculated based on 500 m x 1 h pixels. The pixels with ~~a probability~~ an occurrence frequency above the median value were selected, referring to the most possible air mass transportation. Only the trajectories included in the predefined pixels were kept (less than 10 % of the total trajectories). These screened trajectories were then used to define 4-dimensional hypercubes, with a doubled resolution as previously used for the frequency pixels, considering the model uncertainties. Each hypercube has eight values of four variables (i.e. the edge values of latitude, longitude, altitude and time). Next, the CALIOP-derived smoke layers (after the quality control, see Sect. 2.2.1) were automatically selected using these 4-D hypercubes to ensure that they are on the dominant air mass pathway.*

While I am not an expert in writing, I feel that this manuscript may not be easily readable. Considering that it's a good study with nice ideas on combining different perspectives, I wish the writing also matched that high standard.

Thank you for the comment. We have revised the manuscript and made modifications to the text to improve the clarity of the manuscript.

Corrections:

Line 49: "…vastness of the boreal region; it cannot be covered with advanced ground-based…" (added a ';', removed the second 'based')

Thank you for the suggestions, we have corrected the text:

> *Due to the vastness of the boreal region, it cannot be covered with advanced ground-based ~~based~~ observations or flight campaigns.*

Line 79: works -> studies

We have corrected the word.