Larry Thomason RC#1

We would like to thank Mr. Larry Thomason for his thorough and detailed review. Our responses to each comment are provided below. Reviewer comments are highlighted in blue, and our responses are presented in regular font. The text in quotation marks and shown in italics denotes the text that features the submitting manuscript.

Response to General Comments

This manuscript presents a potentially interesting analysis of the standard CALIPSO stratospheric products over the European continent with potential connections to climate and other change phenomena. The author's methodology could be applied to other geographical areas and thus is readily extensible. However, this manuscript is poorly written and at times incomprehensible and it needs substantial revision before it would be in a suitable form for publication. Some of the errors are egregious such as suggesting that stratospheric aerosol is only present occasionally. I am sure the authors know that is correct, but it is suggestive of the lack of care made in putting this paper together.

It isn't clear to me what is accomplished by including a longitudinal component to the data grid. One, none of the analyses strong a particularly persistent difference in the longitudinal direction. Episodically, there are such differences but, given the temporal box size, it is not clear what those differences mean. To me those differences are rather muddled and of limited value. A possible exception is for the northern most bins as it is well known that PSC occurrences have a longitudinal dependence associated with orographic affects. In addition, breaking December out by itself seems to be a solution to what is not a problem. There's nothing important that happens at the year boundary so DJF crossing a year boundary isn't such a problem. This solution is unique.

Referencing is generally not well done. Many of the references in the list are missing basic information like journal, doi, etc. This made checking whether the references difficult, particularly since there are clear instances where references are wrong or I was unfamiliar with those cited. There's also a tendency to reference off-brand material where more canonical and well-known references exist. This normally isn't a big deal but at times it's kind of odd and led me to refer to the reference list which led me to find deficiencies noted above.

Once more we appreciate the effort by Mr. Thomason as well Referee #2. The manuscript has undergone a thorough revision to address all the issues raised. We improved readability, streamlined the text, and corrected issues related to the formatting and the references. Regarding the primary scientific concerns raised in the general comments, we fully agree that stratospheric aerosols are a permanent feature of the atmosphere rather than an occasional phenomenon. However, the focus of this study is on episodic stratospheric aerosol events and their characteristics, rather than on the background stratospheric aerosol layer. This feature is underscored now throughout the submitting manuscript.

Concerning the comment on the longitudinal component of the grid, Mr. Thomason raises a valid point regarding the lack of persistent longitudinal differences across most bins. Therefore, we removed the longitudinal grids and aggregated the data zonally.

Furthermore, we appreciate the reviewer's perspective regarding the seasonal grouping (December versus DJF). While separating December was an approach tailored to our specific dataset, the section discussing seasonal contributions has now been removed from the submitting manuscript in accordance with the reviewer's suggestion.

Finally, we apologize for the confusion created by the mistakes in the References section. Now, the reference list has been thoroughly revised and corrected. Missing DOIs, journal names, and other publication details have been added and verified.

Response to specific comments

Line 26, 'they have different'
This is now corrected.

Line 27, This statement is incorrect. Aerosol associated with the Mt. Pinatubo eruption persisted in the stratosphere for nearly 10 years.
We thank the reviewer for the comment. We have revised the manuscript to accurately reflect the multi-year residence time of volcanic aerosols in the stratosphere. The following sentence is added in the submitting manuscript, in lines 37-38 of the revised manuscript: *"The injection of volcanic sulfur dioxide (SO_2) leads to the formation of sulfate (sulfuric acid) aerosols that reside for multiple years in the stratosphere* (Toohey et al., 2025)*."*

Line 28, This statement is correct, but sulfuric acid aerosol (commonly but erroneously called sulfate aerosol) but should also note that sulfuric acid aerosol absorbs significantly in the infrared and is often associated with heating in the stratosphere.
We fully agree that this distinction is not acknowledged in the submitted manuscript although as pointed out sulfuric acid aerosol is commonly considered as the same as the sulfate aerosol. The following sentence is added in the submitting manuscript:
lines 37-41 of the revised manuscript: "*The injection of volcanic sulfur dioxide (SO_2) leads to the formation of sulfate (sulfuric acid) aerosols that reside for multiple years in the stratosphere* (Toohey et al., 2025). These particles primarily scatter incoming solar radiation, resulting in a net negative radiative forcing at the top-of-the-atmosphere (TOA) and a consequent cooling at the surface (Aubry et al., 2021; Schallock et al., 2023), their significant absorption in the infrared spectrum simultaneously leads to localized heating within the lower stratosphere (Niemeier and Schmidt, 2017)."

Line 29, I am pretty sure that a paper by Vernier (with Deshler and Knepp) showed the persistence of ash in the stratosphere following the Nabro eruption many weeks/months following that eruption. So maybe up to several months might be more accurate than days or weeks.
We updated the text of the revised manuscript adding also the relevant references of Sawamura, et. al. (2012) and Vernier et. al. (2016). The following sentence is added in the lines 41-43 of the revised manuscript.
"In contrast, volcanic ash particles are relatively large and sediment after some days to weeks generally resulting in only minor climatic significance compared to the sulfate aerosols (Boucher, 2015; Schallock et al., 2023; Sawamura et 2012; Vernier et al., 2016)."

Line 53, the authors should mention the PSC classification work by Pitts and others.
We now also mention PSC classification by Pitts et. al. (2009) in line 65 of the revised manuscript.

Line 61, Halofsky et al., must be the wrong reference because it says nothing on the topic of this sentence.

The reference is correct, however, we recognize that the original wording could be interpreted. To improve clarity and better reflect the scope of the cited publication, we have removed the word "stratospheric". The sentence has been revised and now reads (lines 72–73 of the revised manuscript): *"In recent years, extreme wildfire patterns have become more prevalent across the North Hemisphere (NH), leading to notable changes in the composition and dynamics of aerosols* (Halofsky et al., 2020)*."*

Line 62, it would be more circumspect to describe the Australian wildfire outbreak as comparable to a moderate volcanic eruption (which remains poorly defined) but not refer to it using "can now rival" suggesting something new which is unproven at this point.

We agree that this sentence needs revision. We changed the sentence in the revised manuscript, Lines 73-77: *"Massive smoke injections from extreme pyroCb outbreaks, such as the 2019–2020 Australian wildfires can produce stratospheric perturbations comparable to volcanic eruptions* (Fromm et al., 2021; Peterson et al., 2018)*, in terms of optical depth and longevity, significantly impacting the stratospheric aerosol load and its subsequent climatic feedback* (Hirsch and Koren, 2021)*."*

Line 94, The description of how the SAOD is computed is a mess. Why not say that SAOD is computed by integrating the aaer profile from the tropopause to some appropriate altitude.

The description of the retrieval of sAOD is reintroduced. The description is given in Lines 107-108 of the revised manuscript: *"and the stratospheric aerosol optical depth (sAOD), derived by integrating each a_{aer} profile from the tropopause up to 30 km."*

Line 95-98, more description of the uncertainty in the extinction coefficient values particularly considering the change in composition size distribution of aerosol is really important to understanding what follows. I am pretty sure that the lidar ratio can vary in a single profile as well as from profile to profile. Data without a discussion of uncertainty is pointless.

We thank the reviewer for the comment. We agree that a robust discussion of uncertainty, particularly regarding the variability associated with the lidar ratio, is crucial for contextualizing the retrieved extinction coefficients. CALIOP, being an elastic backscatter lidar, cannot independently measure backscatter and extinction (Winker et al., 2010). Therefore, it relies on the prescribed extinction-to-backscatter ratio, also known as lidar ratio, to solve the nonlinear lidar equation. This is an essential but consequential simplification because the lidar depends on the aerosol size, shape, and chemical composition, all of which can vary continuously within a layer and differ from layer to layer (Floutsi et al., 2023).

The optical properties retrieval process works from the top of the atmosphere downward. Specifically, the uncertainty in the backscatter coefficients increases exponentially with range through a feature at a rate determined by the uncertainty in the lidar ratio used in the retrieval and the uncertainties in the overlying backscatter estimates (Young et al., 2013). This means that lidar ratio errors do not produce a constant bias, instead they grow as the retrieval penetrates deeper into and below the layer. For low optical depths, the relative error in retrieved optical depth approximates the relative error in the lidar ratio itself. At larger optical depths, errors grow significantly, and the relative error in the lidar ratio constitutes a lower bound for the relative error in the retrieved optical depth. The predefined lidar ratio is therefore among the largest sources of uncertainty in optical retrievals, with increasing impact lower in the atmosphere (Young et al., 2013). Critically, lidar ratio selection errors can cause large AOD biases of either sign, and these biases can cancel out or enhance biases from missed layer detections (Burton et al., 2013). This bidirectional nature distinguishes lidar ratio errors from calibration errors, which tend to bias systematically in one direction, making lidar ratio-induced biases particularly difficult to correct.

The lidar ratio selection scheme used in CALIPSO's Level 2 aerosol retrievals has been shaped fundamentally by the Aerosol Robotic Network (AERONET). Omar et al. (2005) performed a cluster analysis of a multiyear dataset to identify six globally representative aerosol types, whose derived optical models underpin the CALIPSO automated aerosol classification and lidar ratio selection algorithm (Omar et al., 2009), with the goal of constraining the uncertainty in the lidar ratio to no more than 30%. This target reflects the natural spread of true lidar ratios within a single aerosol class, but the real atmospheric situation is more complex. A single profile may contain multiple vertically stacked layers of different aerosol types, each assigned a different constant lidar ratio, or a single layer whose composition varies with altitude but to which CALIOP nonetheless applies one fixed value. Misclassification of aerosol type higher in the atmosphere then produces errors that systematically propagate into all underlying layers.

In the Version 4 algorithm (Kim et al., 2018), lidar ratios for each aerosol subtype were updated based on measurements, modelling, and renewed cluster analysis of the multiyear AERONET dataset. For example, the uncertainty on the elevated smoke lidar ratio was reduced from 70±28 sr in V3 to 70±16 sr in V4, reflecting improved constraints from field campaigns and independent retrievals. V4 also introduced four dedicated stratospheric subtypes — polar stratospheric aerosol, volcanic ash, sulfate/other, and smoke — subsequently refined by disentangling sulfate/other into sulfate and unclassified subtypes among other refinements (Tackett et al., 2023).

For stratospheric aerosol specifically, when a layer is optically thin and isolated — as is typically the case in the background and post-volcanic stratosphere — uncertainties are dominated by the lidar ratio uncertainty alone and are approximately equal to it, with limited error propagation from overlying features (Young et al., 2013).

The stratospheric aerosol in Kim et al. (2018) is classified primarily under the sulfate/other or sulfate subtype, a near-spherical, chemically well-characterised aerosol with a 532 nm lidar ratio of 50 sr and an associated uncertainty of approximately 20–30%. At the small optical depth characteristic of the stratosphere (typically AOD < 0.05 outside of major volcanic episodes), this places the expected relative uncertainty in retrieved extinction coefficients in the range of 20–30%, substantially lower than the 50–60% typical for tropospheric smoke layers or the near-100% values near the surface in thick, complex scenes (Young et al., 2013). This more constrained uncertainty range is an inherent advantage of focusing on stratospheric aerosol.

The revised text in the manuscript (lines 122-137) reads as follows:

" As an elastic lidar, CALIOP retrieves aerosol extinction using predefined, constant lidar ratios for specific aerosol subtypes—derived from empirical data, modeling, and AERONET datasets (Omar et al., 2005; Young et al., 2018; Winker et al., 2010)*. Because actual lidar ratios vary continuously based on aerosol size, shape, and composition* (Floutsi et al., 2023)*, this fixed assignment introduces significant uncertainty. Operating top-down, CALIOP's retrieval errors propagate exponentially with range. Unlike systematic calibration errors, lidar ratio errors are bidirectional—capable of either canceling or amplifying biases from missed layer detections, making them difficult to correct* (Burton et al., 2013; Young et al., 2013)*. While the automated classification algorithm targets a 30% uncertainty constraint based on six global aerosol clusters* (Omar et al., 2009)*, real-world complexities like vertically stacked layers or intra-layer composition changes cause misclassifications that systematically corrupt all underlying retrievals* (Young et al., 2013)

For stratospheric aerosols, however, this uncertainty is highly constrained. In these typically thin, isolated layers, error propagation from overlying features is minimal, meaning the relative error in retrieved optical *depth directly approximates the error in the lidar ratio itself* (Young et al., 2013). *The stratospheric aerosol evaluated by Kim et al., (2018) is primarily classified as the sulfate/other subtype—a near-spherical, well-characterized aerosol with a 532 nm lidar ratio of 50 sr and an associated 20–30% uncertainty. At typical stratospheric optical depths (AOD < 0.05), this limits the expected relative uncertainty in retrieved extinction coefficients to just 20–30%. This is an inherent advantage over the troposphere, where complex, thick scenes can drive uncertainties up to 50–100%* (Young et al., 2013).*"*

Line 106 (Figure 4), Using latitude and longitude to define the boxes means that they change size from north to south and aren't really square at all. Since the distribution of points is unlikely uniform within a box (more points the further north in a box particularly at high latitudes due to orbital considerations), unless some weighting is applied the effective central latitude of a box is north of the latitude mid-point. Has this been checked or is this accounted for?

This spatial sampling asymmetry is accounted for in our data aggregation methodology. Because CALIPSO is a polar-orbiting satellite, orbital track convergence inherently creates a higher density of profiles at higher latitudes within any given zone. To prevent this non-uniform data distribution from biasing the zonal averages, each orbit-mean profile is weighted by the number of individual 5-km profiles contributing to it. This weighting scheme ensures that every individual 5-km measurement profile carries equal weight. The resulting zonal values are mathematically equivalent to a direct arithmetic mean of all profiles within that latitude band, neutralizing the effects of both shrinking grid-box areas and increased high-latitude orbital density.

The following sentences are added in the revised paper to address the issue raised, in lines 230-243:

"Because CALIPSO operates in a sun-synchronous, near-polar orbit, its ground tracks converge toward higher latitudes, resulting in a greater sampling density in the northern zones. To account for this spatial sampling asymmetry, the orbit-mean profiles, $\bar{x}_{orbit}$, were aggregated into monthly and annual means, weighted by the number of the 5-km profiles contributing to each orbit average as:

$$\bar{x}_{monthly} = \frac{\sum_{orbits} N_{orbit}\bar{x}_{orbit}}{\sum_{orbits} N_{orbit}} \quad (1)$$

This approach ensures that the resulting values are equivalent to a direct arithmetic mean of all individual profiles within a given latitude zone and prevents regional statistics from being disproportionately influenced by variations in satellite sampling density. To characterize the overall stratospheric state, both monthly and annual averages were computed alongside their respective aerosol type contributions for each zone. Finally, a pan-European assessment was performed to represent the whole study area. "

Line 110-121, The description of what data is used looks ok but some discussion of what happens if you use even subtly different criteria like -40 or 0 with the confidence that something is aerosol. Is it possible that different event types require different filters?

We now expand on this by adding a paragraph explaining better the decision on the selected criteria thresholds. The revised manuscript in Lines 157-164 reads: *"In our analysis, we retained only layers with CAD scores between −100 and −20. By setting the lower bound at −20, we exclude features with low-confidence classifications (scores between 0 and −20) from clouds or polar stratospheric clouds (PSCs)* (Kar et al., 2019; Wainker et al., 2013; Tackett et al., 2022). *The choice of a −20 CAD threshold represents a calculated balance between data retention and feature contamination. Relaxing this threshold toward 0 would include lower-confidence aerosol features, but at high altitudes or during high-latitude winters, it* *drastically increases the risk of PSC contamination being misclassified as aerosol. Conversely, tightening the threshold to −40 or −50 would ensure a higher-confidence aerosol sample, but at the cost of filtering out highly dense pyroCb smoke plumes or thick volcanic ash layers."*

Line 126-132, I am not a fan of the typing mechanism used in Calipso since it suggests that aerosol is always one type or another and neglects the likelihood that aerosol is a mixture of compositions. When the signal is large, one can 'type aerosol" confidently but in more subtle situations it could be a trap for understanding their radiative effects.

We fully agree that the aerosol type assignment is a consequential simplification as described earlier (comment lines 95-98). In reality, aerosol layers frequently contain mixtures of aerosol types, and their composition may vary vertically or horizontally. The aerosol classification scheme simply assigns the most probable aerosol subtype based on the observed optical signatures and other auxiliary data such as underlying surface and layer altitude, for the purpose of selecting a lidar ratio. CALIPSO accounts for dust mixtures in the troposphere with two specific aerosol subtypes. The aerosol classification algorithm outputs dusty marine and polluted dust besides desert dust (Kim et al., 2018). Polluted dust intends to represent mixtures of dust and smoke and dust and polluted continental aerosols. Whereas dusty marine accounts for dust and marine aerosols. CALIPSO is an elastic backscatter lidar, it cannot independently retrieve aerosol mixing state or composition, and therefore cannot represent mixtures explicitly. In strong-signal scenes this subtype assignment is generally robust, but in tenuous layers the discrete typing can introduce additional uncertainty in the inferred radiative properties.

The Version 4 algorithm partly mitigates this issue by introducing additional aerosol subtypes (Kim et al., 2018) and by refining stratospheric aerosol categories (version 4.5 by (Tackett et al., 2023)), but the fundamental limitation remains: the lidar ratio is applied as a constant within each layer, regardless of internal variability or mixing.

Line 156-160, I don't know about the spike filter. Not being a lidar expert, I thought you had to be really careful computing averages since the noise is Poisson. Is this approach ok? Why is it relevant only to SAOD and not all the parameters?

We thank the reviewer for raising this point. The lidar signal noise is governed by Poisson. Because of this, standard random noise manifests as both positive and negative fluctuations centered around a true mean, which should normally be preserved during averaging.

However, the "spike filter" we implemented does not target standard random noise fluctuations. Instead, it is designed to isolate and eliminate rare, extreme, unphysical artifact pixels (often referred to as "hot pixels" or misclassified features) that have backscatter values orders of magnitude higher than both the stratospheric background and standard noise envelopes. Leaving these unphysical spikes in the dataset severely skews arithmetic means during spatial or temporal aggregation. By establishing a conservative 3σ threshold (0.005 $km^{-1}sr^{-1}$), we successfully isolated and removed these distinct artifacts—which accounted for only 0.8% of the data—without altering the underlying Poisson noise distribution of the valid data.

Regarding the second point, these artifacts distort all subsequent optical parameters, not just sAOD, this is now better stated inside the text. The original phrasing was misleading. We have removed the specific mention of sAOD in this sentence and revised the text to clarify that this filtering process provides a cleaner, more robust baseline for all stratospheric parameters analyzed throughout the study. The text below features the submitting paper in Lines 210-211: *"This process was applied across the entire vertical extent of each nighttime stratospheric profile, as these anomalies distort all derived optical parameters."*

and lines 217-224: "The clipping threshold was established at 0.005 $km^{-1}sr^{-1}$, corresponding to the $\mu+3\sigma$ of the distribution. While random lidar signal noise fundamentally follows a Poisson distribution—producing symmetric fluctuations around a true mean when signal-to-noise ratios are sufficient—the targeted artifacts are isolated, extreme positive outliers that do not resemble physical features or standard noise distributions. Applying a strict 3σ threshold successfully isolated these unphysical anomalies, resulting in the removal of 1,063 data points (0.8% of the total dataset). The impact of this refinement is evident in the shift of the dataset's central tendency: the initial mean of 0.0004 $km^{-1}sr^{-1}$ was reduced to a clipped mean of 0.0003 $km^{-1}sr^{-1}$, providing a more robust baseline for all the subsequent parameter analyses."

Line 166-170, There's nothing particularly physical about year boundaries, so why is aligning with a calendar year important? I am not sure I find this discussion very convincing.

We agree that the alignment with calendar years does not have a particular physical significance. In the revised manuscript, however, we have removed the section discussing seasonal contributions, including the text corresponding to Lines 166–170 of the initial manuscript. As a result, the issue raised by the reviewer is no longer relevant to the current version of the manuscript.

Line 174, I would probably delete "robust" as it falls into the same category as "very", it weakens rather than strengthens an argument.

We deleted "robust" as suggested.

Line 195-199, Why repeat, with fewer references, the material in Table 1?

We have revised the text in Lines 266–289 and removed the references from this paragraph since they are already fully detailed in Table 1.

Line 204, I am not sure you can say this. Yes, the signature of volcanic aerosol from tropical events can be significantly weaker than one in northern midlatitudes. However, it can't be said that they never have impacts and can't be confused with output from other events as the atmosphere effectively acts as a time filter. For instance, Pinatubo made an early appearance in the northern hemisphere in the lower stratosphere, but it was much later in the year before it arrived in all its glory.

We agree that tropical eruptions could impact northern mid-latitudes via large-scale stratospheric transport, and that over longer timescales, atmospheric mixing and delayed transport can make accurate attribution even more confusing. Our intention was not to imply that tropical events have no mid-latitude impacts, but rather to explain that our event selection (Table 1) prioritized eruptions with more rapid, coherent, and highly distinct transport pathways to Europe, allowing for immediate and clearer attribution to each event.

To prevent any misunderstanding and to correctly reflect the nature of tropical plume transport, we have revised lines as follows, in lines 277-281 of the revised manuscript: *"While tropical plumes could potentially transport to northern mid-latitudes via large-scale atmospheric circulation, this transport is highly dispersive and delayed, making the immediate, discrete attribution of specific aerosol layers more challenging at European latitudes* (Vernier et al., 2011)*. This study prioritizes events where the transport pathways and aerosol optical signatures were most clearly and rapidly distinguishable."*

Line 221-232, A similar problem as Table 1, repeating material in the table in the text with a little new material. Too much “notable” and “significant”. These are qualitative words that don’t really mean anything.

We have revised the text in Lines 298–315 and removed the references from this paragraph since they are already fully detailed in Table 1.

Line 233, I don’t think this section heading makes any sense. Maybe “Observations of the vertical distribution of stratospheric aerosol”

Thank you. We changed the title to “Observations of Stratospheric Aerosol Occurrence and Vertical Structure”.

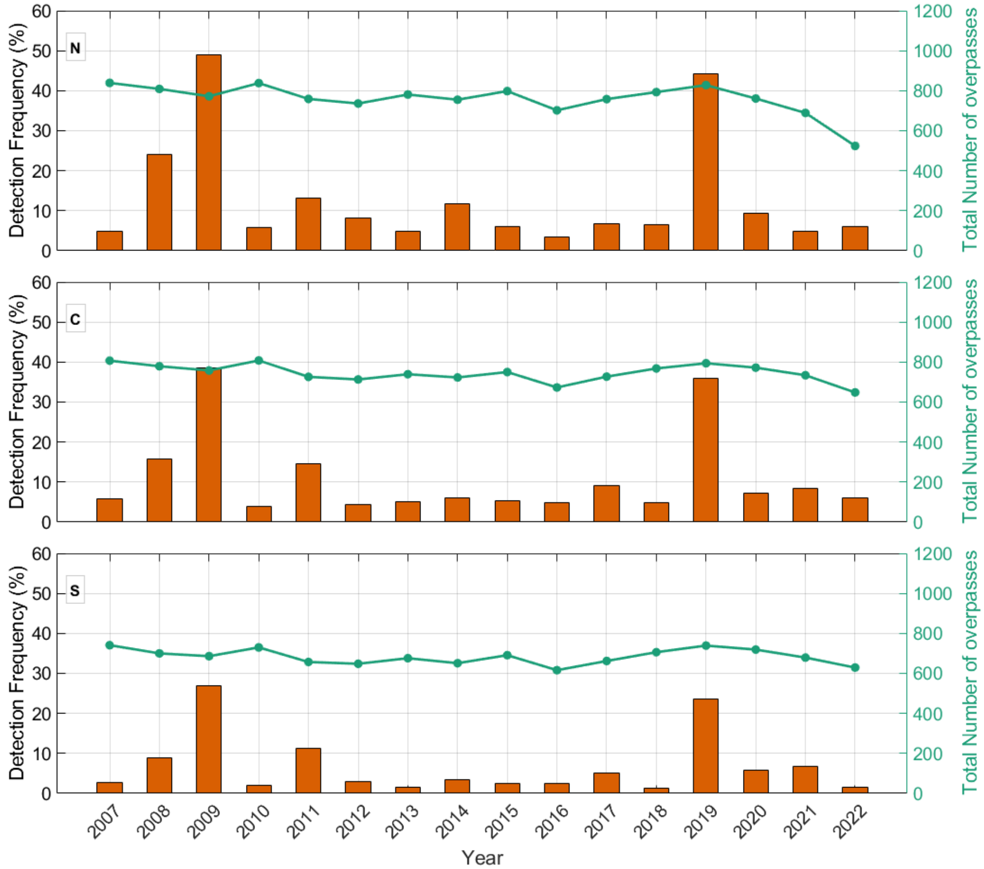
Line 234-235. ALL the profiles contain stratospheric aerosol. I assume you mean some sort of measurements containing enhanced aerosol rather than the background like what appears in Thomason et al 2021 and 2023. This continues through the following discussion and must be corrected.

We thank the reviewer for pointing out this. Our original phrasing was inaccurate. The background stratospheric aerosols are permanently present across all profiles. We intended to refer to the number of profiles containing detected, enhanced stratospheric aerosol load rather than the baseline background state described in Thomason et al. (2021, 2023). Following the reviewer’s recommendation, we have revised the text and the caption of Figure 4 throughout this section.

Line 248-247, Figure 4 shows absolute numbers which can be distorted when the bins contain different numbers. Relative numbers would be more interesting though again, area differences between the boxes is also important and should be considered. I suspect it would make the frequency at high latitudes even more pronounced. That wouldn’t be a surprise but still good to see. It would also be interesting if there’s any difference in the fraction of events tossed out between bins. Also how does the timing of events affect these frequencies? Most events do not conveniently occur at a seasonal boundary so the frequency of enhanced aerosol from those events may occur predominantly in one year or season or be split between them.

We thank the reviewer for this. We have addressed each point step-by-step and updated both the figure and the text accordingly:

As suggested, we updated Figure 4 to show the relative percentage of profiles with enhanced aerosols (orange bars) alongside the total overpasses (green lines). We agree that the differences in geographic surface area between latitude zones are crucial. Because grid cells contract toward the poles, the high-latitude frequency becomes even more pronounced when considering relative density. We have added the updated figure and text in lines 323-357 of the revised manuscript:



*"**Figure 4: Zonal variations in nighttime stratospheric aerosol observations (2007–2022). Orange bars indicate the percentage of profiles containing enhanced stratospheric aerosols (left axis), while green lines mark the total number of nighttime overpasses (right axis) for the North, Central, and South zones.***

CALIOP stratospheric observations across Europe show distinct spatial and temporal variations when aggregated into the three latitudinal zones. The most pronounced peaks occurred in 2009 and 2019, with all three latitudinal zones registering their highest detection frequencies. The north-to-south gradient is evident across the study period. For instance, the N zone showed, as expected, the highest values, reaching a maximum detection frequency of 48.96% (378 overpasses) in 2009 and 44.20% (366 overpasses) in 2019. The C zone generally exhibited intermediate values, maintaining a long-term average detection frequency of 10.98% and peaking in 2009 (38.52%) and 2019 (35.89%). In contrast, the S zone consistently showed the lowest overall metrics, averaging a 6.79% detection frequency across the 16-year timeline. However, it presented the same peak years, showing spikes in 2009 (26.82%, 184 overpasses) and 2019 (23.68%, 175 overpasses).

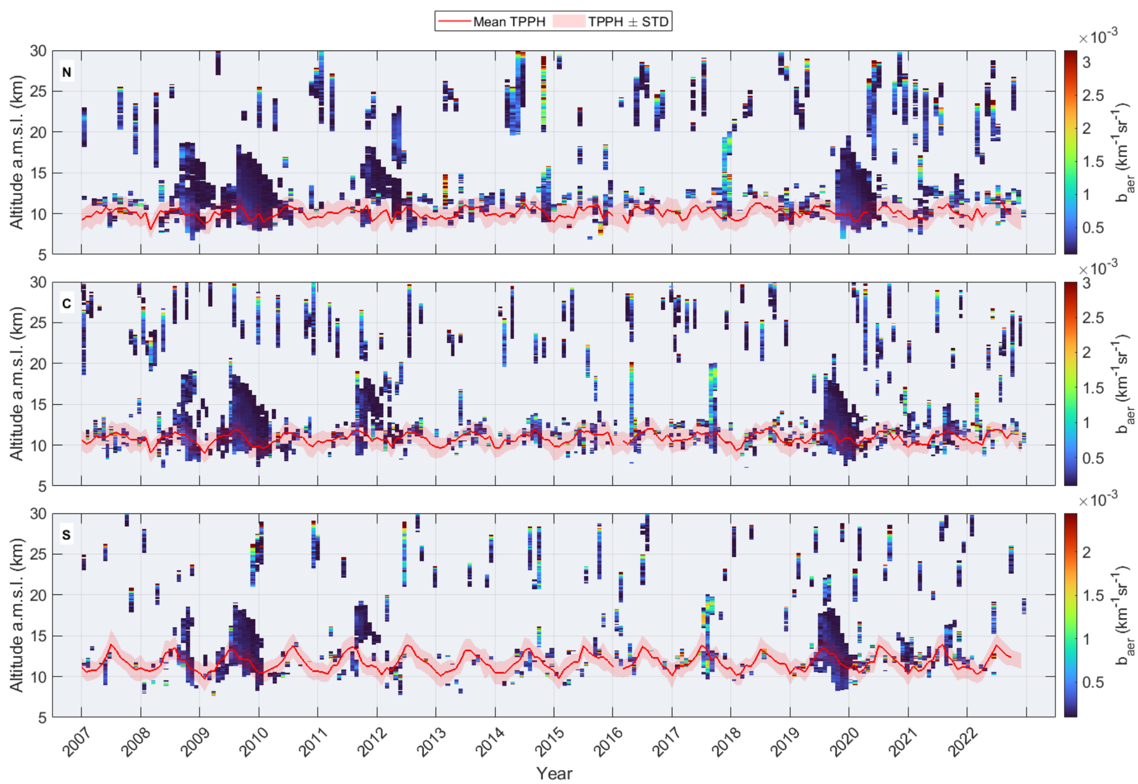
Conversely, multiple years showed widespread low points, with all three zones reporting sharp declines during unperturbed stratospheric periods. Frequencies dipped significantly in 2016 across both the N and C zones, dropping to 3.42% (24 overpasses) and 4.90% (33 overpasses), respectively. The S zone reached its lowest value during these periods, with detection frequencies below 1.5% in 2013 (1.63%), 2018 (1.42%), and 2022 (1.43%). Nevertheless, the temporal distribution of these frequencies is inherently influenced by the timing of individual events. Because each event rarely aligns with calendar or seasonal boundaries, a single event's aerosol plume can be split across consecutive bins. This time-splitting effect can smooth out *the observed frequencies, potentially underrepresenting the immediate seasonal peak of an eruption while artificially prolonging its apparent timeline.".*

Indeed, volcanic eruptions and wildfire events occur independently of calendar or seasonal boundaries. Because stratospheric aerosols stay in the atmosphere for a long time, an event happening late in a season or year will inevitably distribute its optical impact across multiple time periods, often showing peak frequencies in the following year or subsequent seasons. To address this, we have added a discussion paragraph clarifying how to interpret these prolonged aerosol signals, so they are not mistaken for new, separate injections.

Moreover, while we did not explicitly track the exact fraction of discarded events between individual bins, all zones were subjected to the exact same rigorous quality control protocols. We acknowledge that regional variations in cloud cover could slightly alter data retention across zones, but we expect any resulting impact on the broad latitudinal gradient to be minimal.

Line 250-1, Figure 5. The time scale on this figure is unnecessarily complicated. Please use a linear time scale. It might be interesting to see a contour plot of the enhanced aerosol extinction coefficient or backscatter. I know it would have gaps but still be interesting. I wonder if the very isolated events are noise particularly at higher altitudes attaching magnitudes would clarify that.

In the revised manuscript, Figure 5 has been removed and replaced with a contour plot of the backscatter coefficient using a linear time scale, which provides a clearer representation of the temporal and vertical evolution of enhanced stratospheric aerosol layers.



"Figure 5: Vertical distribution of zonal stratospheric aerosol backscatter coefficient (b_{aer}) from CALIPSO nighttime observations through the years 2007 to 2022 across three latitude bands: Northern Zone (N), Central Zone (N), and Southern Zone (S). Overlaid red lines track the monthly mean tropopause height (TPPH), enveloped by a shaded band representing standard deviation (STD)."

The corresponding discussion has been revised and expanded to reflect the updated results (lines 366–388 of the revised manuscript).

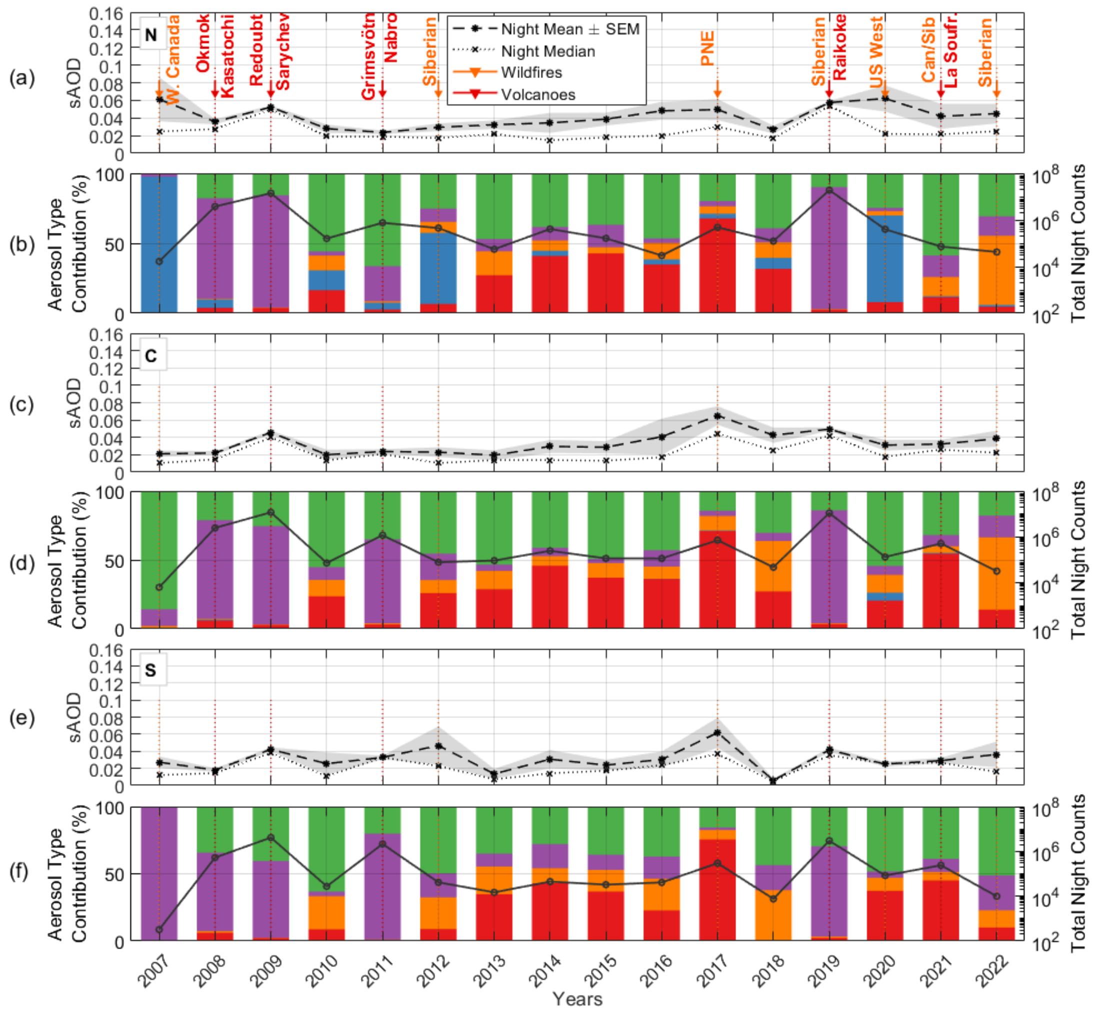
Regarding the possibility that isolated features at higher altitudes may represent noise, the contour representation allows the magnitude and vertical extent of the backscatter coefficient to be directly assessed. The observed high-altitude features are characterized by enhanced backscatter values extending across multiple altitude bins and/or consecutive time periods, indicating coherent aerosol layers rather than isolated random retrieval artefacts. The revised figure therefore facilitates a more robust distinction between genuine aerosol enhancements and potential noise.

Line 273-275, I don't see how Figure 6 shows anything about aerosol loading. I assume that the solid line (which is not described in the figure caption) is total number of enhanced aerosol measurements.

We thank the reviewer for pointing out this issue. We agree that Figure 6 did not provide information on aerosol loading, and the use of the term "loading" in the original text was therefore inappropriate.

In response to this comment, we have substantially revised this part of the manuscript. Figures 6 and 7 have been replaced by a new Figure 5, which combines information on stratospheric aerosol optical depth (sAOD) and aerosol-type classification. The new figure presents the annual evolution of the mean and median sAOD values, together with the relative contribution of the different aerosol classes (Smoke, PSA, Volcanic Ash, Sulfate, and Unclassified) for each of the three latitude zones. Major wildfire and volcanic events are also indicated along the time axis to facilitate interpretation of the observed variability.

Accordingly, Section 3.3 has been completely revised and renamed "Annual Evolution of Stratospheric Aerosol Loading and Aerosol Types". The associated text and figure caption have been updated to clearly describe all plotted quantities and to improve the interpretation of the long-term aerosol record. Please see the revised Figure 5 and the corresponding discussion in Section 3.3 (Lines 421–498 of the revised manuscript).



"Figure 6: (a, c, e panels): Time series (2007–2022) of the annual sAOD mean (continuous black line) and median (dotted black line) during nighttime observations. Shaded areas correspond to the uncertainty of the mean value (SEM). Major wildfires and volcanoes are marked along the timelines with red and orange color. (b, d, f panels): Aerosol type contribution (stacked bars, %) — Smoke, PSA, Volcanic Ash, Sulfate, and Unclassified— to the total number of classified profiles, the total stratospheric nighttime CALIPSO profile counts per year (black line) or each of the three European latitude zones: Northern Zone (N), Central Zone (C), and Southern Zone (S)."

Line 281 There is a whole section of text missing above this point (possibly hiding below the figure).

Yes, a small part of the text was hidden behind the figure. This is now corrected in the revised manuscript.

Line 281-282 There is a reference to the difference between day and night counts but I am under the impression that the authors only use nighttime measurements. This seems unrelated to the remainder of the paragraph.

Correct, we are using only nighttime data. This reference is now removed.

Line 300, Figure 7. I assume that the authors are using tropopause height from CALIPSO which are model based and probably no better than +/- 0.5 to 1.0 km. How does the proximity of the enhanced aerosol to the tropopause affect the magnitude of the perturbations? Have the authors compared their optical depth values with those from other instruments?

Regarding the tropopause height (TPPH) and its inherent uncertainties, we have expanded our text in Lines 108–117 of the submitting manuscript to explicitly acknowledge these limitations:

"The tropopause height (TPPH) from CALIPSO APro L2 product, used for this study, is obtained from the Modern-Era Retrospective analysis for Research Applications, version 2 (MERRA-2) meteorological reanalysis product (Gelaro et al., 2017). TPPH uncertainty in the MERRA-2 reanalyses primarily stems from high-altitude biases, poor vertical resolution in the upper troposphere/lower stratosphere (UTLS), and discrepancies in tracking sharp thermal boundaries (Weyland et al., 2025; Manney et al., 2014). The main effect of this uncertainty on the sAOD calculation is concentrated in the UTLS region; an overestimation or underestimation of the tropopause height by MERRA-2 can artificially exclude lower stratospheric aerosol layers or mistakenly include upper tropospheric feature enhancements in the integration. Given the high concentration of transported plumes near the UTLS boundary over Europe, this can introduce a minor localized variance in the absolute sAOD values, though it does not significantly alter the long-term relative trend (Xian and Homeyer, 2019). "

Furthermore, the impact of the uncertainty in the MERRA-2 derived tropopause height may affect the classification of aerosol layers located close to the tropopause, potentially resulting in the assignment of some stratospheric aerosols to the troposphere, or vice versa. In such cases, the derived stratospheric aerosol perturbations may be slightly reduced or enhanced. Nevertheless, because the major aerosol perturbations discussed here are typically associated with vertically extensive layers that extend well into the lower stratosphere, small variations in tropopause height are unlikely to substantially affect the overall magnitude of the calculated perturbations.

We compared the CALIPSO sAOD values against the GloSSAC climatology (see below, Fig. X). However, we chose not to include this comparison in the manuscript. Our CALIPSO dataset is regionally restricted to the -10° to 30° longitude sector, whereas GloSSAC provides zonal-mean values averaged over all longitudes (0–360°), which dilute these regional peaks. Given this fundamental spatial mismatch, a direct quantitative comparison would be misleading and fall outside the scope of this study.

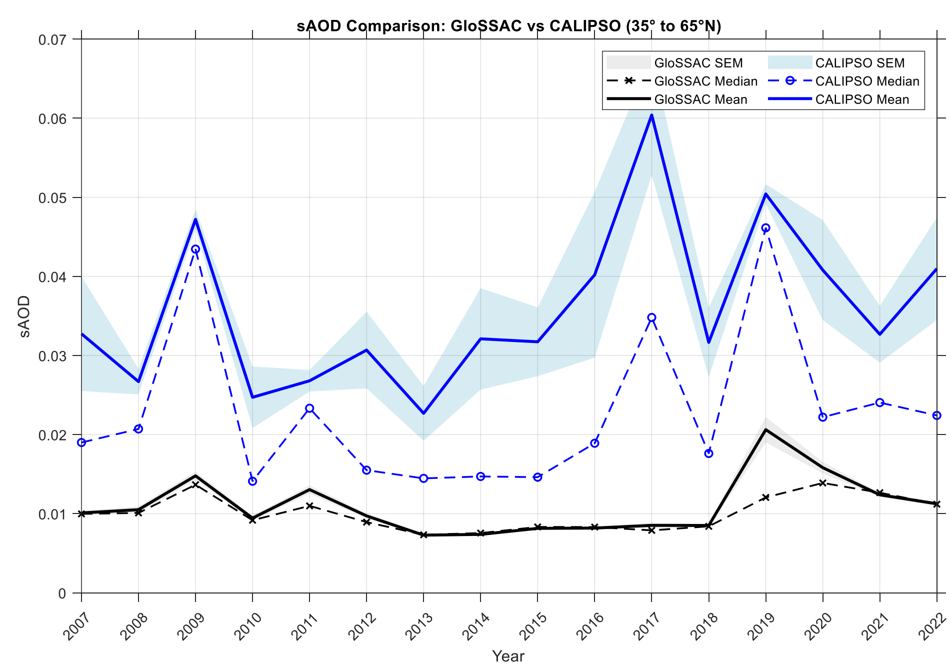
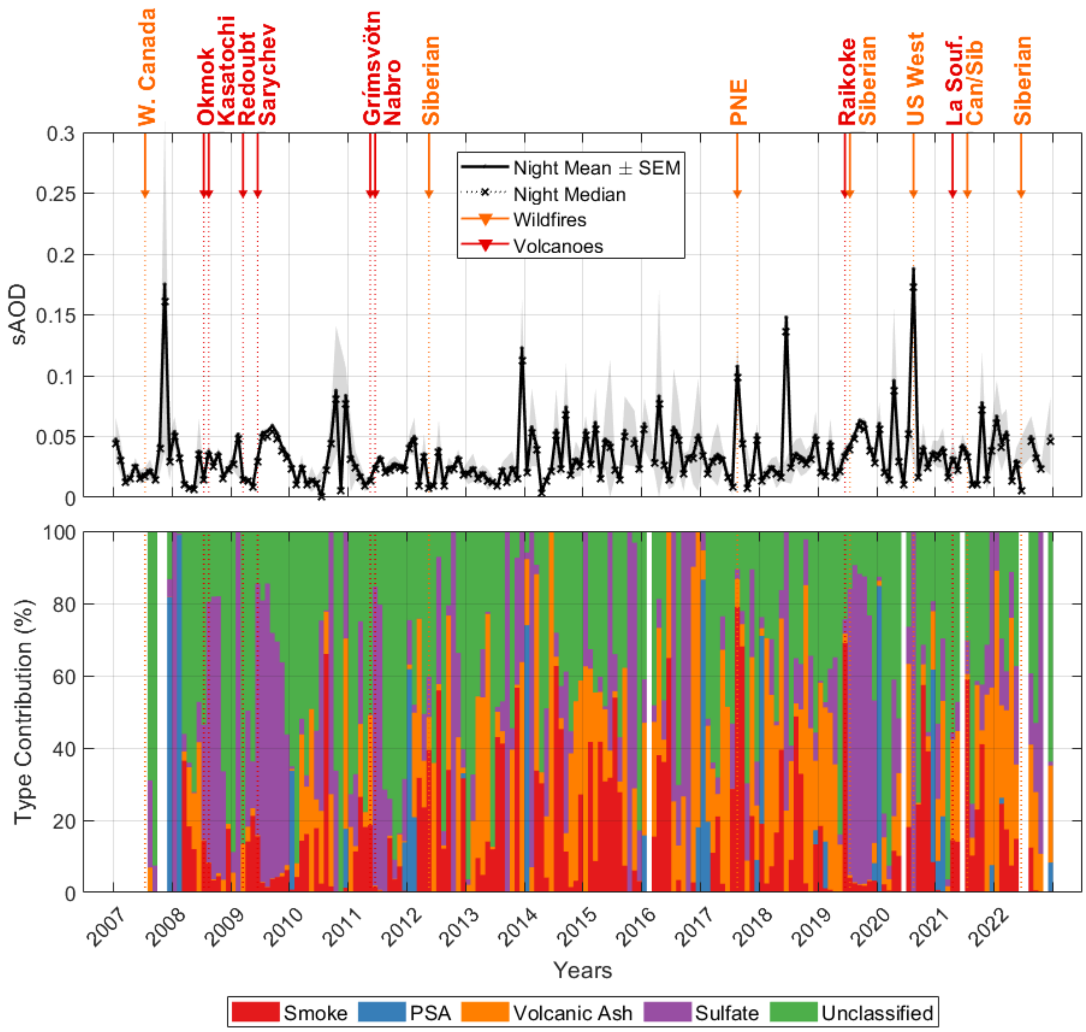


Figure X. Comparison between CALIPSO nighttime sAOD over the European sector (10° W–30° E) and the GloSSAC zonal-mean climatology.

The comparison is shown for illustrative purposes only, as the datasets represent different spatial domains. While CALIPSO captures regional variability over Europe, GloSSAC provides longitude-averaged (0–360°) values, which smooth and dilute regional aerosol enhancements associated with episodic volcanic eruptions and wildfire-smoke events. Consequently, direct quantitative comparisons should be interpreted with caution.

Line 333, Figure 8 (all of section 3.4). I have no idea what this section shows. It would help if the authors showed how these parameters were computed. I can't think of a reason for this material to be included in the paper.
We apologize for the confusion that this section created. Given the reviewer's concerns regarding the presentation and relevance of the results, we have removed Section 3.4 and the corresponding Figure 8 from the revised manuscript. This change has helped streamline the manuscript and improve its overall focus and readability.

Line 360-379, Figure 9. Annual averages, for what is mostly a collection of quite small events, results in a plot (top of Figure 9) that is unremarkable and, given the reported standard deviations (in the figure), it is arguable that none of the events are statically significant. Maybe showing the results in higher temporal resolution would help. The means are based on the distribution of individual optical depth values and perhaps the uncertainty in the mean rather than the population STD would be a better statistic to show. On the other hand, in the text following, the optical depth uncertainties are much smaller, so is there a disconnect between what the figure shows and the text discusses? I do not like the split axis for the lower two figures, why not use a log scale? I also do not like the counts as latitude sampling distortion could be corrupting the meaning of the relative fractions. The authors make increase attributions to their various sources that are not at all clear from the figure. That argument is very weak.
We agree that the annual averaging used in the original Figure 9 obscured much of the temporal variability associated with episodic stratospheric aerosol events. We also acknowledge the reviewer's concerns regarding the interpretation of the uncertainties, the split-axis representation, and the attribution of aerosol sources based on the original figure. Therefore, we have completely revised this section. The original Figure 9 has been replaced by a new Figure 7, and the section has been rewritten and renamed "3.4 Monthly Evolution of Stratospheric Aerosol Loading and Aerosol Composition over Europe". Instead of annual averages, the new figure presents the monthly evolution of stratospheric aerosol optical depth (sAOD) from 2007 to 2022, allowing individual volcanic and wildfire-related perturbations to be more clearly identified. The figure now shows monthly mean and median sAOD values, with the shaded region representing the standard error of the mean (SEM), as suggested by the reviewer. Major volcanic eruptions and wildfire smoke injection events are explicitly marked along the time axis to facilitate interpretation of the observed variability. In addition, the lower panel has been redesigned to show the relative contribution of the CALIPSO aerosol subtypes through time, replacing the previous presentation. The associated discussion has been fully revised, and all interpretations have been re-evaluated in light of the new figure. The new Figure 7 and the revised discussion are in Section 3.4 (Lines 584–635) of the revised manuscript.



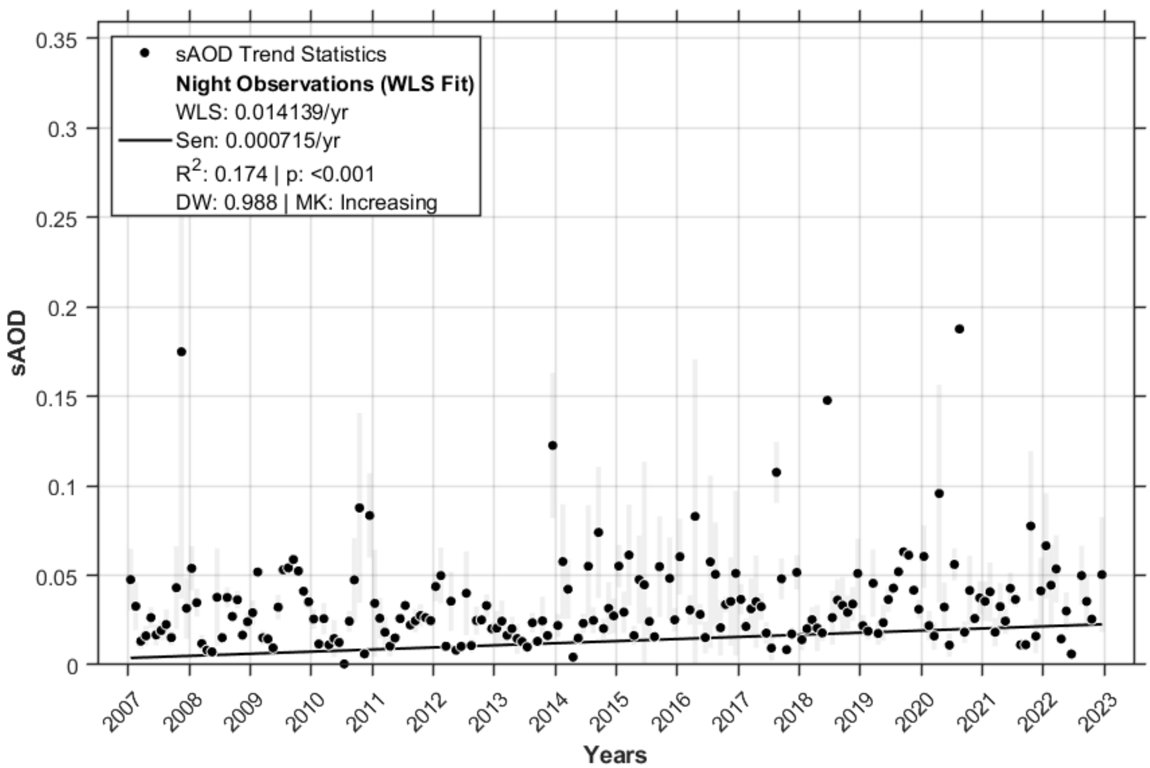
"Figure 7: Nighttime stratospheric aerosol evolution over the European region (2007–2022). (Top) Monthly mean and median sAOD observations, with the shaded area representing the SEM. Vertical arrows indicate major volcanic eruptions (red) and wildfire smoke injections (orange). (Bottom) Relative percentage contribution of CALIPSO stratospheric aerosol subtypes. Vertical dotted lines correspond to the chronological markers of the annotated events."

Line 380-437, Similar to the discussion regarding Figure 8, all the material associated with Figure 10 is incomprehensible to me. The authors must see something that makes these figures interesting, but they are not conveying it to the reader. And one can also include the following subsections (3.5.1-3.5.4) where I do not see why these discussions are interesting. They are, at best, readings of the figures and provide little insight.

We agree that the material presented in Figure 10 and the associated subsections (3.5.1–3.5.4) did not effectively convey the intended scientific message and largely consisted of a descriptive reading of the figures. To improve the clarity and conciseness of the manuscript, we have removed Figure 10 together with Section 3.5 and its subsections from the revised version. The manuscript has been reorganized accordingly to focus on the main findings and their interpretation.

Line 438, trend work is interesting but maybe a reference to past trend work like Hofmann's papers and Deshler's paper. I am not familiar with all the statistical tests they mention, references would be helpful and short explanations of how they work.

We have substantially revised this section, which is now presented as Section 3.5, "Long-Term sAOD Trends over Europe". Rather than annual values, the trend analysis is now based on the complete monthly nighttime sAOD record (2007–2022), providing a more robust basis for the statistical assessment. The original Figure 12 has been replaced by a new Figure 8 showing the monthly mean sAOD time series and trend statistics. Following the reviewer's suggestion, we have also incorporated discussion of previous stratospheric aerosol trend studies, including the works of Hofmann and Deshler, to place our results in the context of the existing literature. These references and their relevance to our findings are discussed in Lines 742–753 of the revised manuscript However, the entire trend-analysis section has been rewritten (Lines 700–765).



"Figure 8: Time series of the nighttime mean monthly sAOD (2007–2022) across Europe. The shaded areas denote the SEM. The line represents the weighted least-squares (WLS; w=N/s^2, where s the uncertainty of the mean) fit and the inset panel summarizes the trend statistics, including the WLS slope, Theil–Sen slope estimate, Durbin–Watson (DW) statistic, and Mann–Kendall (MK) trend test."

Line 466-468. All volcanic and fire events are isolated or episodic events and their effects are always transient. That's their nature. However, they can cluster together randomly which can cause temporary sustained elevated levels (and the reverse at times too). To me, that is what their second discussion highlights and it is as likely to disappear in the next few years as continue. If the authors wish suggest that there's a fundamental change beginning in 2016 associated with a durable geophysical sources, they need to back that up better than they have so far.

We fully agree with the reviewer's perspective. Volcanic eruptions and extreme wildfire injections are inherently episodic, and their atmospheric impacts are fundamentally transient. It was not our intention to claim a permanent geophysical shift or a durable new baseline regime. As the reviewer rightly points out, the post-2016 upward inflection in the CUSUM analysis is the result of a dense temporal clustering of these independent, episodic events, which could just as easily dissipate back to lower baseline levels in subsequent years. In the revised manuscript, we have removed the discussion based on the CUSUM analysis that could be interpreted as implying a sustained change in the aerosol baseline. However, we have retained the broader interpretation that periods of elevated stratospheric aerosol loading can arise from the temporal clustering of independent volcanic eruptions and wildfire events. We now emphasize that the observed variability reflects the occurrence and clustering of episodic events rather than evidence for a persistent long-term shift in the stratospheric aerosol system.

References

Aubry, T. J., Staunton-Sykes, J., Marshall, L. R., Haywood, J., Abraham, N. L., and Schmidt, A.: Climate change modulates the stratospheric volcanic sulfate aerosol lifecycle and radiative forcing from tropical eruptions, Nat. Commun., 12, 4708, https://doi.org/10.1038/s41467-021-24943-7, 2021.

Boucher, O.: Atmospheric aerosols, in: Atmospheric Aerosols: Properties and Climate Impacts, Springer Netherlands, Dordrecht, 9–24, https://doi.org/10.1007/978-94-017-9649-1_2, 2015.

Burton, S. P., Ferrare, R. A., Vaughan, M. A., Omar, A. H., Rogers, R. R., Hostetler, C. A., and Hair, J. W.: Aerosol classification from airborne HSRL and comparisons with the CALIPSO vertical feature mask, Atmos. Meas. Tech., 6, 1397–1412, https://doi.org/10.5194/amt-6-1397-2013, 2013.

Floutsi, A. A., Baars, H., Engelmann, R., Althausen, D., Ansmann, A., Bohlmann, S., Heese, B., Hofer, J., Kanitz, T., Haarig, M., Ohneiser, K., Radenz, M., Seifert, P., Skupin, A., Yin, Z., Abdullaev, S. F., Komppula, M., Filioglou, M., Giannakaki, E., Stachlewska, I. S., Janicka, L., Bortoli, D., Marinou, E., Amiridis, V., Gialitaki, A., Mamouri, R.-E., Barja, B., and Wandinger, U.: DeLiAn – a growing collection of depolarization ratio, lidar ratio and Ångström exponent for different aerosol types and mixtures from ground-based lidar observations, Atmos. Meas. Tech., 16, 2353–2379, https://doi.org/10.5194/amt-16-2353-2023, 2023.

Fromm, M. D., Kablick, G. P., Peterson, D. A., Kahn, R. A., Flower, V. J. B., and Seftor, C. J.: Quantifying the source term and uniqueness of the August 12, 2017 Pacific Northwest PyroCb event, J. Geophys. Res. Atmos., 126, e2021JD034928, https://doi.org/10.1029/2021JD034928, 2021.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Clim., 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Halofsky, J. E., Peterson, D. L., and Harvey, B. J.: Changing wildfire, changing forests: the effects of climate change on fire regimes and vegetation in the Pacific Northwest, USA, Fire Ecol., 16, 4, https://doi.org/10.1186/s42408-019-0062-8, 2020.

Hirsch, E. and Koren, I.: Record-breaking aerosol levels explained by smoke injection into the stratosphere, Science, 371, 1269–1274, https://doi.org/10.1126/science.abe1415, 2021.

Kar, J., Lee, K.-P., Vaughan, M. A., Tackett, J. L., Trepte, C. R., Winker, D. M., Lucker, P. L., and Getzewich, B. J.: CALIPSO level 3 stratospheric aerosol profile product: version 1.00 algorithm description and initial assessment, Atmos. Meas. Tech., 12, 6173–6191, https://doi.org/10.5194/amt-12-6173-2019, 2019.

Kim, M. H., Omar, A. H., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., Hu, Y., Liu, Z., Poole, L. R., Pitts, M. C., Kar, J., and Magill, B. E.: The CALIPSO version 4 automated aerosol classification and lidar ratio selection algorithm, Atmos. Meas. Tech., 11, 6107–6135, https://doi.org/10.5194/amt-11-6107-2018, 2018.

Manney, G. L., Hegglin, M. I., Daffer, W. H., Schwartz, M. J., Santee, M. L., and Pawson, S.: Climatology of upper tropospheric–lower stratospheric (UTLS) jets and tropopauses in MERRA, J. Clim., 27, 3248–3271, https://doi.org/10.1175/JCLI-D-13-00243.1, 2014.

Niemeier, U. and Schmidt, H.: Changing transport processes in the stratosphere by radiative heating of sulfate aerosols, Atmos. Chem. Phys., 17, 14871–14886, https://doi.org/10.5194/acp-17-14871-2017, 2017.

Omar, A. H., Won, J., Winker, D. M., Yoon, S., Dubovik, O., and McCormick, M. P.: Development of global aerosol models using cluster analysis of Aerosol Robotic Network (AERONET) measurements, J. Geophys. Res. Atmos., 110, D10S14, https://doi.org/10.1029/2004JD004874, 2005.

Omar, A. H., Winker, D. M., Kittaka, C., Vaughan, M. A., Liu, Z., Hu, Y., Trepte, C. R., Rogers, R. R., Ferrare, R. A., Lee, K. P., Kuehn, R. E., and Hostetler, C. A.: The CALIPSO automated aerosol classification and lidar ratio selection algorithm, J. Atmos. Ocean. Technol., 26, 1994–2014, https://doi.org/10.1175/2009JTECHA1231.1, 2009.

Peterson, D. A., Campbell, J. R., Hyer, E. J., Fromm, M. D., Kablick, G. P., Cossuth, J. H., and DeLand, M. T.: Wildfire-driven thunderstorms cause a volcano-like stratospheric injection of smoke, npj Clim. Atmos. Sci., 1, 30, https://doi.org/10.1038/s41612-018-0039-3, 2018.

Sawamura, P., Vernier, J.-P., Barnes, J. E., Berkoff, T. A., Welton, E. J., Alados-Arboledas, L., Navas-Guzmán, F., Pappalardo, G., Mona, L., and Madonna, F.: Stratospheric AOD after the 2011 eruption of Nabro volcano measured by lidars over the Northern Hemisphere, Environ. Res. Lett., 7, 034013, https://doi.org/10.1088/1748-9326/7/3/034013, 2012.

Schallock, J., Brühl, C., Bingen, C., Höpfner, M., Rieger, L., and Lelieveld, J.: Reconstructing volcanic radiative forcing since 1990, using a comprehensive emission inventory and spatially resolved sulfur injections from satellite data in a chemistry-climate model, Atmos. Chem. Phys., 23, 1169–1207, https://doi.org/10.5194/acp-23-1169-2023, 2023.

Tackett, J. L., Kar, J., Vaughan, M. A., Getzewich, B. J., Kim, M.-H., Vernier, J.-P., Omar, A. H., Magill, B. E., Pitts, M. C., and Winker, D. M.: The CALIPSO version 4.5 stratospheric aerosol subtyping algorithm, Atmos. Meas. Tech., 16, 745–768, https://doi.org/10.5194/amt-16-745-2023, 2023.

Toohey, M., Jia, Y., Khanal, S., and Tegtmeier, S.: Stratospheric residence time and the lifetime of volcanic stratospheric aerosols, Atmos. Chem. Phys., 25, 3821–3839, https://doi.org/10.5194/acp-25-3821-2025, 2025.

Vernier, J.-P., Thomason, L. W., Pommereau, J.-P., Bourassa, A., Pelon, J., Garnier, A., Hauchecorne, A., Blanot, L., Trepte, C., Degenstein, D., and Vargas, F.: Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade, Geophys. Res. Lett., 38, L12807, https://doi.org/10.1029/2011GL047563, 2011.

Vernier, J.-P., Fairlie, T. D., Deshler, T., Natarajan, M., Knepp, T., Foster, K., Wienhold, F. G., Bedka, K. M., Thomason, L., and Trepte, C.: In situ and space-based observations of the Kelud volcanic plume: the persistence of ash in the lower stratosphere, J. Geophys. Res. Atmos., 121, 11104–11118, https://doi.org/10.1002/2016JD025344, 2016.

Winker, D. M., Tackett, J. L., Getzewich, B. J., Liu, Z., Vaughan, M. A., and Rogers, R. R.: The global 3-D distribution of tropospheric aerosols as characterized by CALIOP, Atmos. Chem. Phys., 13, 3345–3361, https://doi.org/10.5194/acp-13-3345-2013, 2013.

Weyland, F., Hoor, P., Kunkel, D., Birner, T., Plöger, F., and Turhal, K.: Long-term changes in the thermodynamic structure of the lowermost stratosphere inferred from reanalysis data, Atmos. Chem. Phys., 25, 1227–1252, https://doi.org/10.5194/acp-25-1227-2025, 2025.

Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Flamant, P., Fu, Q., Hoff, R. M., Kittaka, C., Kubar, T. L., Le Treut, H., McCormick, M. P., Mégie, G., Poole, L., Powell, K., Trepte, C., Vaughan, M. A., and Wielicki, B. A.: The CALIPSO mission: a global 3D view of aerosols and clouds, Bull. Am. Meteorol. Soc., 91, 1211–1230, https://doi.org/10.1175/2010BAMS3009.1, 2010.

Xian, T. and Homeyer, C. R.: Global tropopause altitudes in radiosondes and reanalyses, Atmos. Chem. Phys., 19, 5661–5678, https://doi.org/10.5194/acp-19-5661-2019, 2019.

Young, S. A., Vaughan, M. A., Kuehn, R. E., and Winker, D. M.: The retrieval of profiles of particulate extinction from Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data: uncertainty and error sensitivity analyses, J. Atmos. Ocean. Technol., 30, 395–428, https://doi.org/10.1175/JTECH-D-12-00046.1, 2013.

Young, S. A., Vaughan, M. A., Garnier, A., Tackett, J. L., Lambeth, J. D., and Powell, K. A.: Extinction and optical depth retrievals for CALIPSO's version 4 data release, Atmos. Meas. Tech., 11, 5701–5727, https://doi.org/10.5194/amt-11-5701-2018, 2018.