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.

Some specific comments of varying importance:

Line 26, 'they **have** different'

Line 27, This statement is incorrect. Aerosol associated with the Mt. Pinatubo eruption persisted in the stratosphere for nearly 10 years.

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.

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.

Line 53, the authors should mention the PSC classification work by Pitts and others.

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

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.

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

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.

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?

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?

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.

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?

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.

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

Line 195-199, Why repeat, with fewer references, the material 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.

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.

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

Line 234-5. 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.

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.

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.

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.

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

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.

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?

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.

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.

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.

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.

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.