Response to Referee 1 Comments

We would first like to express our sincere thanks and appreciation to Referee 1 for their thorough and detailed review. The comments identified flaws and unclear points in the article, providing an excellent opportunity to improve its overall quality.

Our responses follow the structure of the review document and are divided into three sections: 1) responses to major comments, 2) responses to specific substantive issues, and 3) responses to minor issues. Referee comments are written in black and authors answers are in blue.

Major comments:

Point 1: The term "ozone mini-hole" has a specific meaning – it refers to a transient natural synoptic-scale phenomenon that arises, mainly in midlatitudes, through dynamical and transport processes (a combination of uplift and horizontal advection of ozone-poor air). Total column ozone decreases rapidly during a mini-hole event but returns to its initial levels as the weather systems pass. Ozone mini-holes are unrelated to photochemical processes. Thus, the region of low ozone described and attributed to Hunga in this paper is NOT an "ozone mini-hole". This wording needs to be changed throughout the manuscript, including the title.

Response 1: Indeed, the ozone anomalies discussed in the manuscript are primarily chemical, not dynamical. To clarify this, we now refer to them as "Transient Ozone Depletion Event" and have updated the title and manuscript accordingly.

Point 2: The authors acknowledge that the Hunga plume adversely affected the DIAL and SAOZ ozone retrievals, and therefore the post-eruption data from those ground-based instruments are not used in their analyses. The impact of the extreme stratospheric hydration from Hunga on MLS retrievals is also discussed (although that description needs some clarification, as noted in the specific points below). In contrast, while cloud contamination is noted, the potential effects from Hunga on IASI data are not mentioned, and those measurements are presented with no Hunga-related caveats whatsoever. It is hard to believe that the IR measurements from IASI would be completely unaffected by either the enhanced gas-phase SO2 following the eruption or the sulfate aerosol that formed from it within the first week. Indeed, as the manuscript shows, the region of low ozone is highly aligned with the region of initially high SO2. The conversion to sulfate is then inferred from the reduction of SO2 in the region of low ozone. Whether this is a real atmospheric feature or a measurement artifact is not clear. It is essential that discussion of the IASI data quality in the wake of the eruption be added.

Response 2: The Hunga eruption caused significant disturbances in the IASI spectra, particularly within the first two days after the eruption, when pronounced temperature and water vapor anomalies were directly observed in the radiance spectra (see Wright et al., 2022). Increases in SO2 were also detected by IASI, with conversion to sulfates occurring more rapidly and efficiently compared to other volcanic eruptions. In contrast to UV-visible instruments like OMI, which reported significant ozone perturbations following the eruption (later attributed to interference from SO2/H2SO4), no similar disturbances were observed in the IASI O3 retrievals immediately after the event. Since the spectral ranges of ozone and SO2 do not overlap in the IASI ozone retrieval, there is no reason to expect any bias. While sulfate aerosols may share some spectral range with ozone, the retrieval algorithm can distinguish between the two, as sulfate aerosols exhibit strong absorption features, and ozone variations are directly measured through its absorption lines.

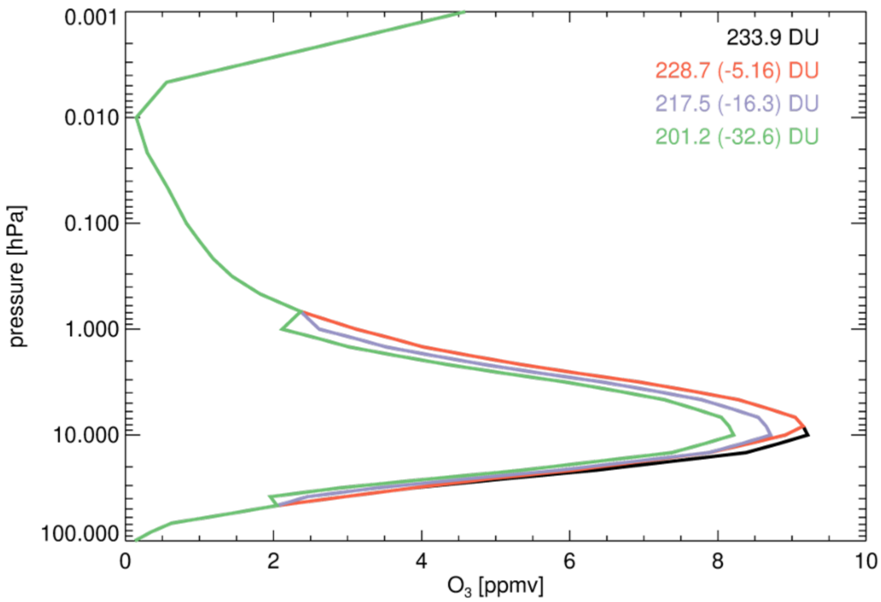
It is important to note that the retrieval algorithms for H2O, SO2, O3, and other gases rely on a priori profiles and error covariance matrices to constrain the possible concentration values. For unusual events like the Hunga eruption, some covariance matrices may be inadequate, particularly for H2O, since water vapor is not typically found at such high altitudes. However, for ozone, the algorithm was specifically designed to monitor the ozone hole, so it should account for the vertical variability between 15 and 40 km effectively.

Elements of this discussion were added into the manuscript following the IASI description.

Point 3: Further to the preceding point, I find Figure 5 and the associated discussion unconvincing. I have several technical criticisms of the figure/text, detailed in the specific points below. But the big-picture issue is that the depiction of anomalies in the IASI maps is not compelling. Anomalies of apparently comparable magnitude can be seen in many parts of the displayed area, including in the vicinity of Reunion on 15 January before the arrival of the Hunga plume, so the anomalies being spotlighted by the authors hardly stand out. Most of the maximum anomalies quoted in the text are marginal, and some are not significant at even 1σ. Moreover, the focus on maximum anomalies is puzzling. Since in most cases the exact location of these points is not specified, it is not even certain that they occurred in the region near Reunion and not elsewhere in the study area. It is not clear why a regional average anomaly on each day was not computed and related to the passage of the plume.

Response 3: In the original version of the manuscript, all negative anomalies below -10 DU were shown, regardless of their associated uncertainties. Many of these anomalies were not statistically significant at the 1σ level, which compromised the interpretation of the results and led to confusion. Regarding the update of Figure 5, we must highlight an important point. During the review process, we discovered that the IASI anomalies had been incorrectly calculated, resulting in significant overestimations. Specifically, the IASI data files analyzed during the review included ozone anomalies expressed in two distinct units: mol./m^2 (consistent with the column ozone unit) or a relative unit, representing a percentage of the column ozone amount. The anomaly unit was not uniform across all data files. As the first author, I mistakenly assumed that all anomalies were expressed in mol./m^2. Consequently, the anomalies were overestimated by a factor of 10. After correcting this issue, the column ozone anomalies are now accurately calculated and fall within the 1–1.5 % range. For the Hunga-related ozone anomalies, we now obtain values that are significant at the 2σ level, greatly improving both the quality and relevance of Figure 5. We have revised the text associated with Figure 5 to enhance clarity and ensure that the location of the anomaly is explicitly specified whenever it is mentioned. However, in response to the referee's final comment: given that the transient depletion event changes both position and shape each day, and that the evolving atmospheric conditions result in varying observations and a different number of data points, it is not feasible to compute a regional average for this anomaly. At best, we could indeed derive an average total ozone content for the entire study region or centered around specific latitudes. However, since the number of available IASI data points varies each day, such an average would not accurately represent ozone variability related to the passage of the plume.

Point 4: The maximum anomaly in January 2022 TCO from IASI (about –39 DU) was linked to the average ozone anomaly in Hunga-influenced profiles measured by MLS, which peaked at 15 hPa. There are several issues with this aspect of the study, starting with relating maximum anomalies in TCO to average anomalies in vertically resolved ozone. In addition, the 15-hPa average ozone anomaly from MLS is not significant (-0.4 $\pm$ 0.7 ppmv, 1σ), so these results are even less convincing than those based on column ozone. Most importantly, it is not possible to reconcile the magnitudes of the two



sets of anomalies, as illustrated in the figure embedded below. The black line shows a climatological MLS ozone profile (for 2005, a representative year) calculated over the region 10°S–30°S; its associated stratospheric / mesospheric (100–0.001 hPa) burden is 233.9 DU. The red line shows the same climatological profile perturbed with an anomaly like the one indicated in this study (-0.4 ppmv at 15 hPa). This modified profile has an ozone burden of 228.7 DU, only about 5 DU less than the original profile. The purple line shows the climatological profile with 0.5 ppmv subtracted between 40 and 1 hPa, effectively perturbing the bulk of the stratospheric ozone layer. The associated burden for this profile is 217.5 DU, a reduction of about 16 DU. Finally, the green line shows the results for 1 ppmv subtracted from the climatological profile over 40-1 hPa. In this case, the reduction in the burden is 32.6 DU, still about 6 DU less than the maximum anomaly in TCO reported in this paper. The key point is that the entire stratospheric ozone layer would have to be substantially perturbed to achieve an anomaly in total ozone of the magnitude asserted here. If indeed TCO was truly reduced by as much as 39 DU, then the decrease must have occurred in the troposphere rather than the stratosphere, in which case it is very unlikely to have been related to the Hunga eruption.

Response 4: Indeed, a direct comparison between IASI total column ozone measurements and MLS stratospheric ozone measurements is not appropriate, as the two instruments sample different atmospheric layers and use distinct observation geometries and techniques. We answer this comment in two main points:

- 1) The referee's comment prompted us to refine the selection criterion for better representation of the impacts of the sulfate aerosol clouds on ozone. Upon revisiting the MLS profiles selected by the initial criterion, we realized that it resulted in ozone profiles with anomalies occurring in two distinct pressure ranges. These profiles are displayed in Figure 1 below (page 4). Profiles exhibiting a significant ozone decrease around the 15 hPa level (approximately 28.5 km) are shown in the left panel, those with a decrease near the 26 hPa level (approximately 24.8 km) are shown in middle panel, while those with a decrease at both levels are shown in the right panel.
  This observation aligns with findings from Legras et al. (2022), particularly in their Figures 2 and 4e, which indicate that "the aerosol plume was initially formed of two clouds at 30 and 28 km" on January 15. CALIOP data in their Figure 2 suggests that by January 28, these clouds have

descended to approximately 27 and 25 km, indicating a loss of roughly 3 km in altitude after 14 days. Our MLS data, collected between January 16 and 23, captures the aerosol cloud during this interval, suggesting that the distinct ozone loss regions observed in Figure 1 are associated with these two aerosol clouds.
The initial selection criterion produced ozone anomalies across these distinct regions, regardless of whether they were caused by one aerosol cloud or the other, resulting in averaged anomalies that masked the individual impacts (hence the non-significant mean ozone anomaly that we found). To emphasize the ozone loss from a specific aerosol cloud, we modified our selection criterion. The new selection criterion represents a sub-ensemble of the previous one. As before, locations with water vapor mixing ratios exceeding 100 ppmv between 10 and 100 hPa were identified. The corresponding ozone profiles were then categorized into four groups: 1) negative ozone anomaly at 15 hPa (associated with the upper aerosol cloud), 2) negative anomaly at 26 hPa (associated with the lower aerosol cloud), 3) anomalies at both 15 and 26 hPa (both aerosol clouds), and 4) no negative anomaly. This method yields significant negative ozone anomalies in the first three cases. This refined criterion allows us to characterize and quantify the impact on each sulfate cloud on ozone levels over the Indian Ocean between January 15 and 23. This significant improvement has been added as one of the article's objectives within the introduction.

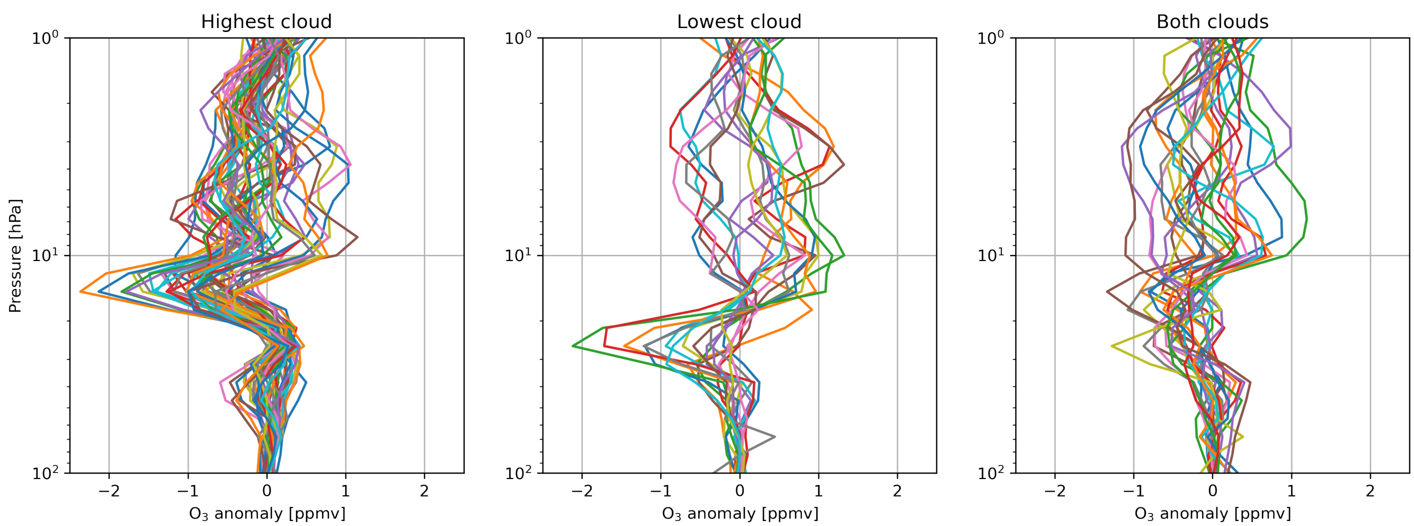


Figure 1: Anomaly profiles selected from a modified criterion, showing ozone decrease at the 15 hPa level (left), the 26 hPa level (middle) and at both levels (right).

- 2) Our goal is not to directly correlate IASI anomalies with MLS ozone anomalies at $\sim$15 hPa, nor to suggest that all ozone loss observed by IASI should match MLS measurements in magnitude. Previous research has already demonstrated stratospheric ozone loss following the eruption; our focus is to show how ozone observations can capture this impact within the stratosphere, where the bulk of the volcanic plume was injected. Because of the difference in observation geometries and techniques between IASI and MLS, the magnitude of the anomalies between both instruments might not match perfectly. However, they do match in showing clear ozone loss in the same pressure range. Our revised Figure 6 shows that ozone loss was particularly present in the 10–30 hPa range; but seems to extend up to $\sim$ 1 hPa. Similarly, IASI partial ozone columns in these pressure ranges show a clear ozone decrease. This result is illustrated in Figure 2 of this document, which shows IASI ozone partial columns for different pressure ranges (specified in the upper right corner of each panel) for January 21 only, where the transient ozone depletion is best seen. These maps do not show ozone anomalies, but they do consistently show lower ozone levels over an extended area of the Indian Ocean. Note that the minimum and maximum values for the colorbar change for each panel. The pressure ranges shown in this figure (from 25.5 hPa to 6.6 hPa) are the ones where the impact on ozone is most visible, but it could be extended to lower and higher pressure levels. However, the troposphere does not seem to show signs of ozone reduction, at least not as much as the stratosphere. This is illustrated in Figure 3 of this document (page 6), where we show the tropospheric and stratospheric ozone columns for January 21. These columns were separated using IASI's estimation of the tropopause height. The region with the lowest stratospheric ozone column corresponds to the location of the ozone depletion event. However, the IASI pixels within this area do not coincide with the lowest tropospheric ozone values.

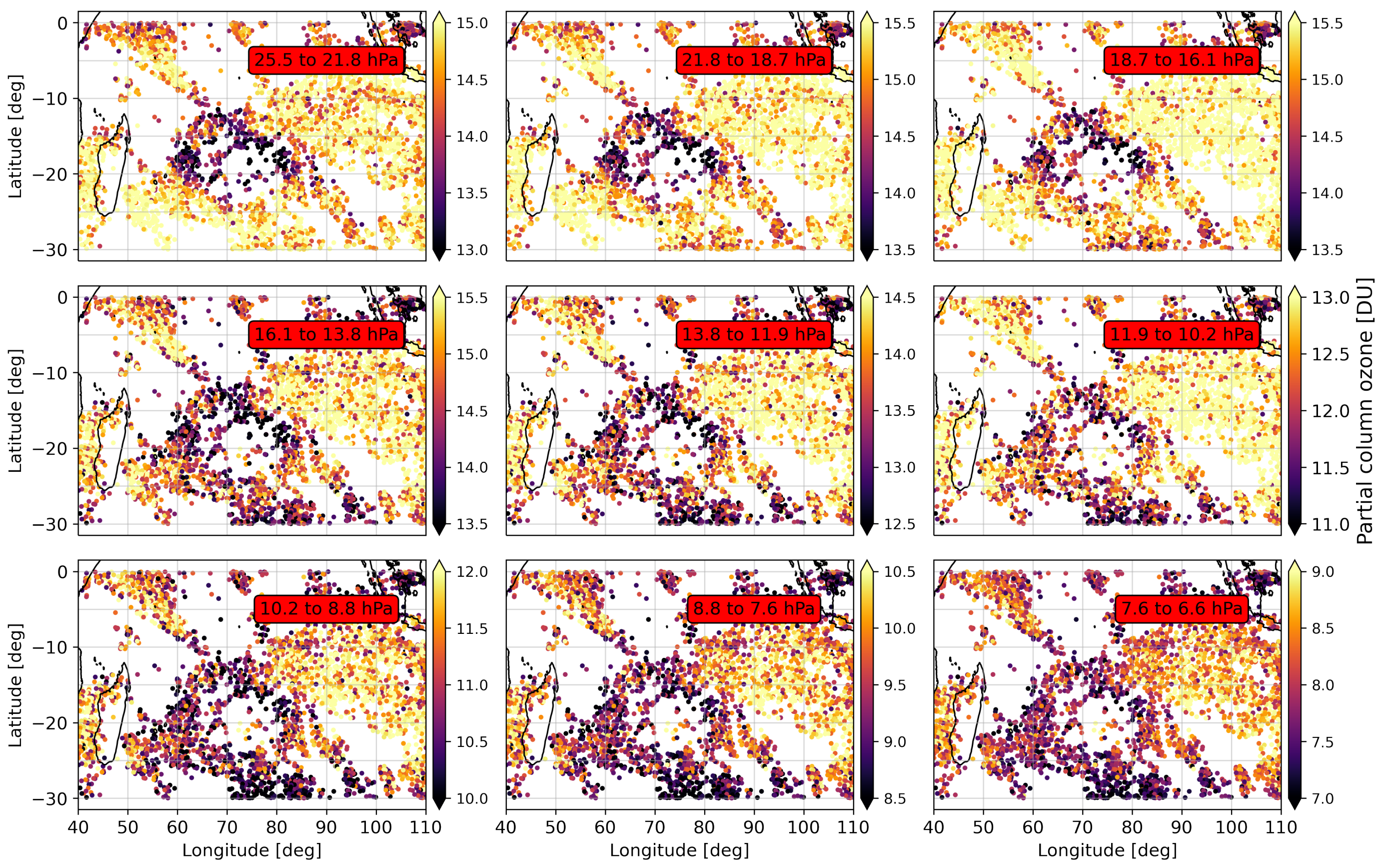


Figure 2: IASI partial column ozone for different pressure ranges (specified in the top right corner) for January 21, when the TODA is best captured.

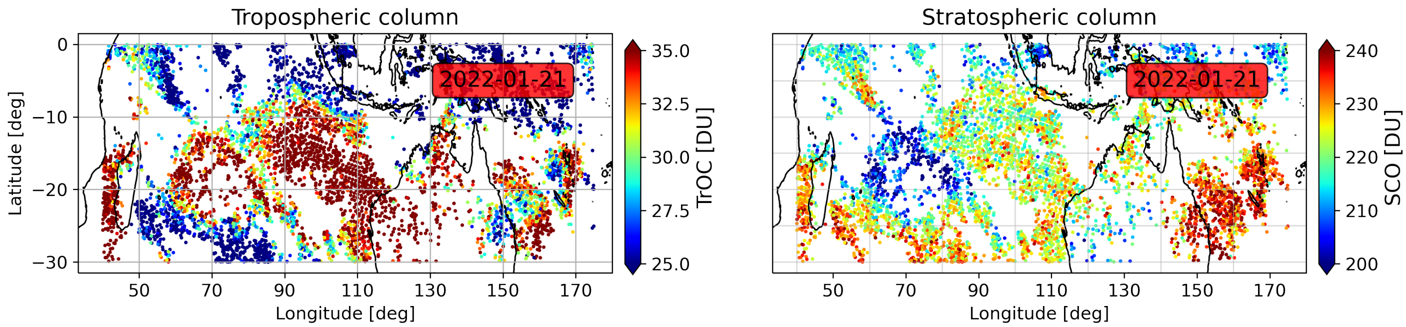


Figure 3: Tropospheric (left) and stratospheric (right) ozone columns derived from IASI on January 21.

Point 5: I fail to see the point of much of the discussion in Section 3.5 on the transport of Hunga-influenced air masses over the Indian Ocean. Several previous studies tracked the early dispersion of the plume, including Millán et al. (2022), Legras et al. (2022), and Khaykin et al. (2022); moreover, its presence over Reunion within a week has already been established by Baron et al. (2023) and Evan et al. (2023). Even if the authors felt that further confirmation was needed, the HYSPLIT trajectory calculations would have been sufficient. Instead, maps of ERA5 PV are shown and the fact that they reveal no "marked discontinuity" in the PV field during this period is argued to be evidence that east-to-west isentropic transport at 600 K was possible. It is not clear what kind of atmospheric feature it is thought may have impeded such transport. An issue that is overlooked in this discussion is that ERA5 does not assimilate water vapor measurements, and thus it did not accurately capture post-eruption perturbations in stratospheric circulation, as discussed for MERRA-2 by Coy et al. (2022).

Response 5: The ERA5 PV maps were included to provide an alternative dynamical perspective alongside the HYSPLIT simulations. However, it is true that they do not significantly enhance the understanding of the dynamics and dispersion of the plume, as this has already been addressed in the studies cited by the referee. Therefore, as suggested, we have decided to exclude this result.

Point 6: Fundamentally, the raison d'être for this manuscript is not clear. Much of the analysis centers on evaluation of MLS and IASI ozone data through comparisons with DIAL and SAOZ measurements made at Reunion under background conditions. The statement is made "Based on the excellent correlation and agreement between satellite (MLS and IASI) and ground-based instruments (stratospheric lidar and SAOZ) over Reunion, it appears relevant to use satellite ozone products to investigate the changes in the distribution of ozone over the study region." But this is hardly a surprising result – both MLS and IASI are very well characterized data sets that have already been employed extensively in similar kinds of studies, including over the region in question. In fact, arguably the entire intercomparison portion of this study was unnecessary. On the other hand, validation of the satellite measurements – in particular those from IASI – under perturbed post-eruption conditions would have been valuable, but that was not possible using the Reunion data as noted above. Furthermore, this work seems to have provided no additional scientific insights beyond those already presented in the papers by Baron et al. (2023), Evan et al. (2023), and Zhu et al. (2023). Indeed, as the authors note, Baron et al. (2023) presented the lidar data and talked about the passage of the Hunga plume over Reunion. Evan et al. (2023) presented MLS (and other) data, including ozone, over Reunion during the same timeframe. Evan et al. (2023) and Zhu et al. (2023) elucidated the mechanisms giving rise to the observed low ozone (conclusions that this paper makes no attempt to add to). Nothing in this current study is new, other than the addition of total column ozone measurements, whose reliability in this particular region at this particular time has not been adequately addressed, as noted above.

Response 6: The intercomparison is essential for obtaining reliable results on ozone depletion during the Hunga event, given the specific nature of lidar data. It also validates the agreement between Maïdo DIAL and MLS ozone profiles, which had not been previously assessed. Calculating averages and variabilities for the background period (2013–2021) is crucial for accurately characterizing Hunga-related anomalies, as these cannot be derived without a clear and coherent representation of mean observations. This study adds a new perspective by identifying ozone decreases linked to each of the aerosol clouds, while also presenting IASI SCO maps for the first week following the eruption, which had not been done before. Furthermore, in contrast to previous studies–such as Baron et al. (2023) and Evan et al. (2023), which focused on static balloon profiles, or Zhu et al. (2023), which uses models–this research provides a global perspective for the Indian Ocean, relying solely on observations. The goal is not to elucidate the mechanisms behind low ozone, but to demonstrate how ozone anomalies are revealed in IASI and MLS data, and how they relate to the two initial sulfate aerosol clouds.

Point 7: Throughout the manuscript, numerical results are reported with what seems to me to be an unjustifiably high degree of precision. As just one example, the maximum anomaly in IASI TCO is stated to be -38.97 $\pm$ 25.39 DU. This "false precision" needs to be removed.

Response 7: Numerical results have been rounded to avoid false precision where necessary. However, double precision was maintained for AOD and sAOD results because their background values typically fall within the order of 10^{-1} or 10^{-2}. Similarly, statistical results (such as coefficients of determination and regression slopes) were retained at double precision for accurate comparisons between datasets.

Specific substantive issues:

Point L1: Most of the aerosols of stratospheric significance were not emitted directly by the volcano, but rather arose through subsequent SO2 conversion to sulfate.

Response L1: We acknowledge the referee's point that mentioning the volcano's emission of aerosols could be misleading in the context of ozone loss, as only stating "aerosols" implies ash, which is less relevant compared to sulfates or SO2. We have therefore removed all references to this in the revised manuscript.

Point L16: It is not clear why the 2018 WMO Ozone Assessment is referenced for this general statement, rather than the most recent Report from 2022, which is cited elsewhere in this manuscript.

Response L16: This point has been amended.

Points L30-42: I have several comments on this paragraph:

- **Point 1**: It is stated in the first sentence that eruptions can influence tropospheric ozone, but the rest of the paragraph does not elaborate on this point at all, and it is not clear why it is relevant to this paper (unless the observed reduction in ozone is in fact occurring in the troposphere). The connection to tropospheric ozone needs to either be explained better or omitted altogether. Moreover, for clarity, in L36 "contribute to ozone depletion" should be "contribute to stratospheric ozone depletion".

- **Response 1**: Since our article does not concern tropospheric ozone, we recognize that adding this information may be misleading. We omitted this information in the new version of the article.

- **Point 2**: It is stated that eruptions release substantial amounts of aerosols, but the volcanic aerosols of most consequence for the stratosphere are those formed subsequently by the conversion of SO2 to sulfate, not those (e.g., ash) emitted directly by the volcanoes.

- **Response 2**: Similar to Point L1, this comment has been amended.

- **Point 3**: Literature citations are inadequate. It is not sufficient to cite only Tie and Brasseur (1995), Hofmann and Solomon (1989), and McCormick et al. (1995) for these points – many more references than these would be relevant in each case. At the very least, an "e.g.," needs to be added in front of all of these references.

- **Response 3**: Of course, we agree many more references could be cited to support paragraph. As suggested, we have added 'e.g.' to indicate that the references provided are examples.

- **Point 4**: It is not clear what is meant by the sentence "*Additionally*, reactive anthropogenic chlorine compounds may be enhanced in volcanically perturbed regions, leading to *further* ozone depletion" [emphasis added]– how is this different from "the activation of chlorine compounds on volcanic particles", "ozone depletion through heterogeneous chemistry", and "relationship between SO2 and chlorine in causing ozone decline post-eruption" that have already been mentioned in the preceding three sentences?

- **Response 4**: We thank the referee for pointing out this repetition. In the revised manuscript, this paragraph was shortened to avoid repetitive information.

Point L59: Wright et al. (2022) is not the most suitable reference for the Hunga aerosol perturbation; in addition to Sellitto et al. (2022), other appropriate work to cite for this point include Khaykin et al. (2022) and Taha et al. (2022) – both already cited elsewhere. Wright et al. (2022) is pertinent to the statement about the comparative energy release by Hunga, so it should be moved to that part of the sentence.

Response L59: We thank the referee for pointing this out. This point has been amended.

Point L61: Sellitto et al. (2022) is not really the best reference for the magnitude of the Hunga water vapor injection; it should be replaced here by Khaykin et al. (2022) and Vömel et al. (2022, https://doi.org/10.1126/science.abq2299).

Response L61: This point has been amended.

Point L63-65: The sentence "As a result of the main austral summer stratospheric circulation and the prevalent phase of the QBO, the first signs of the Hunga aerosol plume's passage over Reunion were noticed only 4 days after the main eruption" is problematic for several reasons. First, it's not clear what "main" means in this context (and the word "main" is used in three other places in the paragraph in reference to the eruption, so it is confusing). Second, the QBO is mentioned, but its influence is not made clear – was the QBO in an easterly or westerly phase at the time of the eruption, thus did it delay or accelerate the plume's arrival over Reunion? I believe that the authors mean that the prevailing westward flow brought the plume to the region of Reunion very quickly, such that it could be observed by instruments there within a short period of 4 days, but the wording is ambiguous and could be misinterpreted. Third, this is the first mention of Reunion in the main text. Since its importance to this work has not yet been established, it comes out of the blue and is a bit jarring. A lead-in sentence introducing Reunion and giving the reader a hint about its role in this work would be good. Otherwise, the relevance of the following information is unclear.

Response L63-65: The referee's interpretation of the sentence is correct. Our original intent was to highlight that the combination of the austral summer stratospheric eastward circulation and the QBO easterly anomaly allowed Hunga's plume to reach Reunion quickly. However, we chose to omit this detail because: 1) it was not an essential information (it was the only mention of the QBO in the article), and 2) including it would require additional, unnecessary context. Specifically, we would need to explain (see Stocker et al., 2024) that the QBO phase during the eruption likely induced an easterly anomaly above ∼30 km, which accelerated the plume detected at ∼34 km over Reunion four days later (Baron et al., 2023). Below 30 km, the QBO produced a westerly anomaly. We recognize that the word "main" was overused and have replaced it with more precise terms. Additionally, we have added a preceding sentence to better explain the significance Reunion in the context of our study, making it clearer for the reader.

Point L67-79: This discussion of the results of Evan et al. (2023) and Zhu et al. (2023) could be better organized – it jumps back and forth between heterogeneous and gas-phase reactions, making it difficult to follow. More importantly, some of the results of those studies are misstated. First, the Hunga-induced stratospheric cooling enhanced heterogeneous reaction rates but was not a factor in the rapid conversion of SO2 to sulfate aerosols, as is implied by the current wording (L69-70). (Also, a reference to the earlier paper by Zhu et al. (2022, `https://doi.org/10.1038/s43247-022-00580-w`) should be added for the impact of abundant OH from the Hunga hydration on the rapid sulfate formation.) Second, I was puzzled by the emphasis on photolysis of Cl2 (L76), as this is not part of the conclusions about gas-phase chemistry reported by Zhu et al. (2023) as is suggested, but then I found a similar sentence in the paper by Evan et al. (2023). However, Evan et al. are talking about the negative HCl anomaly arising from *heterogeneous* chlorine activation on sulfate. Their statement about Cl2 photolysis is made in connection with the colocated positive anomaly seen in daytime ClO. It is not correct that this is a "key gas-phase mechanism contributing to ozone loss".

Response L67-79: The information regarding Hunga-induced stratospheric cooling was not accurately conveyed and may have implied incorrect facts. We have rephrased it for clarity. We acknowledge the confusion caused by intertwining the explanation on 1) heterogeneous chemistry and 2) gas-phase chemistry. We have revised the text to provide a clearer, more distinct explanation of ozone loss. We acknowledge that we misinterpreted the information from Evan et al. (2022) and mistakenly included it in our study. As suggested, we have corrected this by deleting it from the "key gas-phase mechanism" list.

Point L80-93: Although the longer-term evolution of the Hunga water vapor and aerosol plumes is certainly interesting, it is not clear what relevance any of this has to the ozone distribution in the first week following the eruption, which is the focus of this study. If such discussion is retained, then it needs to be much more comprehensive in its summation of the existing literature on Hunga's radiative impact. Moreover, if the radiative effects from Hunga in subsequent months are covered here, then why are its chemical effects ignored?

Response L80-93: Indeed, the radiative impacts of the Hunga eruption are not relevant to our study. Since this information does not contribute to the analysis, we have removed this section.

Points L95 and L136: Livesey et al. (2008), which is a conference proceeding, is not a suitable reference for Aura MLS. The paper by Waters et al. (2006) is sufficient.

Response L95 and L136: We replaced the "Livesey et al. (2008)" reference by "Waters et al. (2006)" in these two lines.

Point L98-99 and L102: The phrase "dynamics of its advection" seems strange to me, since "dynamics" and "advection" are essentially synonyms. I suppose that the authors mean that they will show details of the plume's transport, but this should be clarified. Moreover, it is not clear what "its" in this sentence is referring to – grammatically it does not make sense.

Response L98-99 and L102: We thank the referee for pointing out this wording issue. This phrase has been revised in the article's objectives to more clearly convey the aim of showing the zonal displacement of the anomalies.

Point L111-116: It is not sufficient to simply state that the temporal and vertical resolution of the lidar data is "high". This information should be specified, especially the vertical resolution. Moreover, it is not appropriate to characterize a data set consisting of a total of 470 profiles obtained over a 9-year period as having "high" temporal resolution.

Response L111-116: In the revised version of the manuscript we included information regarding the lidar's vertical resolution, which varies with altitude. We omitted the information relative to temporal resolution as we recognize it was not appropriately articulated.

Point L132-135: Aspects of the MLS description need to be improved. The term "consistent measurement frequency" is ambiguous – initially I thought it was referring to spectral frequency. Thus, "spatial sampling" would be better. Also, it is not clear what "consistent" means in this context (and the MLS orbit ground tracks do differ slightly from one day to the next) – I would delete this word. It is not quite correct to refer to MLS as "a radiometer" (the instrument actually consists of seven radiometers); this is an unnecessary detail that it would be better to omit.

Response L132-135: The word "consistent" was omitted. Our intention was to convey that, unlike most ground-based instruments, satellite observations enable regular, constant data acquisition, thus providing continuous monitoring. We also removed the term "radiometer" for clarity.

Point L136-140: The recommendations of Millán et al. (2022) are slightly mischaracterized. That paper stated that the reliability of MLS measurements *inside* the Hunga plume (not "close to" it) was degraded in the first few weeks immediately following the eruption, because of the enormous enhancement in H2O concentrations. The statement that "MLS v4 relies only on profile retrievals from O2 signals whereas v5 also uses the H2O line" is unclear – this statement refers specifically to how information about *instrument pointing* (required for the retrieval of atmospheric composition profiles) is obtained in the two versions. This should be clarified. In addition, Millán et al. (2022) indicated that the standard MLS data quality screening protocols should NOT be implemented for the v4 H2O data during that initial post-eruption period. On the other hand, such filtering should still be performed for the O3 measurements, whose quality, as noted in L139, was unaffected by Hunga. The description of the MLS v4 data handling is unclear on this issue – since both the v4 and the v5 MLS Data Quality Documents are referenced in L157, the implication is that the v4 data (both H2O and O3) were screened, but the data filtering recommendations should be followed and the approach taken should be stated explicitly.

Response L136-140: We have replaced "close to" with "inside" to align with Millán et al. (2022). Additionally, we have clarified the explanation of the differences between the v4 and v5 profiles. In response to the referee's last point, we assure that we did not implement the screening protocols for the v4 H2O data, and we appreciate the referee for highlighting this oversight in our manuscript. While all screening procedures were properly applied to the v4 ozone profiles, as well as the v5 H2O and ozone profiles, they were not applied to the v4 H2O profiles.

Point L144-146: Some rearrangement of this discussion is needed. The numbers of MLS profiles being examined here – 113 influenced by Hunga and 2190 in unperturbed conditions between 15 and 23 January – only make sense in the context of a limited region (since MLS measures ∼3500 profiles per day). However, the information that the comparison is restricted to a 5-degree radius around Reunion is not provided until much later in the paragraph.

Response L144-146: This discussion has been revised for clarity on this point.

Point L147: Is the standard deviation calculated separately at each pressure level or over the whole 10–100 hPa range? In other words, is the maximum value of 0.05 ppmv quoted here never exceeded at any single level? If the standard deviation is being calculated over the entire profile, then a larger value (at, say, the level of peak ozone anomaly) could be getting "diluted" in the overall profile standard deviation.

Response L147: The standard deviation mentioned in L147 refers to the maximum value observed at any pressure level within the 10–100 hPa range, rather than an average over this range. We clarified this information by changing the sentence to: "The maximum standard deviation did not exceed 0.1 ppmv in any of the 10 to 100 hPa pressure levels". Throughout the manuscript, we updated 1σ standard deviations and uncertainties to 2σ, where possible, to reflect a 95 % confidence level. Thus, for this sentence, the standard deviation of 0.05 ppmv (1σ) was updated to 0.1 ppmv (2σ).

Point L152: I am interpreting the statement "repeated for both ascending and descending nodes" to mean that for the comparison between perturbed and unperturbed conditions, background values were calculated separately for the measurements obtained on the two sides of the orbit. Since ozone at these altitudes does not display large diurnal variations, I am wondering why it was considered necessary to derive both daytime and nighttime background profiles.

Response L152: We included the phrase "repeated for both ascending and descending nodes" to provide a comprehensive description of our methodology for deriving background profiles. However, this was not meant to imply that we processed the ascending and descending profiles separately. In fact, we combined all profiles from both nodes within a single day to produce an average profile, rather than distinct daytime or nighttime profiles. To clarify, we have reworded the sentence to: "All ozone and water vapor profiles within a 5-degree radius of each of the January 2022 impacted profiles were collected, regardless of the satellite's ascending or descending node".

Point L206-207: While the three references cited in this sentence are pertinent to the statement that persistent Hunga effects were confined to the stratosphere, they are completely unsuitable for the point that they immediately follow, which is that the stratospheric circulation is stable and stratified. In fact, no such statement is needed to justify the use of trajectory calculations (a very common technique). I recommend deleting everything in this sentence up to "we used HYSPLIT".

Response L206-207: This point has been amended.

Point L218-219: Similarly, PV is so widely used now for characterizing isentropic transport in the stratosphere that not only is such a list of citations unnecessary, but also the one provided is so seemingly arbitrary and self-referential that it does more harm than good. This sentence should be deleted.

Response L218-219: This point has been amended.

Points Section 2.4:

- **Point 1**: MLS retrievals are output on a pressure grid, whereas the comparison with the lidar measurements uses altitude as a vertical coordinate. How the MLS measurements are placed onto an altitude grid needs to be explained.

- **Response 1**: To place MLS measurements on a pressure grid, we downloaded MLS geopotential height profiles, which associate geopotential height to each MLS pressure levels. Since geopotential height can be easily converted into height above mean sea level (using an approximation of variation of gravity with altitude), we used geopotential height profiles to associate a height above mean sea level (i.e. altitude) to each pressure level. This clarification was added into the manuscript.

- **Point 2**: As noted above, the vertical resolution of the lidar measurements is not given, but I presume that it is much higher than that of the MLS ozone profiles (which is ∼2.5–3 km in the lower stratosphere). Simply sampling the finer profile at the MLS retrieval surfaces is not the best approach. To make a truly fair comparison between high-vertical-resolution profiles and coarser-resolution MLS data, it is necessary to follow the guidance in the MLS Data Quality Document to apply the MLS averaging kernels to and perform a leastsquares "smoothing" of the high-resolution data set (Sections 1.8 and 1.9 of the MLS Quality Document, respectively). Although performing such a procedure may not make a substantial difference to the bottom-line results, this issue is nevertheless worth some investigation. At the very least, an experiment in which the lidar profiles are smoothed over ∼1.5 km (i.e., boxcar smoothing) should be conducted to explore the impact on the comparisons with MLS.

- **Response 2**: As suggested by the referee, we proceeded to modify the methodology of the comparison between DIAL and MLS profiles by incorporating MLS averaging kernels. Doing so reduced the number of points available for the comparison and changed the appearance of the mean relative bias profile (Figure 4 of the manuscript). However, it did not undermine the agreement between both datasets. In fact, this approach has improved statistical metrics, notably reducing the RMSD and increasing the correlation.

- **Point 3**: The notation in the numerator of Equation (1) is confusing: O3 MLS - O3 DIAL(z)(z). Why is the first term not written O3 MLS(z) (as in L240), rather than putting both "(z)"s at the end?

- **Response 3**: We thank the referee for pointing out this typo. The equation was corrected in the revised version of the manuscript.

Point L267 and Figure 1: It is difficult to judge where 5°S and 25°S are located on these panels, as the x-axis latitude grid is odd. Instead of placing the vertical lines at the x-axis major tick marks, it would be more helpful to draw vertical lines at 5°S and 25°S. I also request that minor tick marks be added to both x- and y-axes.

Response L267 and Figure 1: Figure 1 was updated to show regular tick labels and marks, as suggested. We also added vertical lines at 5°S and 25°S to better visualize the location of our study region.

Point L281-282 and Figure 2: I do not understand what "This multi-year average represents an average of AOD data which is grouped into months, irrespective of the years" means – does the blue line in Figure 2 show the overall average over the 2003–2021 period, or is it just the January mean over all the years in that period? The uncertainty error bars on the red and black lines in Figure 2b are nearly impossible to see, even when zooming in on the panel. Greater contrast between the lighter and darker shades is needed. I also request that minor tick marks be added to both x- and y-axes.

Response L281-282 and Figure 2: The blue line represents the January mean over 2003–2021. This precision was added into the manuscript. We recognize that the error bars were difficult to see properly and tried a different approach. Minor tick marks were also added to both x- and y-axes.

Point L320-332 and Figure 4: This discussion requires clarification in several respects:

- **Point 1**: Saying "the bias decreases to -3.73" makes it sound as though the bias has gotten smaller, whereas it has changed sign but actually is larger in magnitude.
- **Response 1**: The revised comparison methodology, now incorporating averaging kernels, has slightly altered the results. The average bias in this altitude range is now 0.24 $\pm$ 2.12 %.
- **Point 2**: Given the large oscillations in the differences below 20 km, the average bias carries little meaning, so there is no real benefit in stating it.
- **Response 2**: With the new mean bias profile using averaging kernels, differences below 20 km are significantly reduced and we continue to report these values.
- **Point 3**: Livesey et al. (2022) discuss the presence of known systematic vertical oscillations in the MLS ozone retrievals in the UTLS; these likely play a role in the differences below 20 km seen here. However, they do not explain the rather large relative bias at 20 km.
- **Response 3**: We thank the referee for pointing to this reference.
- **Point 4**: Given the stated caveats on the lidar data, it is not justifiable to say "the MLS mean bias profile seems to under-estimate ozone concentrations by 20.73$\pm$1.89%". All that can be said with confidence is that there is a relative bias of $\sim$21% between the two data sets.
- **Response 4**: The revised comparison methodology, now incorporating averaging kernels, has slightly altered the results. The relative difference pointed out by this comment is now considerably reduced and no longer mentioned.
- **Point 5**: Similarly, the linear regression shows that "MLS profiles tend to slightly over-estimate ozone concentrations" *relative to DIAL*.
- **Response 5**: The revised comparison methodology, now incorporating averaging kernels, has slightly altered the results. The linear regression slope is now below unity, so we simply state that "MLS profiles tend to slightly under-estimate ozone concentrations relative to DIAL."
- **Point 6**: I request that minor tick marks be added to both x- and y-axes.
- **Response 6**: Minor tick marks were added to both x- and y-axes.

Point L345-371 and Figure 5: I have several issues with Figure 5 and the accompanying discussion:

- **Point 1**: The east-to-west movement of the plume during this time period was shown also by Millán et al. (2022).
- **Response 1**: Millán et al. (2022) was among the first to publish on HTHH material, focusing primarily on H2O, with some attention to SO2 and HCl, but no mention of ozone. In contrast, our work examines H2O, SO2, and O3 during the first 10 days after the eruption.
- **Point 2**: Schoeberl et al. (2022) is not an appropriate reference for the rapid conversion of SO2

to sulfate aerosol – Zhu et al. (2022) and Asher et al. (2023) are better citations for that.

- **Response 2**: We have replaced the Schoeberl et al. (2022) reference with the suggested alternatives.
- **Point 3**: It is not clear what "important" means in the context of the negative ozone anomaly.
- **Response 3**: We recognize that the signification of "important" is vague. We simply omitted this word in the revised version.
- **Point 4**: I am confused about what the error bars on the daily minimum TCO and maximum TCO anomaly values represent – please clarify.
- **Response 4**: The error bars on TCO anomalies are measurement uncertainties. This information was added into the manuscript.
- **Point 5**: Why is the IASI TCO anomaly in January 2022 (which is derived relative to the 2014–2021 IASI climatology) being compared to the SAOZ climatological January TCO rather than to the IASI January climatology?
- **Response 5**: The SAOZ climatology was selected to emphasize local data collected over a significant period (1994–2021). For completeness, as suggested, we have also included the January IASI climatology.
- **Point 6**: The color palettes used for the TCO maps need improvement, especially the anomaly one. For one thing, the color bar saturates below -10 DU, making it impossible for the reader to judge the $\sim$20–40 DU maximum negative anomalies noted in the text. Although a bright blue color is used for those largest negative anomalies, the contrast between it and the color used for the negative anomalies with magnitude smaller than 10 DU is too weak to be readily discernible without extreme magnification. In addition, positive anomalies should be easily distinguishable from negative ones – as it is now, the zero line falls in the middle of a continuum of bluish-greenish colors.
- **Response 6**: We have revised the color palettes to address the issues raised. The IASI TCO anomaly palette now enables clear distinction between positive and negative anomalies. Additionally, the anomaly maps now display only significant anomalies at the 2σ level, with reduced scatter point sizes to enhance map readability.
- **Point 7**: The overlaid contours depicting SO2 are red, not blue.
- **Response 7**: We thank the referee for pointing out this typo. This point has been amended.
- **Point 8**: It makes little sense to highlight the MLS profiles with high H2O values in green, when the orbit tracks are overlaid on SO2 maps plotted using a yellow-green color palette.
- **Response 8**: We now use a darker color of blue for the Hunga-impacted MLS profiles.
- **Point 9**: The black and white star denoting the location of Reunion is very difficult to spot.
- **Response 9**: The size of the star was increased and Reunion is easier to spot.

Point L379: While the largest anti-correlation may coincide with the maximum ozone anomaly, the r value a couple of levels below is nearly as large.

Response L379: The results in Figure 6 have been updated with the new criterion. However, as the correlation results between water vapor and ozone yielded low values ($|r| < 0.6$) and were not essential to our study, we decided to omit them. The anti-correlation maxima identified by the referee concerning the previous version of Figure 6 are located at 15 hPa and 26 hPa, directly concerning the upper and lower aerosol clouds. Within the relevant pressure range (10–100 hPa), the new maximum anti-correlation (not shown) is at the 15 hPa level, corresponding to the highest sulfate aerosol cloud. The correlation line for the lowest sulfate aerosol cloud has a maximum at 26 hPa. This is no surprise, as the refined criterion was designed to better capture ozone depletion at these levels. Separating the impacts of each aerosol cloud enhances the accuracy and significance of these results.

Point Figure 7: Why are the trajectories plotted using such thick lines? The individual trajectories are completely indistinguishable. Perhaps that is the point, but it could still be easily made using differently colored lines of more moderate thickness.

Response Figure 7: We reduced the trajectory line widths and applied a color gradient to differentiate between them, and we hope this has improved the figure in the revised version.

Point L392-402 and Figure 8:

- **Point 1**: Of what possible relevance is the location of the *global* average latitude of the subtropical barrier on these dates? In the context of this study, only its location in the region of the Indian Ocean is important.
- **Response 1**: The discussion in Section 3.5 and Figure 8 has been omitted, and this point has been amended.
- **Point 2**: The anomaly does not stay completely north of the subtropical barrier – some solid green dots are clearly present poleward of the barrier on 18 and 19 January (panels (d) and (e)).
- **Response 2**: The discussion in Section 3.5 and Figure 8 has been omitted, and this point has been amended.
- **Point 3**: It is stated that the anomaly exits the region on 22 January, but a couple of green dots still appear on the map on both 22 and 23 January, and on 23 January (panel (i)) they fall south of the subtropical barrier.
- **Response 3**: The discussion in Section 3.5 and Figure 8 has been omitted, and this point has been amended.
- **Point 4**: The red line marking the subtropical barrier is dashed, not solid as implied in the caption.
- **Response 4**: The discussion in Section 3.5 and Figure 8 has been omitted, and this point has been amended.
- **Point 5**: Using red for the subtropical barrier is a poor choice since it is overlaid on contours of similar color (ranging from purple to orange).
- **Response 5**: The discussion in Section 3.5 and Figure 8 has been omitted, and this point has been amended.

- **Point 6**: As in other plots, it is very difficult to spot the black and white star indicating Reunion.
- **Response 6**: The discussion in Section 3.5 and Figure 8 has been omitted, and this point has been amended.

Point L407: As noted earlier, most of the aerosols of stratospheric significance were not emitted directly by the volcano, but rather arose through SO2 conversion to sulfate.

Response L407: Similar to Point L1, this comment has been amended.

Point L417: It is not appropriate to say "the ozone reduction occurred at the level of the ozone layer". The ozone layer is a broad feature, extending over roughly 15–35 km altitude. MLS showed substantial (but not significant) anomalies only in a narrow layer around 15 hPa.

Response L417: We thank the referee for highlighting this inappropriate wording. We now specify the pressure level of maximum ozone anomaly instead.

Point L431-441: The statistical quantities described here are common and widely used, so I am not convinced that this Appendix is really needed.

Response L431-441: Statistical equations were removed from the Appendix.

Minor points of clarification, wording suggestions, and grammar / typo corrections:

We thank Referee 1 for their careful review and for suggesting numerous grammar and typo corrections. All of these minor points have been addressed in the revised version.