Answer to Referee Comment 2

June 11, 2019

We thank referee 2 for valuable comments and suggestions. Our answers are given below. The original referee comment is repeated in **bold**, changes in the manuscript text are printed in *italic*.

This paper presents interesting observations of hno3 from aircraft observations in the Arctic winter of 2016 indicating nitrification of the Arctic polar vortex in the 10-14 km range, the maximum altitude of the observations. While the observations are of interest and the overall analysis convincing, that nitrification of the lower most stratosphere occurred, the paper is poorly written, much too long, and should not be accepted in this form. Throughout the descriptions of the individual flights, which are used to introduce the observations, the authors claim that the observations show redistribution, enhanced hno3 layers and nitrification. This is before the methods are explained, or the reference profiles discussed. The reader has to guess how these conclusions are made. The description of the figures often resorts to a recitation of numbers in the figures. The authors specify vortex air in the figures by stating what is not vortex air. Figure caption 2 confuses by not describing the panels in order. Waypoints marked in figures are not used. Reference is made to NOy*, but it is not used further, or defined. The claims of "distinct differences," with some of the CLAMS sensitivity tests, are not well supported by the figures. The paper could be published, but only after major revision.

We thank referee 2 for pointing out that the presented observations are of interest and for the very helpful and constructive criticism which lead to major changes described under the specific items below.

The analysis should begin by describing that ozone will be used as a tracer for the air sampled and to describe why that works for 2016. Thus the majority of section 5 should appear before any aircraft data are shown. Only two detailed aircraft profiles need to be shown, first the reference profile in December which currently is not shown, and an example of the measurements in January.

We agree that the structure of our study is not stated clear enough at the beginning and understand now that our intentions are somewhat difficult to extract. As correctly pointed out by referee 2, the central result is the quantification of nitrification of the LMS ("research question 2"). However, our goal is also to show how 2-dimensional vertical HNO_3 distributions in the LMS are structured at different stages of the winter and how these patters compare with an established chemical transport model ("research question 1"). For clarification, we modified the manuscript as follows:

- *How are HNO_3 distributions structured in the LMS during the course of the cold Arctic winter 2014/16? How do HNO_3 distributions, which are affected by nitrification, compare with the stratospheric tracer ozone? How do observed small-scale spatial patters compare with a model (CLaMS)?*
- *Do tracer-tracer correlations constructed from GLORIA HNO_3 and O_3 indicate nitrification of the LMS? How does nitrification inferred from the GLORIA observations compare with that inferred from CLaMS? Can we identify a critical model parameter which results in a significant overall improvement of the agreement?*

Thus, our manuscript is structured in the following way: In section 2, we introduce the measurement campaign and the data used. In section 3, we introduce the specific characteristics of the data products and the methods and approach of the analysis.

In section 4, we present GLORIA cross-sections of HNO_3 and CLaMS simulations to address the first block of research questions. Thereby, we discuss observed patterns in the HNO_3 distribution only qualitatively; here, the underlying assumption is that nitrified regions can be identified as patterns differing qualitatively from patterns in the ozone distribution, which experiences no vertical redistribution associated with PSCs.

In section 5, we finally use the previously introduced GLORIA data to derive nitrification of the LMS using tracer-tracer correlations, and test the sensitivity of the model to modifications of different parameters (second block of research questions).

In our opinion, the discussion of individual flights provides valuable insights into the quality of the data and how nitrification patterns are structured in the LMS. Therefore, we think that the results presented in section 4 are the prerequisite for the analysis presented in section 5.

We agree that the method of quantification of nitrification for research question 2 should be provided already in the methods section and revised the manuscript accordingly. The method is now explained in the new section 3.4.

These two examples are enough to set the stage for the relative normalized frequency distribution (RNFD) discussion, and Figure 6, which is the key figure of the paper. A figure showing an example of RNFDs would also be of interest. While the authors seem to be keen on showing all of the aircraft data in detail, this really distracts from the main point of the paper. The authors should find another venue to do that and to stick here to the science which can obtained from the data.

We agree that the presentation of a larger number of flights draws away the reader's attention from the key results and removed flight 6 from the revised manuscript. However, we think that the remaining flights are important to address the first block of research questions. Furthermore, we now provide a rationale for the flight selection (see comments to referee 1). In addition, we added a figure (new Figure 1) showing an exemplary RNFD together with the single correlation points.

Further detailed comments follow by page and line number.

3.6-8 What is meant by stating that "the vertical HNO3 redistribution may be saturated"? How is a vertical redistribution saturated?

We changed the manuscript to:

At a later time, ~~the vertical HNO$_3$ redistribution may be saturated as due to the lower HNO$_3$ mixing ratio, no additional NAT particles can be nucleated.~~ in order to nucleate new NAT particles in denitrified air, lower temperatures are needed because of the already decreased HNO_3 mixing ratios. This results in a maximum potential denitrification for a given temperature. .

5.28-29 and Fig 1 caption. In the text "the identified vortex region (indicated by non-shaded areas in Fig. 1a)" and the figure caption, "light grey shading: areas that are not associated with the polar vortex", are at best confusing and at worst contradictory.

We removed this flight as according to the referee's suggestion.

The figure caption's version is more consistent with the figure, but then there is a meandering split of the vortex into a western and eastern half. The uniform width of this split seems to indicate more than a filament of non-vortex air.

We removed this flight as according to the referee's suggestion.

6.1-2 "Only above the British Isles, southern Scandinavia and north-west of Norway patches of air masses do not fulfil this filter criterion." How is the reader to interpret this statement? Is it an apology that these regions aren't also included as vortex air, thus doubting the Nash criterion? Is "southern Scandinavia" meant to indicate the southern Baltic Sea? The aforementioned region of air nearly splitting the vortex cannot be characterized as "patches of air masses". And with the criterion indicated as vortex a majority of the a/c observations are in this "patch".

We removed this flight as according to the referee's suggestion.

6.10-7.2 "In summary, the observed HNO3 structures exhibit a much larger spatial variability than those observed in the ozone distribution, indicating their formation due to redistribution processes." This statement is not acceptable. First the figure does not show a "much larger spatial variability in HNO3 than in ozone. Between 11:30 and13:00 in the flight data, both gases vary: HNO3 from 3-7 ppbv, a factor of 2, ozone from 0.4 -1.2 ppmv, a factor of 3. Second if there were a difference how does that immediately lead to the conclusion of a redistribution process? There must be some additional explanation to make this leap this early in the paper.

We removed this flight as according to the referee's suggestion.

7.11-12 "during in..." "Nash criterion, a relatively"

Done.

7.13 redundant with 7.11, please don't repeat.

We have deleted this sentence and added PSC occurrence to 7.13

7.15-16 "a number of GLORIA observations where sorted out by cloud-filtering" What does this mean? Sorted out and put where? Do the authors mean remove? I do not understand what sorting out means.

We changed **sorted out** to *removed*.

7.20 NOy* has not been explained and there is no reference and it is not used again. Is it important?

We changed "NO_y*" to "total NO_y"

7.20-23 What is particulate HNO3? Perhaps these are particles vaporized in an inlet and the hno3 gas measured, or ???

Yes, NOy containing particles were vaporized and measured as gas phase equivalent.

Why are the data not corrected for enhancement efficiency, insufficient information, small correction,...? Is it important that they are not corrected?

They are not corrected because of insufficient information on this parameter for this specific flight configuration. Here we use the data only to show qualitatively that NO_y-containing PSC particles were present at flight altitude. This information can be clearly inferred from the data, despite the uncertainty of the inferred absolute values of particulate NO_y which rely on assumptions.

Why do the presence of HNO3 containing PSC particles need to be confirmed? Confirming compared to what? What other kind of PSC particles could be present besides hno3-containing particles?

The presence of HNO_3-containg particles needs to be confirmed to distinguish from potential cirrus cloud particles. Here, we aim at showing (i) that particles (i.e. a PSC) are indeed found at these low altitudes and (ii) that the particles contain NOy and are therefore likely to be linked with designated nitrification patterns seen in the GLORIA observations.

The first particle maximum doesn't coincided with a GLORIA maximum, but should it, if the hno3 has condensed?

In a simplified view, maxima in particulate NO_y should coincide with regions of low gas-phase HNO_3 (GLORIA), as particulate NO_y that is transported downward during denitrification from higher layers is still in the solid phase. However, this simplified assumption does not necessarily hold depending on the previous de/nitrification history of the observed air masses. Therefore, here our main point is that we find NO_y-containing particles (probably transported downward by gravitational settling during the denitrification of higher layers) in the vicinity of highly structured local maxima of NO_y (which differ from patterns of the "dynamical" tracer O_3), which suggests that nitrification was going on in the observed air masses.

7.30 What is meant by "band-like structures"?

By band –like structures we referred to the connected/coherent structures tilted with altitude. We changed **band-like structures** to *coherent structures tilted with altitude* in the manuscript.

7.33 "distribution, indicating their formation by redistribution of HNO3" The authors again jump to their major conclusion without presenting any reasons.

The indication of nitrification is based on the assumption that ozone and HNO_3 are affected by the same dynamical processes. We added an explanation at the beginning of the section to support this conclusion (Page 7, Lines 11-16):

The observed patterns in the HNO_3 distributions are compared with the observed patterns in the ozone distribution. Since ozone and HNO_3 are effected by the same dynamical processes, the different patterns in the observed distributions are likely caused by processes that effect only one species (i.e. nitrification due to sublimation of NO_y-containing particles sedimented from higher altitudes). Therefore, the local HNO_3 enhancements seen in comparing adjacent air masses at a given height level and the deviations of their pattern from the pattern seen in the ozone distribution are interpreted qualitatively as a result of nitrification.

Although we are convinced that strong hints for nitrification can be inferred from qualitative comparing the GLORIA HNO_3 and ozone distributions (which provide the foundation for the subsequent quantitative tracer-tracer analysis), we agree that the main conclusion should not be stated already here and therefore changed "indicated" to "*suggest*".

7.34-35 Here for the first time the authors make an argument for their conclusion, but it is very brief and none of their data has included temperature relative to equilibrium temperatures with respect to NAT. Such information would help the reader understand why in some regions there are particles and in other regions gas phase hno3? In the regions where GLORIA data are shown, should the reader assume these are cloudfree?

Prior to the retrieval process of trace gases, the GLORIA data are cloud-filtered according to the cloud index method by Spang et al. (2004) since too strongly influenced spectra by clouds would lead to larger errors in derived trace gas abundances. However, this is a qualitative method for cloud filtering, and under conditions of low cloud index values (i.e. a high influence of particulates in the spectra, as observed here), the presence of cloud particles at flight altitude cannot be excluded. While the particulate in situ NO_y observations (i.e. measured total NOy=(gas phase+particles evaporated in the

instrument; forward inlet) minus gas phase NO_y (gas phase only, backward inlet) are sensitive to cloud particles, with the GLORIA observations only HNO_3 in gas-phase is derived; however, the simultaneous presence of particulate NOy in the same air volume cannot be excluded.
The GLORIA temperature data (not shown, see supplement to Johansson et al., 2018) indicate temperatures close to the equilibrium temperature of NAT (and well above the equilibrium temperature of ice) in the vicinity of the HNO_3 maxima and where particulate NO_y is detected. Thus, the GLORIA data support the processes of evaporation and persistence of NAT particles at temperatures close to saturation versus the NAT phase. However, the accuracy of the GLORIA temperature observations is not sufficient to explain precisely the observed pattern of particulate NO_Y observations and HNO_3 maxima.

Figure 2 caption. The panels need to be described in the order in which they appear. A figure caption is so readers can understand what is in the figure. Why confuse them by listing the contents of each panel out of the order in which they appear?

We changed this according to the referee's suggestion.

Since the interest is in polar vortex air, why do the authors state, here and elsewhere, what is not polar vortex air, rather than what is polar vortex air?

We changed this according to the referee's suggestion.

8.1-6 While the hno3 mixing ratios for CLAMS agree with GLORIA, CLAMS does not show anything like the altitude tilted features appearing in the hno3 GLORIA data. What particle information does CLAMS contain? Does that reproduce the in situ particle measurements?

The setup of the CLaMS run is focused on the stratospheric composition and the resolution below ~9 km altitude is rather poor. This is one of the reasons, why CLaMS does not show detailed structures below ~8 km altitude. The evaporating HNO_3 from the simulated particles is then distributed within a vertical large grid box and therefore hardly visible.

Figure 3 Why is waypoint A marked on the map and then not on the panels and not mentioned in the text.

Waypoint A is not shown in the panels, since this part of the flight is not shown there. We added this information in the figure caption.

9.8 Why call the hno3 mixing ratios "enhanced" and "strongly enhanced"? This language assumes the authors' pre-determined conclusions prior to the arguments being made. The hno3 mixing ratios are what they are, without this qualifying language.

We removed the distinction between enhanced and strongly enhanced values.

10.1-2 "Since those structures between waypoints C and E vary significantly from those observed in the ozone concentrations they most likely originated from nitrification" Is this now the argument to be pursued, using ozone as a tracer? But in fact Figure panels 3b) and 3c) do not support the statement. The structure and the relative magnitudes of ozone and hno3 are quite similar. How do they "vary significantly"?

When looking closely, different patterns can be identified, e.g. the local HNO_3 minimum (~3 ppbv lower compared to ambient mixing ratios) at flight altitude directly around waypoint D which is not identified in the ozone distribution. However, we agree that this aspect is difficult to extract from the figure and does not need to be pointed out here. We therefore removed this statement.

10.3-4 Now the CLAMS hno3 is "enhanced" even though the maximums and structure of the high hno3 regions do not match the observations. What is the significance of pointing out the very narrow high regions of hno3 in the CLAMS hno3? The last sentence describes well how the model and observations compare.

As described in the previous section, the "enhancement" is based on the comparison with GLORIA observations of HNO_3 mixing ratios in adjacent air masses and at the same potential temperature levels (see potential temperature isolines). Therefore, structured patterns in the HNO_3 distributions inside the LMS at the same potential temperature level and differing from the ozone distribution are likely to be affected by an additional non-dynamical process, i.e. nitrification by sublimation of NO_y-containing particles originating from higher layers.

In the used measurement mode, GLORIA measures a profile about every 13 seconds resulting in more than 250 profiles per hour. Further, the measurements during this flight span several 1000 km. Therefore, regions that appear narrow in the cross sections are still representing areas of several tens to a few 100 km and should not be interpreted as single data outliers.

11.6-9 The enhanced language is used again and the claim of structures indicative of nitrification, as if this point is obvious for almost all of the hno3 values measured by GLORIA greater than some number. Perhaps if the reader were shown "non-enhanced" measurements of hno3 they may agree with the authors about the language, but all we see are hno3 values in the range of 5-10 ppbv pretty much in every flight segment shown.

As described in the previous answer, the "enhancement" is based on the comparison with GLORIA observations of HNO_3 mixing ratios in adjacent air and at the same potential temperature levels.

11.22- It would be helpful to show some of the relative normalized frequency distributions. These could be more interesting for the reader than so many flight profiles, and the pointing out of small features in the flight profiles, which now may be removed as outliers, when the data analysis is finally explained.

We added an example for the relative normalized frequency distribution to the methods section. However, as we stated before, small features in the vertical cross sections stretching over several profiles are still representing several tens to a few hundreds of kilometers and therefore are not due to outliers.

12.5- Here finally information on the reference profile which motivated the previous language about enhancements, etc. is offered to the reader. Considering its importance the paper would be better served by showing this nominal reference data for comparison with the more dynamic data later.

We hope that we could clarify above that the qualitative discussion of nitrification described in the previous sections is not based on this reference profile. The local enhancements described in the previous sections are derived qualitatively by considering modulations in HNO_3 distributions on wide ranges at (i) constant potential temperature levels inside the vortex region and (ii) with the corresponding ozone fields. The HNO_3 and ozone data shown in the cross-section provide the foundation for the quantitative tracer-tracer correlation discussed here. While the slopes of the different overall relative normalized frequency distributions (RNFDs) indicate how the overall nitrification state of the LMS develops during the winter, the widths of the RNFDs show the variability of the nitrification state at the dates of the flights (i.e. simultaneous presence of already nitrified regions and less/non-nitrified regions; these are the "differences " in the patterns seen when comparing the HNO_3 and ozone cross sections).

I am not sure the point of sentences listing the numbers of the maximums. Tables are good for numbers. Text is good for describing general features of the figures such as the progression of the descent of the hno3 over January and how CLAMS doesn't capture this, nor the magnitude of the hno3.the

We changed this paragraph to a more descriptive text for better readability.

12.14 An ozone loss of 15% significantly reduces the nitrification, so it needs to be clarified whether this was possible. Was such ozone loss occurring in the LMS? CLAMS should be able to at least estimate this.

Our assumption of an ozone loss of 15% is based on the study by Johansson et al. (2019), where the ozone loss during the Arctic winter 2015/16 was estimated based on CLaMS calculations.

Additionally, we investigated the ozone depletion in the LMS based on CLaMS (see following Figure1 (not shown in the manuscript)). This shows a good agreement with the 15% that we assumed in our study.

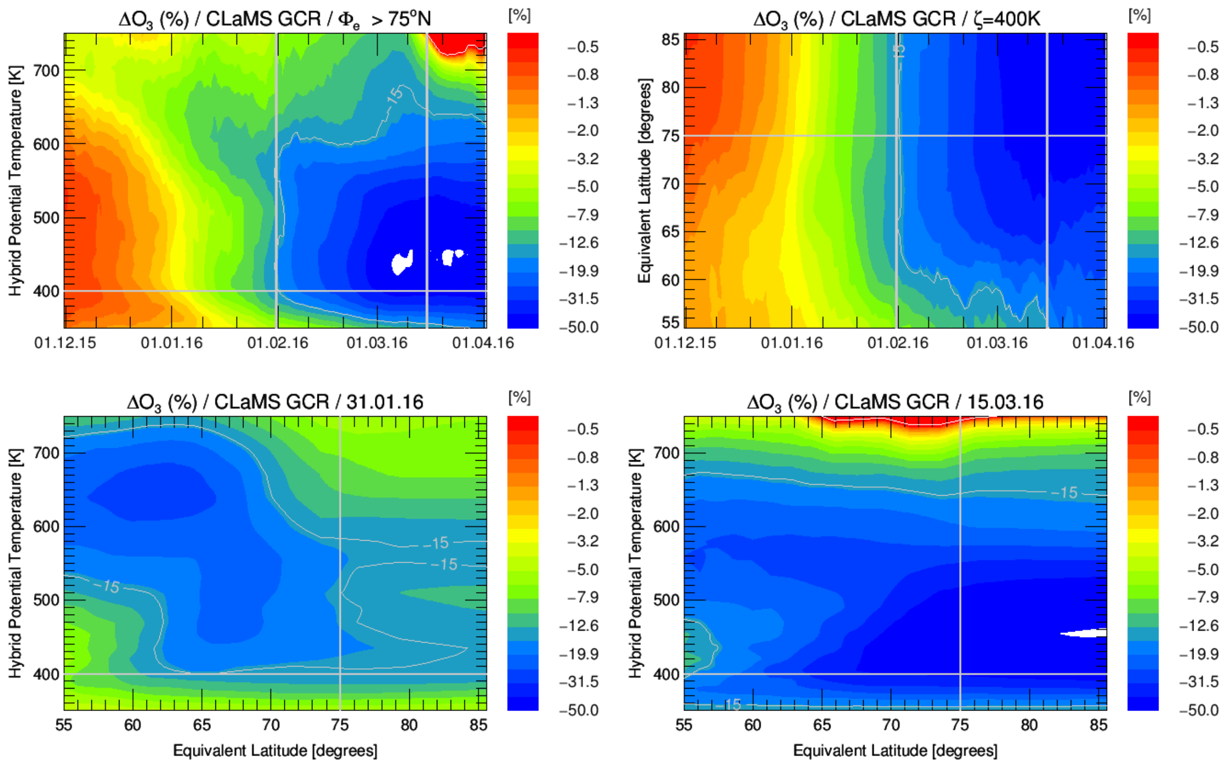

Figure 1: Ozone depletion in the LMS estimated by CLaMS.

For clarification of the given numbers, in the revised figure (now Fig. 4) we constructed the same correlation assuming a potential ozone loss of 15 % (see comments to referee 1).

12.16 I thought the measurements were filtered from non-vortex air in several ways. Are we now to assume that these correlation plots may be affected by non-vortex air?

The measurements were filtered such that non-vortex air was excluded, as described in section 3.3. As we pointed out there, a robust identification of vortex air is difficult in the altitude regions where the measurements were made. Therefore an influence by non-vortex air cannot be excluded entirely.

13.4-9 Another paragraph pointing out the numbers which the reader can more easily obtain from the figure, or if they are important could be put in a table. Text should be reserved for something

more interesting. This whole paragraph and others like it could be mostly removed and the paper would be better for it.

We changed this paragraph to a more descriptive text for better readability.

13.10-12 Yes but CLAMS completely misses the continuing descent of the hno3 observed by GLORIA.

Here, the point addressed by the referee is not clear to us. “Descent” in the sense of “nitrification by sedimented NAT particles” is clearly identified in both the GLORIA and CLaMS data as the change of the ozone-HNO_3 slopes from flight to flight, but the extent is different in the GLORIA data and the simulation (i.e. less nitrification or “slope change” is found in CLaMS data, but the nitrification is not completely missed). On the other hand, “descent” in the sense of “diabatic airmass descent” cannot be read from this plot, as this diabatic airmass descent would affect ozone and HNO3 in the same way and the correlation slope would remain unchanged.

13.27-29 “The comparison is based on the RNFDs depicted for the individual flights in Fig. 6.” And that is all there is to say about the sensitivity analysis of CLAMS compared to the flight data? Amazing, after the paragraph above describing all the different scenarios to test the sensitivity of CLAMS, no discussion of the results which indicate that the CLAMS results are almost insensitive to these perturbations.

This statement was used to refer to Figure 6 which is the basis of the discussion in the following sections 6.1 – 6.4.

As suggested by the referee we decided to remove sections 6.1 – 6.4 and offer a more general discussion in section 6 to the reader (see revised manuscript).

Figure 6 the flight dates should be added to the figure panels for reference later.

We added the flight dates in Figure 6.

13.30-16.6 Without much of a difference observed in the summary RNFD plots for the CLAMS sensitivity tests the authors proceed to discuss the flight 6 cross section and its sensitivity simulation in detail, pointing out fine features/differences in Figure 7. But what is the conclusion reached from this detailed discussion?

As stated before, RNFDs and cross sections offer a different perspective on the data. While RNFDs give a more general picture of the data, the cross-sections give additional information on the representation of spatial fine structures. However, we agree that the observed differences are small and therefore decided to move Figure 7 to the appendix.

Line 14.8, “However, the overall structure in the RNFD is similar to the reference simulation.” Exactly, which is clear from Figure 6a. With this diversion back to detailed discussions of cross sections I gave up on the paper, assuming the same was going to be done for each subsequent flight. Although this is not done, neither is a general discussion of figure 6 offered. Is it really important to go through each model sensitivity difference for each flight when there are so few differences with the reference? A more helpful discussion of the figure would organize it by model sensitivity, and only discuss those sensitivities which make a significant difference with the reference in the direction of the observations. Based on Figure 6 this criteria would shorten the discussion considerably. Figure 7 and the current sections 6.1-6.4 should be removed.

We agree that a discussion based on model sensitivity would be more helpful and therefore changed the manuscript to a more general discussion of Figure 6.

17.11-19 Claiming distinct differences is an overstatement. The improvement of the temperature fluctuation for 31 Jan. is only evidenced by increased hno3 near 800 ppbv, otherwise it matches the CLAMS reference and all simulations lay significantly higher than the observations. The improvements in the T-1K simulation are generally hardly outside the reference except for 20 Jan. The aspherical particle case provides only a slight difference on 12 Jan. If some estimates of precision were placed on the CLAMS reference, most sensitivity simulations would be hardly outside. The sensitivity simulations simply do not support the claim of distinct differences.

We agree that the improvements are small for most of the flights. However, we observe changes for individual flights even though no sensitivity simulation shows a distinct improvement for all flights. Therefore we changed **distinct** to *noticeable*.

Which sensitivity should be chosen to improve the overall agreement with observations over the campaign? There is none.

We agree with this statement. This issue is not solved by this study. We stated in the introduction that we test how well different parameterizations within the model reproduce the GLORIA observations. Based on our results, we conclude that the sensitivity simulations show differences and lead to a sometimes improved comparison. However, no sensitivity simulation can be identified that generally improves the model. Potentially, important processes are still missing in the model, or stronger changes of parameters tested here need to be implemented.