Responses to referee#1 (Kris Wargan)

We would like to thank referee#1 for his constructive review. Our responses are organized as follows. We first summarized the major modifications in the manuscript according to the comments of the referees then we provide detailed responses to each referee's point. Our responses are written in italic green.

Major modifications in the manuscript

1. *The title has been revised in order to be less general, according to comments from referees 2 and 3 and is now: "On the capability of the future ALTIUS UV-VIS-NIR limb sounder to constrain modelled stratospheric ozone"*
2. *We also realized that the ALTIUS simulated profiles have inherited some noise from NR since the EnKF procedure adds small perturbations to each model state at each model time step. This would explain some of the larger variability and lower correlation in the comparison between (AR,NR) and (CR,NR) which was pointed out by referee 1 & 2. This is now discussed in the paper with the help of the error scaling profiles calculated by the Desroziers method for NR and AR which are shown. This is something which has not been anticipated before the first submission of the manuscript.*
3. *Two additional experiments have been added in order to evaluate the impact of the low sampling of ALTIUS during the night, a questions raised by referee 1. They consider the assimilation of MLS all data and MLS daytime data with a system configured as for the assimilation run.*

Responses

This paper describes and discusses an OSSE experiment designed to assess the impacts of ozone assimilation from ESA's ALTIUS instrument scheduled for launch in 2024. The experiment consists of an MLS ozone analysis performed with a version of the BASCOE system as its nature run, a no-assimilation control and several simulated ALTIUS assimilation runs, where different configurations of bright limb, solar, and stellar occultations are used. The results show that ALTIUS (or a similar) sensor can provide sufficient data to constrain ozone in systems such as CAMS and other reanalyses. This is a very encouraging result.

Performing OSSEs prior to new satellite missions is a standard (in some cases even required) practice that allows one to assess the usefulness of these missions and helps scientists and decision makers plan ahead. That makes the topic an important one and in line with the scope of AMT. It is great to see that our community will have an alternative ozone data source in addition to the OMPS series in the dreaded but inevitably approaching post-MLS era. The paper is a pleasure to read. The science is sound, the manuscript is clearly and logically written; the figures are legible, and the supplementary information and plots are helpful. I have only a handful of specific and technical comments and suggestions for minor revisions. Once these are addressed the paper should be ready for publication.

<u>Specific comments</u>

L41: "limb-scattered solar light during daytime" is what's called "bright limb observations" below, right? If that's the case, I'd suggest introducing this term here and using a consistent terminology throughout the paper.
Done.

L79-82: There have been more solar occultation sensors than the ones mentioned here: HALOE, POAM, ILAS. I feel it should be mentioned here that solar occultation is a well-established measurement technique for ozone.
These instruments are omitted because they are not working in UV-Vis-NIR wavelengths. This will be clarified.

Table 1: Please explain what threshold uncertainty is. I take it to mean the worst but still acceptable uncertainty. Is that correct?
This is indeed the meaning, added in the table caption.

L101: There's a new paper now in ACPD that discusses this scheme (at least I think it's the same one): https://acp.copernicus.org/preprints/acp-2020-1261/#discussion ; the authors may want to consider citing it if the new Monge-Sanz paper is accepted by ACP before this one is.
This paper is still under review so it is not cited.

L108: How well are the effects of polar heterogeneous chemistry on ozone represented?
In L104-105, it is stated that "…, COPCAT chemistry has the advantage of providing a better representation of polar ozone depletion… (Monge-Sanz et al., 2011, Jeong et al. 2016)."

Please, refer to these two papers for a exhaustive evaluation (note that we added Monge-Sanz et al. here).

L126: I take it that observation uncertainties are assumed to be uncorrelated (which is fine). Is that correct?
Yes, this is discussed in Desroziers et al.

L141: "15 days later". Is it because that way the initial condition has sufficiently departed from the assimilated state in NR?
Yes. The sentence has been rewritten to clarify this point: "It is running with the same spatial resolution than CR and is initialized 15 days later than NR and CR with the O_3 state from CR as initial conditions (the 15-day delay allows AR to have initial conditions sufficiently departed from NR initial conditions)."

L151: The latest version of the V4.2 data quality document is from 2020 (Livesey et al., 2020). I don't believe the recommendations of estimated uncertainties for ozone changed between 2015 and 2020. If they didn't it would be better to cite the latest available document (https://mls.jpl.nasa.gov/data/v4-2_data_quality_document.pdf). Same applies for all the other instances this document is cited.
Livesey et al. (2015) has been changed by Livesey et al. (2020).

L168: In the upper stratosphere the chemical time scales for ozone drop rapidly with altitude making data assimilation a particularly hard problem. Can you explain the decision to assimilate ozone at those pressures, vis a vis Errera et al. 2019 who imposed a cutoff pressure at 4 hPa for ozone despite using a much more advanced chemistry model? I'm not saying I disagree with that decision, I'm just curious if the authors find the results in the upper stratosphere meaningful.
It is true that ozone assimilation with BASCOE full chemistry can lead to several issues above ~4 hPa. First, the system cannot completely eliminate the FmO bias for ozone. Moreover, in 4D-Var, the system used to introduce negative biases in several species (e.g. HCl, NOx, H2O) to increase the amount of ozone. This the reason why several years ago, we have decided to not assimilate ozone in the upper stratosphere. However, several studies using ozone linearized chemistry displayed good results for ozone assimilation up to 1 hPa (e.g. Geer et al., 2006, ACP). According to the FmO statistics shown in our paper (e.g. Fig. S1 and S2), this seems not to be the case for the BASCOE COPCAT scheme. The reason is likely due to a negative bias in the O3 climatology of COPCAT, this climatology being computed based on multiannual simulation of the SLIMACT CTM, which suggests a negative bias in SLIMCAT as well. This has not been included in the paper where we think it is out of its scope.

L174: Uncorrelated with each other or with the NR?
Obviously, uncorrelated with each other. This is clarified as: "… ozonesondes (Figs. 3 and 4). The uncertainties of ozonesondes are assumed…"

L185: It would make sense to re-emphasize that CR is also driven by ERA-Interim so that the meteorology is consistent with NR.
Done as follows. "The control run (CR) is based on a BASCOE free model simulation (no assimilation) also driven by ERA-Interim but with a lower horizontal resolution than in NR:…"

L217-252: I like it that the authors provided this extensive explanation. It really helps if the reader knows what was done and why. It looks like you found the right balance between what needs to be done in the way of simulating ALTIUS and what can be done. I appreciate it.
Thank you for this positive comment.

L240-243 and Figure 7: The MLS data quality document also contains accuracy estimates (reported as 2-sigma). The plot only shows precisions. Is it because the estimated ALTIUS error standard deviations also represent precision? Is there anything we can say (perhaps the answer is "no") about the expected accuracy of ALTIUS ozone data?
Satellite instruments are calibrated before the flight to expect no bias. So remaining biases are unexpected and are difficult to evaluate before real observations are validated. This is why the MLS and ALTIUS accuracy is not discussed here.

L272 and Fig. 9: This harkens back to my question about the ability of the chemistry scheme to represent polar ozone loss. It looks like CR misses it completely. Can you explain? I understand that this is somewhat tangential to the topic of this paper, so just a couple of sentences of explanation will be enough.
In fact, the CR run is not so bad in representing the polar ozone loss. From Fig. 4, we could see that the agreement with ozonesondes is relatively good (see also Fig. 11 for the SP). The profile in Fig. 9 is shown on Sept. 15 when the ozone loss is the most important and the timing of the loss in CR seems to be delayed from NR and AR. This is clarified by adding at the end of Sect. 4.2. : "However, note the relatively good representation of Antarctic ozone depletion in CR, thanks to the COPCAT chemistry, when compared to ozonesondes (see Fig. 4)."

~L292: While the difference standard deviation is reduced nicely for the most part in AR compared to CR, there is a patch of values <5% in CR-NR between 30°S and 60°N at p<~20 hPa that becomes slightly worse (>5%) in AR-NR. Can the authors comment on that?
This issue has been studied in more details in the revised manuscript. The reason is due to the noise that ALTIUS simulated data have inherited from NR. This is something we did not anticipate we the manuscript was submitted.

L294-296: Why not add hatching to the plot to show the regions of significance? Or is the improvement significant everywhere?
The improvement is significant almost everywhere and adding a hatching would degrade the figure. So the figure has not been updated.

L302-304: Even though AR sits within or not too far from the NR envelope, it shows a lot more variability than NR on daily-to-weekly scales, especially in MS md TLS. I'm looking at the jagged blue line compared to the red line. Can the authors explain what's happening there?
Again, this issue has been studied in more details in the revised manuscript and would be attributed to the EnKF noise that has been inherited by the ALTIUS simulated profiles.

L307-308: Would it be possible to test this? One could run an additional short experiment with double the number of simulated "ALTIUS" observations, e.g. by also simulating observations along the descending night node. Those wouldn't, of course, literally make sense as simulated ALTIUS data but looking at what this does to the spread could help substantiate this claim.
This sentence has been removed according to our response to the previous point. However, it was much easier to assimilate MLS daytime data only instead of trying to simulate ALTIUS nighttime limb profiles from which SNR table are not available.

L355: I would be more specific here: "to constrain ozone in chemical data assimilation systems such as"
Done

Suggestions for technical corrections

L10 and below: I would replace the word "weight" with "impact" ("impact on the analysis"). I reserve "weight" to situations where we talk about different weights given to different data sources in data assimilations, etc. It's just my preference. Up to the authors.
"Weight" has been replaced by "impact".

L18: "signature". I think something like "signing" or "implementation" would sound better.
"Signature" replaced by "implementation".

L24: "limb-looking"; I suggest "limb-viewing" instead.
OK, done.

L25: I suggest avoiding using ellipses ("...") in a scientific text. This sentence (causes affecting the ozone layer and their consequences) doesn't read well. Please, revise.
I did not find something better, so the sentence has not been changed.

L38: Nearly polar -> near polar
Done.

L40: part -> parts
Done.

L67-71: Please use either "Section" or "Sect." consistently; I would opt for the full word "Section".
This is the AMT standard. I quote here from https://www.atmospheric-measurement-techniques.net/submission.html#manuscriptcomposition "The abbreviation "Sect." should be used when it appears in running text and should be followed by a number unless it comes at the beginning of a sentence."

L93: In the interest of being concise, I don't think it's necessary to repeat what the three geometries are, given they were just discussed in detail. I suggest deleting the text in the parentheses.
Done.

L100: “throughout” -> “through” or “via”
“Throughout” replaced by “via”.

L112: “consists in” -> “consists of”
Done.

L286: “reduced below” -> “reduced to below”
Done.

L334: Again, I don’t think “weight” is the best word here. Perhaps “capability” would work better, or “impact on ozone analyses” or something like that.
“Weight” has been replaced by “impact”.