

Response to Reviewer 2

Laura Pan on behalf of all co-authors

Point-by-point responses

This paper investigates the sensitivity of nadir looking satellite sensors to changes in the chemical composition of the UTLS region within the Asian summer monsoon (ASM) as compared to typically adopted limb sounders. It explores both seasonal and day- to-day variability attempting to exploit the high horizontal resolution of nadir sounders to better depict horizontal structures in the ASM distribution of tracers of pollution and stratospheric air (CO and O₃). This study could give a valuable contribution both to the broad issue of exploiting nadir sensors at UTLS altitude, and to the ability of observe and understand this region at fine spatial and temporal scale. However, attempting to tackle both challenges in the same study, the authors have failed, in my opinion, to achieve enough robustness and the study needs substantial improvement before I can recommend it for publication on ACP.

We appreciate the reviewer's criticisms and have made major revisions to address these general concerns. The overview of the revision is given in a separate file.

GENERAL COMMENTS

The manuscript is generally well written and structured, although some editing may improve it. Despite appreciating this effort of promoting a synergetic use of nadir and limb data, I find there is a certain lack of overall clarity on the aims of the paper. The study is introduced as a qualitative comparison of nadir and limb data (see introduction), but with the ambition that their results would support the use of this approach in future research. It is not clear to me whether the study aims therefore at validating the use of the adopted nadir observations under ASM UTLS or trust support from the literature and aims at producing novel results for the ASM UTLS region. They focus on 3 questions: the first two point to a validating exercise which cannot be kept at a qualitative level, the third one to a more general interpretation of the results for the ASM UTLS region, although leaving it with no clear answer.

We agree with the reviewer that the previous version of the paper has significant weakness in addressing the objectives. We have made significant revisions to improve four aspects of the paper as detailed in the overview.

This work, however, was not designed to be a quantitative comparison, as stated in the last paragraph of the introduction. In the revision, we further emphasize that this work is NOT a validation study in its traditional sense, i.e., to compare the measurements to trusted source and quantify the accuracy and precision of the retrievals. Instead, this is an exercise of “process-based retrieval EVALUATION”. Here the goal is not to substantiate the specific values produced in the retrieval, but rather, the representation of a process, i.e. the Asian monsoon dynamics and transport, in the UTLS CO and O₃ field, which is the overarching idea of the three questions. In the revision, we made effort to clearly answer all three questions posed in the introduction.

The authors use the observed ASM atmospheric composition (in comparison with MLS) to support the use of the adopted nadir data and then use the same nadir data to investigate the details of the ASM atmospheric composition at finer scale (going beyond MLS). The observed details and fine scale variability has to be proved to be a real measure of the natural variability

and not a combination of perturbed conditions and a low sensitivity retrieval. The authors appreciably introduce averaging kernels and mention vertical resolution/degrees of freedom for signal (DOFS) but do not use them quantitatively. E.g., the IASI mean averaging kernel for the 12-16 km observation (Fig1) is so broad that contributions from layers outside the range are expected and the high CO cannot be assigned to a shallow layer, contrary to MLS data. The same applies to O3. A vertical resolution of 10-14 km: the a priori profile is likely to be simply scaled to match the average sensed value of the whole tropospheric column. Indeed, the horizontal details presented by the authors are very interesting, but should they be read more like an on/off effect of a perturbation at an unknown tropospheric layer? You need to quantify how much of the results are coming from the measurement and from what layer. In order to support the authors' claims, there is a need for performing sensitivity studies to investigate the response of the retrieved profile to perturbations at varying altitude that mimic the observed ASM behavior, contributions from different layers and possible contaminations (clouds?). To this end, both retrieval simulations and atmospheric model simulations (e.g. the referenced Pan et al., 2016) would be of help. Furthermore, comparison between sounders could be performed more quantitatively, e.g. introducing convolution with the averaging kernels.

We have included new figures to demonstrate the IASI UT sensitivity. Please see point 1) in the overview.

Again, we do not aim to make quantitative comparison. As stated in the paper, the two sensors (MLS and IASI) are observing very different air masses.

I find very unusual the choice of using two different years for the analysis of the two targets. I see no reason not to compare results for nadir CO and O3 for the exact same days and regions (and also O3 from the two nadir sensors) and verify that the small-scale structures you attribute to natural variability are indeed consistent among the two targets. Even then, you may still be seeing a retrieval artefact but in response to sensitivity to different layers of the atmosphere, so giving further support to your approach.

We have re-worked the CO analysis using IASI 2008 data. See point 2) in the overview.

I encourage the authors to consider these main issues and the specific comments below in order to improve the paper to make it a valuable reference for future studies.

SPECIFIC COMMENTS (P=page, L=line)

Abstract:

P1 L14: the use of "information content" is misleading since it is not analyzed in the manuscript. You only quote mean estimates of DOFS from the literature. On the other hand, adding an analysis of the information content from the retrievals would be of great help in understating the sensitivity of the nadir observations under these conditions. How much information for each point of your profiles is coming from the measurement? How much from the a priori? Adding maps with this information could give convincing support. P1 L16: same.

In the revision, we aim to make this point clear that conventionally the phrase "information content" has a specific meaning in retrieval, and it is often quantified by the "degree-of-freedom in signals (DOFS)". In this work, we advocate a "process-based" retrieval evaluation where the information is evaluated for how effective the DOFS is in retrieval results, in this case, to represent the Asian monsoon dynamics and transport. The dynamical consistency we demonstrated with the new figures is a clear indication that although IASI has relatively weak

UT information, the retrieval used the information well to represent the dynamical variability effectively. Figs 7r, 8r, and 9r in particular demonstrate that the effectiveness of the information content is shown by the geophysical consistency of the CO maps and cross sections.

P1 L15: possibly due to a typo, the sentence starting with “Day-to-day behavior” should be rephrased discussing first the seasonal scale analysis, then moving to the day-to-day and finer scales in the following sentence.

Revised.

P1L 20: I would tune this sentence down depending on further support to the analysis of the layers that actually produce the signature in IASI data. P1 L21: The same issue applies to OMI data, for which you actually show no profiles in your work.

For OMI O₃, see point 2) in overview of revision

P1 L24: same.

Introduction P3 L4-9: this is a key statement for justifying the lack of robust analysis during the comparison. I feel the authors failed to extract enough support from the listed literature in order to prove the claims of their paper. See comments to the listed references here below. If the ambition is to show how the synergetic use of nadir and limb sounding can improve the understanding of the ASM UTLS, then you need to quantify the agreement, more carefully analyze the limitations of the two and what parts of the nadir observations can be trusted. You could then propose a strategy for how to merge the horizontally coarser and less frequent vertical information from limb sounders to drive the finer and more frequent picture coming from nadir observations.

Again, it is not the goal of this work to “validate” the nadir sensors’ retrieval through comparisons with MLS data. The goal is to examine the consistency of both limb and nadir data with the dynamical field over the Asian monsoon region. See point 3) and 4) in the overview for the nadir data’s contribution beyond what we can learn from MLS.

P3 L9-14: how can a qualitative comparison help to assess the information of these nadir viewing datasets for ASM UTLS studies? Please explain. P3 L14-16: I am doubtful on what part of the study investigates the information content and how your analysis can help further studies if the horizontal information cannot be located at a correct altitude range.

Vertical range of the IASI CO information is demonstrated in the revision. See Figs 7r and 9r included in the overview of revision.

Data Description

P3 L19-P4 L17: I think the available support for your study from the literature you list has been overestimated. In general, the profiles of CO and O₃ retrieved from IASI and OMI have such a low vertical resolution and DOFS that have almost no sensitivity to differentiate layers in the troposphere, allowing for 1 or maximum 2 independent partial columns (or points in your profiles) throughout the whole troposphere (see e.g. George et al., 2015). Liu et al (2010b) show OMI O₃ vertical profiles have 0-1.5 DOFS in the troposphere, or 14/11 km vertical resolution at 12.5/17.5 km altitude. Kroon et al (2011) show (their Fig. 1 and discussion) that in the troposphere it is almost impossible to distinguish different layers in OMI O₃ (certainly true around the 100 hPa level). Wachter et al (2012) validation reach only 225 hPa for IASI CO, i.e. below the layer adopted in this study. Bak et al. (2013) state that a large smoothing error is introduced in OMI profiles by the retrieval. Safieddine et al (2016) only use 0-6 km columns. Barret et al (2016) indeed compared IASI vertical profiles and

results from a transport model, although working on monthly means only, so leaving the validity of small scale day-to-day data in the UTLS unsupported. Averaging over long time periods or large regions, the resulting profiles are very smooth and agreement with low information content retrievals can be very good. But is this residing on the a priori climatologies? Instead, you intend to use the nadir data under variable conditions and unknown vertical gradients. Conditions of strong vertical variability may be completely missed by nadir retrievals: see e.g. Gazeaux et al. (2013) AMT where IASI O₃ profiles completely miss to reproduce the plumes at various tropospheric altitude observed by ozonesondes over Antarctica. How do these limitations affect your analysis?

The references cited in the data description meant to give the background for this work. The current study does not depend on the previous validation studies, but rather, is motivated by the limited information from validating through direct comparisons. The approach of the process-based evaluation of retrieval information is to complement the direct comparisons between the datasets.

For the ozone analyses, we are not expecting the OMI data to show multiple tropospheric layers, rather, we expect the data to recognize the stratosphere and the troposphere, which is fully supported from the averaging kernels. The result is positive, as shown in the figures in the previous submission. As a result, we have demonstrated the effectiveness of process-based evaluation.

P3 L19-P4 L17: are there detection deficiencies expected for the three sensors in the UTLS region? E.g., how is IASI affected by clouds? Could you also help the reader understanding how the numbers you give as degrees of freedom translate in how much information can be extracted from the observations? To what extent are 0.8-2.4 DOFS in IASI CO sufficient “to capture upper tropospheric variability” as you write? If you can retrieve 1 to 2 independent partial columns, help the reader to understand that you expect only 1 or 2 independent points in your profiles which will then be used to scale the climatological profile.

In the revision, we discussed the effect of cloud. In IASI retrieval, cloud coverage greater than 25% are not retrieved, which leave significant data gap in the monsoon region.

See point 1) in Overview of Revision for the question of “to what extent the IASI CO capture the UT variability”.

P3 L19 –P4 L17: At what local time are the observations taken? Are the fixed local times of the measurements affecting your analysis?

The dynamical variability we are targeting are synoptic scale features and are not affected by the observation time per se.

P3 L23: Could you use the horizontal resolution to predict at what scale you expect MLS data to lose sensitivity to finer scales as compared to nadir observations? Is this reflected in your results (e.g. the 100 km features seen in IASI CO daily maps but missed by MLS)?

Both instruments have their challenges with regarding the data gap. In the revised manuscript, we will provide more discussion on the importance of carefully interpolating the data.

P3 L25: how are these and following interpolations affecting your study? E.g., averages and interpolation tend to add further smoothing errors in the profiles therefore reducing the vertical resolution. How are the vertical profiles changing depending on the horizontal interpolation you adopt? One may expect the agreement between nadir and limb data to improve reducing resolution (and therefore variability).

Specific discussions on interpolation and smoothing are discussed in the revision. In the updated figure, both MLS and IASI are interpolated using the same method (Natural Neighbor method) and followed by 1-sigma Gaussian smoothing. For the 3D structure we again focus on the variability signature not the absolute value of enhancement.

P3 L28: please carefully discuss the vertical resolution for IASI CO (and OMI O3) as this is a key element in your analysis.

We agree. See point 1) of Overview of revision.

P3 L29: the use of a single a priori profile could lead to a bias in the average you perform and could be removed from the final result to reduce its influence.

This is an interesting subject to discuss with the retrieval team.

P4 L1-4: see comments on supporting references above.

P4 L5-17: see comments on supporting references above. L11: you should clearly state that the DOFS in the troposphere are 0-1.5 and that the troposphere is then observed as one single column (Liu et al. 2010b). L14: the fact that there is useful information does not mean you can control where that is coming from (e.g., what about contamination by photochemical production?) Note that for example the averaging kernels in Liu et al. 2010b under tropospheric perturbed conditions peak at 5 km altitude in their case study in the tropics (their Fig. 5): so as long as the signal is coming from the upper troposphere, what you show is sound. But if the signal is coming from different layers then you give a wrong picture, with no ability to distinguish among the two cases. Please find support for the kind of conditions you analyze and provide sensitivity tests. Can you support via e.g. MLS that the whole signal in the region is coming from the upper troposphere and then constrain the nadir observations?

Again, for ozone study, we focus on the separate between stratosphere and troposphere, which OMI data demonstrated excellent consistency with the tropopause data.

P4 L18-23: your use of different years is quite unusual when studying simultaneously different targets since further insight can be achieved by inter-comparison. Without studying at least one year with both targets before showing what happens in a different year, the robustness of your analysis is largely reduced. CO data from IASI on MetOP- A in 2008 were used from George et al. (2015). And why was O3 from IASI not used instead of/in comparison to OMI O3? If the observed small-scale variability is not an artefact of the poor retrieval (i.e., the smoothed response to local composition of the 1 or 2 independent partial tropospheric columns you can retrieve), then the very same structures should be present both in CO and O3. Even then, you may be seeing the same retrieval artefact but it would give further support to your approach since the retrieval of the two targets would be most sensitive to different layers of the atmosphere.

Addressed in Overview of Revision (point 2)

P4 L25: Figure 1a. It would be good to focus only on the ASM region under investigation, plotting the data with actual size of the nadir footprint and an indicative footprint of the smoothed horizontal area defining each MLS measurement. To complete the information, I would also show a companion plot with a vertical cross section passing through the ASM region showing the position of MLS tangent points and the 1 or 2 independent points of the nadir profiles (maybe with error bars indicating their vertical resolution).

Fig 1a is revised to show only the study region. Vertical cross section is in Fig 9r.

P4 L29: I would rephrase “vertical information distribution”. Figure 1b. It would be useful to add the same figure for OMI O3 AKs (and possibly for IASI O3). Or simply discuss whether they are similar.

See Overview of Revision point 1) and 2)

P4 L30: explain whether IASI 12-16 km partial columns are a standard IASI product or you calculate them to match the anticyclone vertical range. Could you compare them to partial columns from MLS too?

Revised to analyze IASI CO data at 150 hPa level. See Overview.

Comparison

P5 L12-14: could you show the vertical extent of the anticyclone in one of your plots (e.g. in the additional panel of Figure 1a I suggested, or in Figure 7)?

Shown in Fig. 9r

P5 L23: it would be helpful to add a conversion to km or hPa for the various levels you adopt (e.g. is 147 hPa for MLS CO about 14 km and layer 18 about 16-17 km?). Have you tried repeating the analysis at different levels? Does it remain consistent?

147 hPa is approximately 14 km and layer 18 is approximately 100 hPa are stated in the text.

P5 L24: see comments above, why not choosing the same year for both targets and prevent a self-consistency comparison?

P5 L28 – P6 L9: the analysis on JJA averages show quite large differences in CO and a much better agreement in O3. Even for O3, the wave pattern at 0/30N is not visible in OMI data. Why? I would expect finer scales to be resolved by nadir? Is that at a layer outside the nadir sensitivity? Can you give a more in-depth interpretation of how to read these results based on the low DOFS nadir data have? P6 L10 . . . :

We will enhance the discussions of the comparison, but the goal of the quantitative comparison in this work is mainly to provide a baseline to look at the variabilities together.

The scatter plots of Fig 3 are very useful: here clearly the correlation of O3 is very good and the sensitivity to CO is weak with likely contamination from other layers. I recall these are 3-month averages: have you performed a similar comparison on shorter time scales? Can you compare the agreement/standard deviation you find to e.g. what is an accepted comparison in validation studies from the literature? Since these are comparisons of 3-month averages and not coincident profiles, it seems

to me that the sentence P5 L17 states more confidence than what is shown by the data.

Again, the quantitative comparisons are made to characterize the differences between the two types of data and to have a baseline to look at the variability together. The scatter plots are daily binned points, i.e. the average of the data that represent the same region and same day, to the extent of different viewing angles allow. These are all points for the three months but it is not a three month average.

Sub-seasonal variability

P7 L19-20, L22-23, L28-32, Fig 5/6: IASI CO is showing enhanced values within the ASM region but also scattered features outside. Why would MLS not pick for example the feature of enhanced CO at 90E and 10S/0 if it were around the 147 hPa layer? How can you distinguish cases with high CO in the upper troposphere from cases with high CO in the middle troposphere (see the high levels of CO in e.g. Fig 7)? The different evolution of MLS CO and IASI CO over the days points toward signals coming from different layers. P8 L1: I think you are underestimating the shortages of this analysis and should be more careful with this statement. There is some sensitivity to UT CO but you cannot distinguish it from sensitivity to the rest of the troposphere. Fig 7 would be more useful if compared to MLS data (in the upper range) and to a model simulation to understand the vertical structures and compare it to that reproduced by the 2 DOFS of IASI CO.

These figures are all updated in the revision

P8 L11-L21, Fig 8: I find it difficult to read the figure as reported in the text. Could you over plot the zero-anomaly contour of one on the other? Or the GPH shape? Could you mark the trajectory of the moving local maxima to highlight the propagation of anomalies? I do not see a convincing agreement. I see it clearly in the model by Pan et al (2016). Again, with sensitivity studies you would have more support not to speculate in L19-21. Please quantify what is coming from the UT.

Revised to include zero anomaly. See Fig. 6r in the Overview for the flavor of new figures.

P8 L22. . . , Fig 9: the regional timeseries look more robust, which seems to me supporting the fact that when you average more data together you remove variability and can find a better agreement. But this is not supporting your attempt to use nadir data to investigate finer scales. Can you help extracting more information from the timeseries? What parts are in robust agreement and why? What happens if you over plot timeseries for the Tibetan and Iranian lobes? What if you over plot timeseries at different altitude?

These figures are changed in the revision.

P8 L32-34: it would be more convincing if the agreement with Pan et al. (2016) was shown adding their data to the figures. Please rephrase “able to detect the impacts of vertical transport”.

See Fig 6r in Overview – the dynamical consistency in satellite data based Hovmoller is comparable but better than the model result.

P9 L13-L21, Fig 10/11: there are large differences that need to be investigated with sensitivity tests. Are they due to different layers being involved? What is coming from measurement of the UT and what is contamination? You need to quantify this to identify what you can extract from the data you show.

The main difference in the structure is from sampling differences and smoothing.

P9 L22-L29, Fig 12: see comments for Fig 8. How can you get different frequencies for the migrating anomalies among the three plots?

Sampling density differences will contribute to the differences in apparent frequency. The figure is revised.

P9 L30: I am not sure how you attribute the differences to the sampling densities rather than (lack of) sensitivity to different layers.

Comment on the vertical sensitivity will be added in revision.

P10 L1-2: Can you investigate more these timeseries and identify when and why they correlate and when not? The reason of the weak correlation needs to be better expressed since to me these timeseries should give the strongest support to your analysis: if they fail to correlate, can you trust the nadir data in the way you adopted them?

The time series figures are no long in the revised paper.

Conclusions

P10 L6- P11 L9: I think the conclusions should better reflect the limitations of the adopted nadir data and the caution needed to deal with them under these circumstances.

See overview of revision for discussion.

FIGURES

General: for some reason, the figures at actual size on screen (especially labels) show poor quality, whereas they have very good quality when enlarged. Could you tune this?

Revised

Fig 1a: I would limit the map to the ASM region. Please add a panel with vertical snapshot – see text. Blue/red is very hard to read. Could you have the symbols with the actual size of the footprints (horizontal resolution for MLS) when limiting the map to the ASM region?

Revised

Fig 1b: increase legend font size. Possibly place the zero aligned. Why are the numbers so different? I would like to see also the same plots for O3.

Revised. See point 2) in overview

Fig 2: it is a bit confusing to have a different color scale for the two. Could you add a remark in the text? Remove “(GPH)” or “geopotential height” from caption as it was already explained. Please mention first “a,b”, then “c,d” in the caption.

Fig 3: could you provide some significance test? Could you add also a comparison of the vertical profiles with their standard deviations too?

Following figures are all revised.

Fig 5 and 6: please join the two figures to allow easier comparison. Fonts at actual size are not readable. Why do you fill missing data? Please let them visible.

Fig 7: could you add an approximate vertical scale in km too since you use both in your text? Could you add the vertical and horizontal extent of the anticyclone? You appreciably pay attention to the horizontal interpolation for MLS data, but to my understanding this plot is produced with only 2 independent points on the vertical. Is this correct? If so, could you add a comment/a sign for this?

Fig 8: could you help the reader over plotting contour of one zero anomaly line on the other? Or of the GPH? In a similar fashion as in Pan et al. (2016).

Fig 10-11: please join the two figures and see comments above. Fig 12: consider over plotting reference contours as in Fig 8.

Relevant technical corrections are made in revised version.

TECHNICAL CORRECTIONS

General: please note that the use of the article “the” is recurrently not consistent throughout the manuscript. E.g., P1 L17: “the ASM UTLS trace gas”, P1 L21 “of ASM anticyclone”, P1 L25 “of the ASM anticyclone”, etc.

P1L13-14: since you first introduce MLS, I would link to it in the following sentence where you state you work on IASI and OMI and avoid the reader to wait to know why MLS was introduced: e.g., P1 L15 “. . . IASI and OMI, IN COMPARISON TO THE MLS LIMB SOUNDER.”

P1 L14: “these type” –> “this type”

P1 L15: “IASI” and “OMI” were not introduced before and needs to be explained

P1 L16: remove “in the UTLS” after “(O3)” as it is repeated at the end of the sentence. P1 L17: “variability is” –> “variability are”

P1 L18: remove comma after “explored” P1 L19: –> “results show”

P1 L26: remove “(GPH)” P2 L4: “;” –> “:”

P2 L12: –> “of the Tibetan mode”

P2 L15: “Asian summer monsoon” –> “ASM” P2 L18: –> “chemical impact.”

P2 L21: –> “are widely used for this purpose”?

P2 L25: this is slightly confusing since in the title you mention “Limb and Nadir...” (limb first), and here you introduce your work stating you work only on nadir data

P2 L26: “Two specific dataset we explore” seems as the reader should expect further nadir data to be used in the study

P2 L28: 2-3 months at what altitude/which layer? P2 L29: “long lifetime” at what altitude/which layer?

P2 L30: “pollution sources” with negligible or not negligible impact? P2 L31: “will be examined” –> “were examined”

P2 L32: remove “the” in front of “MLS”, this is the first time you introduce them P3 L1: remove “the” in front of “UTLS levels”

P3 L15: –> “also aims to”

P4 L17: “which” –> “whose”, “is greater” –> “are greater” P5 L19: “pressure” –> “vertical”

P6 L24: “10-20 day” –> “a 10-20 day”

P7 L9: –> “(Aug 18, when the center of the ...)” P7 L19: –> “Compared to MLS...”

P8 L30: –> “by differenCES over. . .”

P12 L15: George et al reference is wrong

P13 L14: Liu et al misses page/volume number. P13 L 30: McPeters et al. misses page/volume number. Please check other references have the same problem.