Response to the Reviewer 2

H. Sagawa, et al.
sagawa@nict.go.jp

We deeply thank the reviewers for their works and constructive comments on our article. Please find below our general response followed by the responses to each comment.

We found that both reviewers pointed out the weakness in the comparison methodology of the original submitted manuscript. To improve the quality of the paper, we applied the following major updates on the manuscript:

- Ground-based measurements are newly added for the comparison. We would like to include B. J. Connor, who is responsible for the analysis of the ground-based ClO measurements, as a coauthor of this paper. In order to make the paper title consistent with the updated content, we would like to change it to "**Comparison of SMILES ClO profiles with other satellite, balloon-borne, and ground-based measurements**".
- Comparison of the SMILES and Envisat MIPAS ClO profiles has been updated using coincidences of the two datasets while the zonally-averaged data comparison has been removed. With this change, all the comparisons described in this paper are now done with the same comparison methodology, which makes the discussion more straightforward to understand.

We believe these major changes will improve the quality of the article, and they make the paper more useful for readers.

The responses to the individual comments are described in detail below. The original comments from the reviewers are cited with an *italic* style.

#1 Specific comments :

Sec. 2.3 - Error analysis:

p622, Lines 4-5: The reference to the JPL spectroscopic database (Pickett et al., 1998) does not appear to include the details on the measurement of the line position for this ClO doublet. An additional reference should be added to provide this information.

The detail reference for the ClO line parameters on the JPL database can be found in Cohen et al. (1984). We cited this reference in the manuscript.

Sec. 3 - Methodology:

p624, Line 2: I am not sure about the term "a median of the zonally accumulated dataset". Is this a climatological comparison? It seems that the term used to describe the comparison differs through the paper and this should be made more consistent. Also, are there references that can provide additional insight into using this method?

We changed the comparison methodology of SMILES and Envisat MIPAS data to search the coincidence profiles instead of comparing the zonal mean distribution. In the originally submitted manuscript, we had selected the latter approach just because the latest version of the Envisat MIPAS data for the SMILES operated period was not available at the time of the analysis (thus, the coincident measurements were not available). The latest version of MIPAS ClO, covering the SMILES period, became available after the submission and has been used for revising the paper. The differences in the measurement sensitivity of MIPAS and SMILES are taken into account considering the averaging kernels of both instruments.

p624, Lines 8-10: Could you provide a reference or further explanation about this "difference in sensitivity" issue that excluded the ground-based comparisons?

Our original statement was not clear. The ground-based ClO measurements are generally analyzed after integrating the data from several dates, and the subtraction of nighttime spectra is also required. We intended to state these facts as "difference in sensitivity" in the original manuscript. We were also considering the low vertical resolution of the ground-based instruments compared to the limb sounder. Because of all these points and because of the good agreement between the limb sounder datasets, we believed that the comparison with ground-based was not necessary. However we have decided to add tropical ground-based observations in the revised paper in order to make the paper more useful for readers.

p624, Lines 13-16: Are there some general comments that could be made about the coincidence criteria at this point? The justification for the choices was not always clear in the discussion in Sec. 5. On Page 624, Line 15, what is meant by "representative values"? Is this the difference of the median profiles for each instrument or the median of the individual profile differences? The calculation *methods seems to be specified for the average relative difference calculation (Lines 19-22) but is not as clear for the average absolute difference.*

The criteria were determined somewhat empirically in order to keep a certain numbers (not smaller than 50) of coincidences. For the comparison of SMILES and other satellite-borne instruments, we set the basic criteria as the measurement geolocation distance closer than 300 km (which is comparable to the horizontal resolution of limb-geometry observations), the measurement time difference within 1 hr, and the solar zenith angle difference within 3 degree. For the case of Aura MLS, thanks to its dense measurement sampling, we could put more stringent criteria. We added this explanation on the manuscript.

The word "representative values" means "the median of the individual coincident ClO profiles of the two instruments". We improved the sentences on the manuscript, in order to make the meaning clear.

Use of Averaging Kernels and Vertical Resolution:

Page 624, Lines 25-26 and further: The term vertical resolution is used here and further on in the paper. However, it is not defined. Is this the FWHM of each averaging kernel? Is the same definition used for each instrument or are these values (and the profiles shown in Fig. 5) obtained using different metrics? A consistent definition should be used.

The vertical resolution presented in this study is defined by the FWHM of each averaging kernel. We added the explanation on the text.

Page 625-626, Lines 25-2: For the smoothing calculation, it is not clear if a representative averaging kernel is used or if individual ones are used for each coincident pair. This should be clarified.

We have used the individual averaging kernel for each measurement. We added the explanation of the text.

Page 626, Lines 11-14: It was not clear why the triangular functions were used for smoothing for the SMILES/MIPAS zonal median distribution comparisons. Were these applied before or after the zonal median calculation? Were the averaging kernels not available or were only the zonal median distributions provided? It would be good for the authors to clarify the reason for this approach and the impact of using versus not using the smoothing for this comparison.

We consider that this question can be skipped now as we changed the method of comparing SMILES and MIPAS data. The comparison based on coincident data take into account the actual averaging kernels of the MIPAS instrument.

Sec. 4 - Comparisons between SMILES versions:

Page 627, Lines 19-20: Could you be more specific about the differences in the a priori temperature and pressure profiles used for the two retrievals? Are these "first guess" for a retrieval of pT or a fixed input for the ClO retrieval?

They are fixed inputs through the ClO retrieval. The manuscript is updated.

Page 628, Lines 8-9: Is there a reason why the JAXA vertical resolution is higher?

We consider that there should be several reasons for this: for example, the largeness of the a priori error (JAXA uses a priori error of 100% of the a priori ClO profile while NICT uses ~60%), the usage of the full spectral bandwidth (which gains more sensitivity at upper troposphere), and probably different definition on the measurement noise (NICT uses a constant 0.5 K as the measurement noise brightness temperature which is larger than the theoretical value in order to take into account additional errors such as low baseline ripples). The first mentioned difference in the a priori error is now added in the text.

Page 629, Lines 1-16: It would be helpful to clarify which spectroscopic parameters were changed in this experiment and how they were changed. Was it all of those listed in item 3 on Page 627? Were they changed one at a time or all together? Further information would be helpful to provide justification for the comment on Lines 15-16.

We tested one by one of the parameters listed in item-3 on p 627 (i.e., line_frequency, gamma_air, and n_air), and also tested the case of changing them all together. The impact at around 1 – 2 hPa described in the text is the case with the maximum error which was obtained when all parameters were changed.

Page 629, Lines 22-25: Have any tests been done to confirm this suggestion about the limited spectral bandwidth? It would be helpful to be able to describe a preliminary result if it is available.

In order to describe the impact of limited spectral bandwidth, we added a short appendix with a figure showing theoretical sensitivity using the full spectral bandwidth (1.2 GHz) and the limited one (400 MHz). Please see Appendix A on the revised manuscript.

Page 630, Lines 9-12: I am not sure I follow why the latitudinal shift in daytime stratospheric ClO creates a difference in retrievals made from the same set of observations. Could you clarify this point?

Our original statement was based on the interpretation that more ClO abundances give more systematic differences between NICT and JAXA level-2 processing therefore the amplitude of the differences follows the latitudinal distribution of the ClO abundances. Now we found that this interpretation requires further analysis; and, since this section is not the main purpose of the presneted paper, we decided to remove it from the revised manuscript.

Sec 4.2.2: In the earlier discussion, it was noted that the JAXA retrievals were smoothed by the NICT averaging kernels for this. Is there any discussion of the impact of this? It would be useful to quantify the differences.

We have compared the JAXA and NICT profiles before and after smoothing, respectively. Manuscript is updated to explain the impact of the smoothing.
On the original submitted manuscript, a typical averaging kernel was applied to all the profiles. Now we updated to use individual averaging kernel for each profile. And also we revised the data selection algorithm in order to be sure to select the data from inside the vortex. Thus, the comparison result slightly changed from the original version.

Page 630, Lines 15-17 and 21-23: It is not clear what is meant by "MLS-derived meteorological products"? Were they derived for MLS rather than SMILES observations?
Also, how was location of the SMILES profile relative to vortex determined? Did it use data at one altitude, at the ClO peak or throughout the stratosphere? How did the change in SZA criterion for daytime profiles (between lower latitude and polar observations) affect the comparisons? What was the change in the number of profiles available and the median difference? It would also be instructive to examine the impact of this change in SZA criterion using the lower latitude profiles.

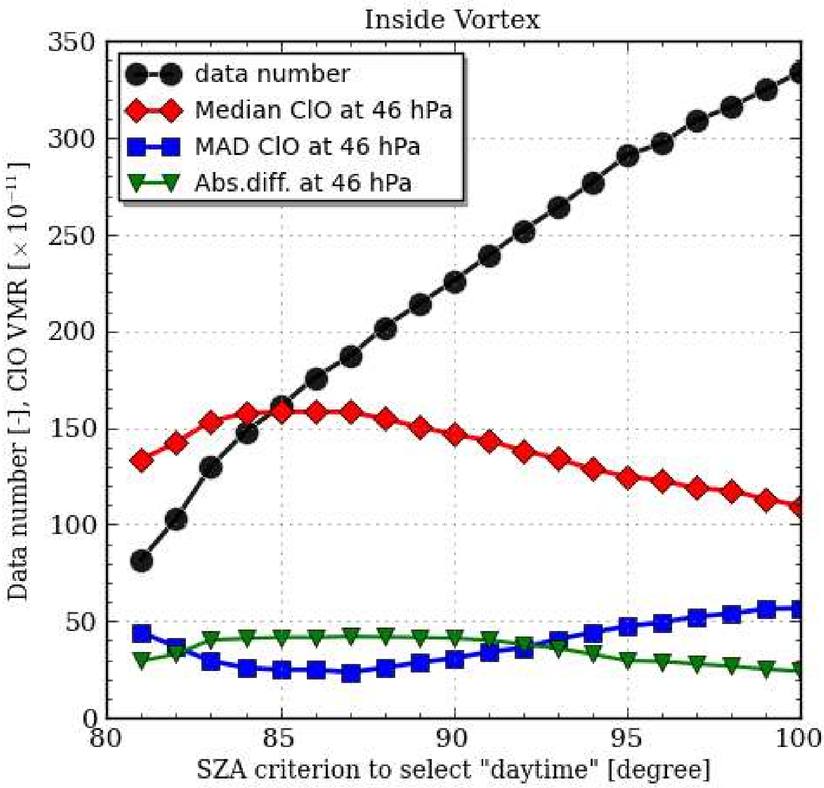
What we used is the information of vortex-geolocation (latitude and longitude) which have been computed in the frame work of the MLS Derived Meteorological

Product (DMP) processing (Manney et al. 2007). For each of the MLS measurement, the MLS DMP provides the distances (in the equivalent latitude coordinate) of the measurement points from the edge of the vortex. By using such information and the geolocation information of the SMILES measurements, we estimated the distances of SMILES measurements from the vortex. The simple reason of using the "MLS"-DMP is because there is no "SMILES"-DMP at the moment.

We used the vortex information at the potential temperature level of 520 K, which approximately corresponds to ~19 km (close to the lowest altitude of the SMILES NICT ClO retrieval). As we wanted to be sure in selecting the SMILES data measured inside the vortex, we selected the data which were located within the "inner" vortex as defined in the MLS-DMP.

The change of the "daytime" definition (SZA < 87 degree, instead of 80 degree) was needed to keep enough data for the comparison. We determined this criterion based on the following requirements: (1) number of the selected data inside the polar vortex, and (2) amplitude of the chlorine activation in the median ClO profile in order to examine the worst case for biases (systematic errors such as the error due to spectroscopy are proportional to the ClO concentration).

The supplemental figure on the next page shows how these conditions change when different SZA criteria are applied for the daytime data selection. Increasing the SZA limit, of course, increases the data numbers; but this means that nighttime measurements are also included into "daytime" dataset, and therefore the median ClO peak becomes lower and systematic errors become lower. Actually, a SZA around 85 – 88 degree seems to be a good choice for selecting the "daytime" measurements. Considering the large horizontal resolution of the SMILES limb sounding measurements (several hundred km), we decided not to go too close to 90 degree to stay away from the terminator. These are the reason why we selected 87 degree for the daytime data selection criterion. Note that we found the absolute difference between NICT and JAXA ClO profiles is rather constant (0.4 ppbv) for any SZA criterion between 83 – 89 degree.



Supplemental figure: Change of the selected data number inside the polar vortex (shown in black circles), median peak VMR of ClO at 46 hPa (red diamonds) with its deviation (median absolute deviation, MAD) (shown in blue squares), and the median of absolute differences between NICT and JAXA level-2 processing (green triangles). Note that the MAD increases when the SZA limit is increased to be higher than 90 degree; indicating the mixture between daytime and nighttime ClO .

Sec. 5 - Comparisons with Other Instruments

Sec. 5: For each of the instruments, slightly different sets of parameters are given. To provide the appropriate background for the reader, it would be best to include consistent information (e.g. antenna dimensions, spectrometer resolution, spectral range or bandwidth, number of measurements per day, measurement duration/integration time, scan altitude range). Also, no description is given (in this section or in the SMILES data descriptions) of the vertical coordinate used for each of the retrievals and the method used to change all of the observations to a common grid. For example, in Fig. 11, the change in vertical coordinates is obvious between the three panels but it is not discussed in any detail in the text.

We revised the manuscript with the intention to be concise and to provide the relevant references for each instrument where the details can be found.

Information about the vertical coordinate of the ClO retrieval is added in the relevant subsections for each instrument. Figure 12 (we consider the above commented Fig. 11 is typo) is now re-plotted in the pressure coordinate.

Page 634, Lines 1-3: It is unclear if the time criterion is also changed in these comparisons (versus what was used for MLS).

Only the geolocation distance criterion is changed. An explanation is added in the text (in the methodology section).

Page 634, Lines 3-7: If the SMR measurements cluster near local sunrise and sunset, why were the comparisons not split this way as the day and night ones were for MLS? Does this make a difference in the comparisons? Did the coincidence criteria also ensure that the SMILES and SMR measurements were taken on the same "side" of the sunrise or sunset (e.g. both before or both after)?

The rapid ClO decrease/increase around the Odin-sampling local times (sunset and sunrise) induces significant errors in the coincident comparison. Such (random) errors can be reduced by increasing the number of data. This is the reason why both sunrise and sunset coincident data are analyzed together (and also the reason of not requiring the both instruments to be at the same side (before or after) of the terminator). It is also commented that the amount of ClO and its variation rate are similar in both sides.

As additional information, we would comment that the recent paper by Khosravi et al. (2013) shows the comparison between SMILES and SMR using a different method compared to the presented analysis (i.e. coincidence comparison): they compared the two datasets for each local time separately (Fig. 3 on Khosravi et al. (2013)). Their result did not find any significant differences between the sunset and sunrise conditions.

Page 634, Lines 14-21: Were the SMR averaging kernels used to smooth the SMILES result for the mesospheric comparisons? An experiment to look at the effect of smoothing should be considered in revising the comment made on Lines 20-21.

We don't consider the smoothing for the SMILES and SMR comparison because their vertical resolutions are similar as shown in Figure 3. The observed negative VMR in the SMR ClO data at mesosphere is not due to the change of the limb scanning scheme, but most likely due to the limitation in the SMR instrument (e.g., systematic bias coming from the spectral baseline problem). To avoid the confusion of readers, we removed this sentence from the manuscript.

Page 635, discussion of MIPAS data versions and errors: It is not clear which version (high-resolution or reduced-resolution) data is being described and used. In particular, the "offline-reprocessing high-spectral-resolution version" description (Line 21) does not seem to match with the time period and description of when MIPAS measured with reduced spectral resolution (Lines 12-14). This should be clarified further.

"Offline-reprocessing high-spectral-resolution version" was our mistake. It should be "Offline-reprocessing high-spatial-resolution version". In order to avoid the confusion, we removed "high-spatial-resolution" from the text.

Sec. 5.3 methodology: There are some details on this zonal median distribution comparison that should be clarified. The pixel size (latitude by pressure levels or altitude bins) needs to be defined (here or Page 262, Line 5). Why were the two data sets not sampled over the same/more similar time period to remove this issue in the comparison?
Could the data have been filtered into vortex and extra-vortex measurements to reduce complexity for the Arctic? This technique would allow the comparisons at higher and lower latitudes to be made independently and could help with some of the more challenging issues. I think that these comparisons could be useful but need to have a number of uncertainties and inconsistencies reduced to better understand the differences.

We consider that this comment can be skipped now as we changed the method of comparing SMILES and MIPAS data. The revised comparison is limited to the mid-latitude daytime conditions.

Balloon Comparisons
Sec. 5.4 measurement details and discussion: There is a lot to take in from Fig. 10 and it would help to provide a bit more information on the scan range, rates and other measurement characteristics of TELIS and MIPAS-B. The measurement location for each instrument is not provided at the same altitude and none seem to match the level used for the PV field. Could some comment be made on the "mismatch" in the comparison arising from this?

A concise description of the TELIS and MIPAS-B instruments together with the data analysis is given in Section 5.4. We newly added some information on the vertical scan range and rate. For space-borne instruments like SMILES and MLS, the latitudinal and longitudinal drifts of tangent points during a single vertical scan are very small compared to those of TELIS and MIPAS-B. For example, the tangent points of SMILES observation moved no more than 0.4 degree in latitude when SMILES scanned the limb from 15 to 35 km (altitude range considered in this section). Thus we plotted only one geolocation for each scan of SMILES and MLS. We removed unnecessary SMILES observation points from the plot in order to make the figure less busy.

The PV field is plotted for the purpose of general understanding on the polar vortex at the time of the observations. We selected the isentropic surface of 530 K from the ECMWF defined isentropic levels, which is the closest level to the peak altitude of ClO enhancement (around 28 hPa). From the PV field map, one could expect that there is a significant difference in the atmospheric states between the TELIS and MLS measurements (as MLS measured the atmosphere of larger PV compared to TELIS). However, the two selected profile of SMILES (one is close to the MLS measurement's geolocation and the other is close to TELIS tangent points) showed a consistent ClO profile. This may suggest that the PV difference presented on Fig. 11 of the PV between the TELIS and MLS is not a critical issue in the comparison. Moreover, as we described in the manuscript, the relatively poor horizontal resolution of MLS and SMILES should be taken into account when discussing the impact of the PV inhomogeneity.

Page 639-640, Lines 27-5: It is not clear which specific SMILES and MLS profiles are being considered as having high gradients and contributing to difficult comparisons. I think that this section would benefit from the use of output from a chemical transport model to assist with the interpretation, as mentioned by the authors.

Our intention is that both of SMILES and MLS are affected by the high gradient of the atmospheric variability at the edge of the polar vortex because they both have horizontal resolutions of several hundred km along the line-of-sights. We fully agree with the reviewer's comment that the chemical transport model is needed for further convincing discussion on this comparison. But we still would like to keep the current content of the section as it is because it shows a good agreement between the four instruments (SMILES, MLS, TELIS, and MIPAS-B) for the first order comparison.

Conclusions:

Some care should be taken to revisit these after doing further work to ensure that they are not over stated in terms of "good agreement". It is better to quantify what is meant. The last statement sort of downplays the usefulness of the current dataset (issues that are "not fatal"). If it will get better in the next version, what is the appropriate use for this version?

Description is improved following this reviewer's comment.

The future versions of the NICT L2 processing is planned to include the lower stratosphere in the retrieval analysis. Because of the good agreement between NICT and JAXA products from the middle stratosphere to the mesosphere, we believe that the new version will not significantly improve the results at these altitudes. The conclusion of the paper is that this version of NICT SMILES data can be used above ~25 km for scientific analysis. This has been clearly stated in the conclusion.

Tables & Figures

Table 1: This needs a bit more work with the text to explain where the values came from. Did all of these values come from Sato et al., 2012? There need to be some more references. Also, some listings do not describe an assumed uncertainty they describe the two configurations that are contrasted. These should be made more consistent.

We updated the description and references on Table 1.

Table 2: Consistency in the capitalization of headers needs to be fixed. Also, is precision the term used mostly in the text or is it random error. One term should be used consistently.

The first letters are all capitalized.

Also we did the checks on the homogeneous-wording through the manuscript.

Fig. 2: Which ozone isotopologue appears in this spectral window? It would be helpful to explain that all of these analyses came from one scan (date, time, location). Are the averaging kernels for altitudes or pressure levels?

Generally, the figures need a bit of work to make sure that the captions provide sufficient information and that they do not contradict the text.

It is ^{17}OOO (O_3-asym-O^{17}) at 649.275 GHz. We updated the figure caption.

The retrieval processing is carried out in an altitude grid. In Figure 2, we plotted them in a pressure grid since we wanted to make comparison studies in pressure levels. This is now explained in the manuscript.

#2 Technical Corrections and Suggestions for Improvements:

We considered all of the technical corrections pointed out by Reviewer-#2, and reflected them to the manuscript revision. We believe the updated manuscript is now improved a lot. Below we put our responses to the comments which may require some specific answers.

Page 618, Lines 24-25: Does the direction of scan influence the results or is this not necessary for this discussion?

The direction of the scan has no impact to the studies discussed in this paper.

Page 621, Line 25: It is not clear in the latitude range that was used for the nighttime assessment. This should be clarified.

Mid-latitude case was used. This was clarified in the manuscript.

Page 625, Lines 9-10: Is the reduced FOV for TELIS due to being in the atmosphere or really because the instrument is closer to the tangent altitude being measured?

The submm antenna of TELIS has a diameter of 40 cm in the vertical direction, which is an equal size with SMILES. Being within the atmosphere makes TELIS closer to the tangent point; and this means that the same viewing angle results in a smaller vertical resolution at tangent height for instruments which are closer to the tangent point.

Page 627, Lines 6-8: Is the offset parameter retrieved in this process?

Yes. We retrieve a global frequency shift by fitting the line position of ClO.

Page 629, Lines 19-20: Is it more appropriate to discuss this here once the JAXA *comparisons have been introduced?*

We consider that the discussion about the "actual" (not forward-model simulation based) negative bias in the nighttime lower stratosphere is important at the error analysis section. We would like to keep the current manuscript.

Page 634, Line 17: Do you mean -0.05 ppbv rather than -0.05 hPa?

It was a mistake in the original submitted manuscript. Thanks for the correction.

Page 634, Line 28: Which is the ascending and descending equator crossing time for Envisat?

10:00 a.m. is the descending node. We updated the manuscript.

Page 639, Lines 15-20: It might be nice to show the difference plots as well as the individual profile comparisons.

In order to avoid the increase of numbers of figure (with a similar content), we decided to embed one sub-plot showing a closeup at the ClO-enhanced altitude. We consider that the visibility of the plot has been improved.

References:

Manney, G. L. et al. (2007), "Solar occultation satellite data and derived meteorological products: Sampling issues and comparisons with Aura Microwave Limb Sounder", Journal of Geophysical Research, **112**, D24S50, doi:10.1029/2007JD008709.

Khosravi, M. et al. (2013), "Diurnal variation of stratospheric HOCl, ClO and HO_2 at the equator: comparison of 1-D model calculations with measurements of satellite instruments", *accepted to Atmos. Chem. Phys.* (the discussion paper was published as *Atmos. Chem. Phys. Discuss.*, **12**, 21065-21104, 2012)