Response to anonymous referee #1

We thank the reviewers for their constructive comments and suggestions which helped us to improve our manuscript. We have addressed their questions as follows:

This is a well-written and detailed manuscript describing initial spectral fitting results from the TROPOMI sensor on Sentinel 5P. The images in the manuscript show dramatic improvement in identifying small-scale features that are allowed by the higher resolution of this new satellite sensor. The authors nicely describe their methods and selections that were made to avoid biases by other absorbers. I feel that the manuscript is appropriate for AMT, but have a few comments that I think should be addressed in a revised version.

Broader comments:

The authors show that formaldehyde causes a particularly challenging species to retrieve along with BrO, particularly when using narrow spectral windows. In the end the authors make what appears to be a good compromise to avoid the larger absorbers and other issues, but one really wants to quantify all absorbers and not simply try to avoid them and then hope that the selection was good enough. Therefore, I think the manuscript should indicate that this window is a good choice, but future work needs to be done to continue to improve quantification of all the absorbers that occur. This window is a good start, but quantitative results will be improved by further research, which should be encouraged in the manuscript and conclusions.

We have added a few more sentences that emphasize the need of further studies on the optimum fitting window and DOAS settings to solve the problems presented in this study as you suggested:

In section 3.1.6 (P12 L9 in the revised manuscript):

Although the choice of the optimal fitting window may seem arbitrary to some degree, the analysis of several different relevant scenarios for many possible combinations of fitting windows described above demonstrates that this is an overall robust selection.

{+ "However, further studies are needed to address the remaining challenges identified through the sensitivity tests, in particular the possible spectral cross correlation of BrO with HCHO around the selected fitting window."}

In conclusion (P24 L23):

Based on the information gathered from the sensitivity tests for different measurement scenarios, 334.6-358 nm was selected as the optimal wavelength range for TROPOMI BrO retrievals for most of the possible measurement situations.

{+ "This fitting window yields reliable BrO retrieval results with small fitting errors, but future studies on quantitative assessments and cross correlations between BrO and all the interfering absorbers are encouraged to further improve retrieval results."}

The manuscript shows dramatic images of small scale BrO features that are amazing to see and will increase our scientific understanding. However, it is important that the manuscript points out that albedo effects are really important for quantification of lower-tropospheric BrO. Therefore, some things that are evident in the images using the simple geometric air mass factor should be pointed out as likely due to albedo effects. The reason for this is that I worry that a reader less aware of these issues may not take proper care of the albedo and misinterpret the images. An example that I think is likely an albedo effect is the sharp edge of small BrO evident North of Scandinavia/Russia in Figure 9a. That feature could be caused by the sea ice edge (http://nsidc.org/arcticseaicenews/2018/04/), with the high albedo of snow/sea ice allowing you to observe tropospheric BrO, while the low albedo of open ocean tends to prevent observation of lower tropospheric BrO. These albedo effects may also be important in some of the Rann of Kutch observations shown in Figure 13. The authors should point these effects out and indicate that they need to be taken into account for quantitative work.

We agree with the reviewer and have revised and added more text on uncertainties and biases of retrieved BrO vertical columns applying geometric AMFs in several parts of the manuscript:

In section 4.1 (P15 L4):

“The number of photons detected at the satellite is larger over bright surface areas than over dark surface areas. Therefore, the use of a simple AMF which does not consider the sensitivity to surface albedo can underestimate BrO vertical columns over dark surfaces such as the ocean in comparison to high surface albedo regions such as the Antarctic region, north of Russia and Canada. In addition to the surface albedo, clouds also affect signals detected at the satellite. The light path length and intensity are significantly changed by the cloud top height, cloud thickness and cloud fraction. Using an AMF that does not take into account the cloud effects can therefore result in errors in the computed vertical columns, as can be seen from the slightly higher BrO VCDs in the subpolar regions where cyclones are frequently observed due to the sub-polar low pressure system (Fig.11). Consequently, an improved AMF reflecting the sensitivity of surface albedo, cloud properties and BrO vertical profile should be calculated to obtain more accurate vertical column densities (Theys et al., 2011; Sihler et al., 2012), and this will be investigated in detail in a follow-up study using surface albedo and cloud information from the operational satellite products as they become available.”

In section 4.3.1 (P20 L10):

“Convective clouds can be formed around open leads due to the supply of water vapor and enhanced vertical mixing, but computed BrO enhancement over clouds may have an error because of the use of AMFs which do not consider the effects of clouds.”

In section 4.3.2 (P22 L16):

“It should be noted that the AMFs used in this work do not consider surface albedo and cloud effects, and therefore BrO VCDs may be overestimated over the bright salt marsh.”

In conclusion (P25 L20):

“In this demonstration study, a simplified air mass factor and no stratospheric correction were applied because the main purpose of this study is to find the optimal DOAS retrieval settings for BrO that reduce systematic biases by minimizing effects of interfering absorbers and to assess the consistency with previous satellite results. However, for future quantitative studies of tropospheric BrO explosion events, stratospheric correction and improved air mass factor calculation taking into account the observation conditions are essential. In particular, investigation and evaluation of high resolution input data applicable to the unprecedented small footprint of TROPOMI should be performed, which will be a subject of further work.”

The comparison of these retrievals to those of OMI and GOME-2B assist in assuring consistency between these different satellite sensors. However, they do not assure "accuracy" of the measurements; they show consistency. Therefore, on page 15, line 7, it should say "To evaluate the consistency of TROPOMI BrO..." Additionally, on page page 2, lines 26-28, the text should read "TROPOMI BrO columns were compared with those from the two existing satellite instruments, GOME-2B and OMI, with the consistency of these data sets were investigated." These comparisons do not either "validate" this method, nor do they assure "accuracy". They are more "precise" and "consistent", which can be pointed out and are nice advances in the field.

We agree with your comment on the problem of “accuracy” word usage because we didn’t validate our retrieval results using independent observations but identified consistencies of BrO retrievals between TROPOMI and previous satellites. Thus, the term “accuracy” has been removed in the revised manuscript as you suggested.

P2 L27: TROPOMI BrO columns were verified -> TROPOMI BrO columns were assessed

P2 L28: with their accuracy -> removed

P17 L7: To evaluate the ~~accuracy and~~ consistency -> To evaluate the consistency

I believe from the text on page 15, line 8 that the authors have used spectra from OMI and GOME-2B and retrieved BrO using the same processing methods applied here for TROPOMI data to compare with their TROPOMI measurements. Both OMI and GOME-2B have operational products, but my reading of this sentence is that those operational products were not used in this consistency check, but a single retrieval method (described here and using Table 2 settings) was used on spectral data from the three satellites. It needs to be made more clear whether spectra (Level 1) data or slant column (Level 2) operational products are being used. The reason it needs to be made more clear is that one might read that this paper says there is good correlation between this method and OMI, and take that to mean that the OMI operational product (OMBRO) agrees with this product, which is not what I think is plotted in Figure 10 panels. Alternatively, if the operational OMI and GOME-2B products are being used, they don’t appear to be cited appropriately. Any references in the abstract and/or conclusions also needs to make clear what this "consistency" check is.

We have revised the text to make clear the data used in the intercomparison of satellite retrievals as:

P14 L14: “Applying the retrieval settings described in Table 2, BrO vertical columns have been computed from TROPOMI, OMI and GOME-2B Level 1 spectra.”

P17 L8: “To evaluate the consistency of TROPOMI BrO retrievals with those from other satellites, a comparison was performed using GOME-2B and OMI retrievals obtained by applying the same retrieval setting (Table 2) to Level 1b data.”

Specific comments:

Page 1, line 18: This should say "...reactions that deplete ozone..." (replace "which" with "that").

This has been corrected in the revised manuscript.

Page 2, line 27: Remove "verified" and "accuracy". Clarify if you are using operational products or re-analysis of spectra.

Thanks for pointing this out. We have removed the expressions “verified” and “accuracy” in the revised manuscript.

P2 L27 TROPOMI BrO columns were verified -> TROPOMI BrO columns were assessed

P2 L28: with their accuracy -> removed

P17 L7: To evaluate the ~~accuracy and~~ consistency -> To evaluate the consistency

Page 3, line 2: Should say either "... which allows it to monitor..." or "... which allows monitoring of ..."

This has been corrected in the revised manuscript.

Page 10, line 15: The word "dependency" should be "dependence"

This has been corrected in the revised manuscript.

Page 12, line 2: Although this is a "robust selection", it would be valuable to point out that further work is warranted, particularly to retrieve both HCHO and BrO.

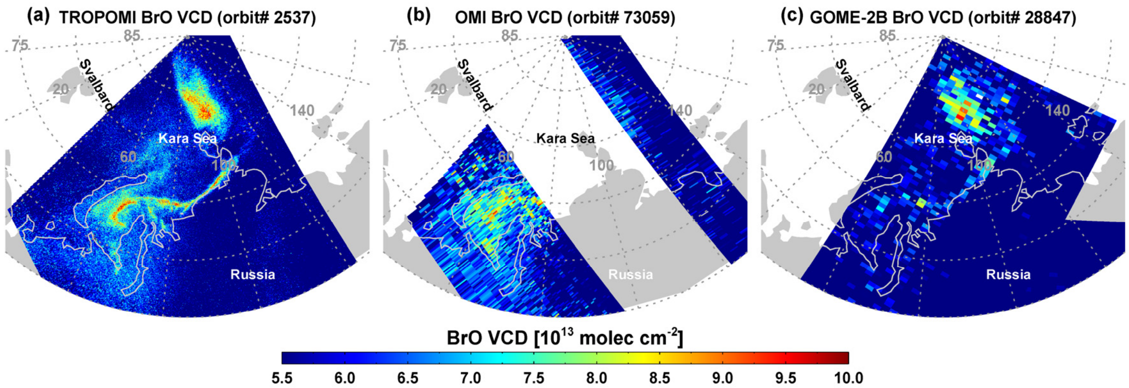
We have added more text to emphasize the need for further studies in section 3.1.6 as:

“However, further studies are needed for the remaining challenges identified through the sensitivity tests, in particular the possible spectral cross correlation of BrO with HCHO around the selected fitting window.”

Page 17, Figure 11: As a courtesy to the reader, please point out the geographic region where this coastal edge occurs. I think it is the Kara Sea in the Russian Arctic. Potentially pointing out some island or landmass would help.

Figures have been modified to high resolution with adding texts denoting the region in the revised manuscript.

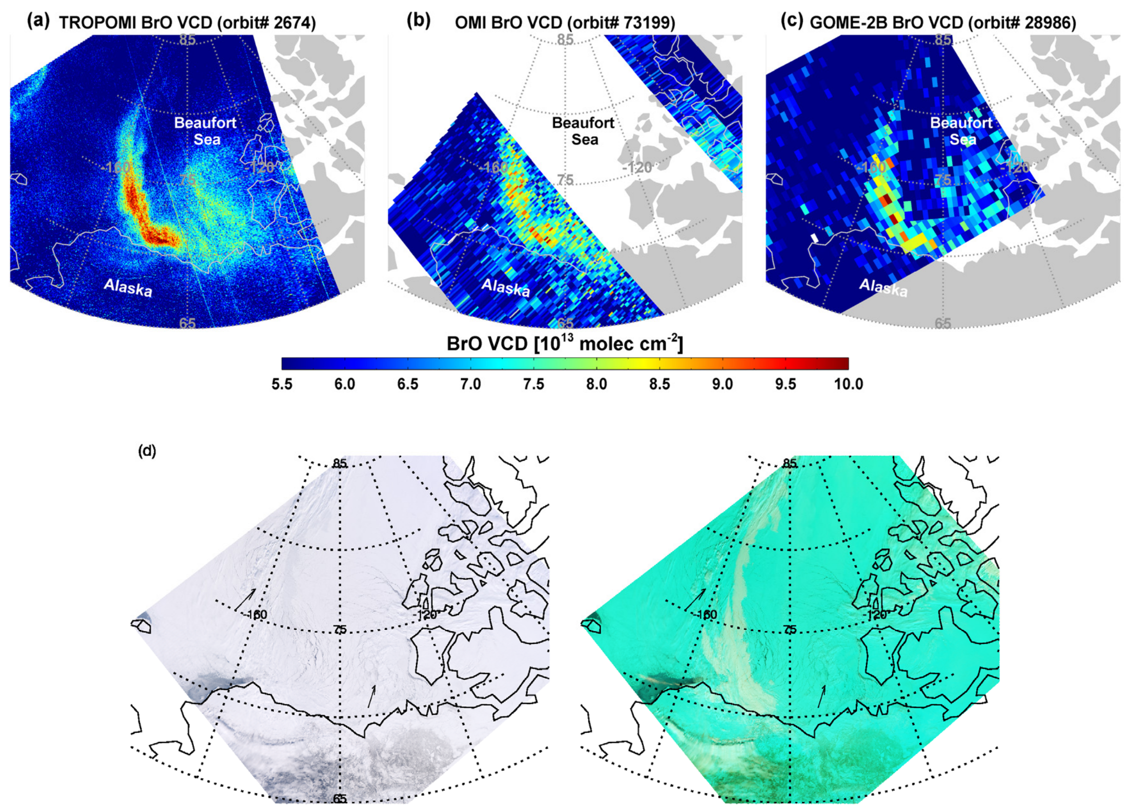
- Figure 13 in the revised manuscript:



Page 18, Figure 12: Again, it would be nice to see where this image is – I think it is the Beaufort Sea, North of Alaska.

Figures have been modified to high resolution with adding texts denoting the region in the revised manuscript.

- Figure 14 in the revised manuscript:



Page 20, Line 26: It is possible that using wavelengths longer than 362 nm had poorer results because of poor fitting of O4. O4's spectrum is temperature dependent and a warm temperature was chosen. If you cut the wavelength region at 362 nm, then the polynomial can remove some aspects of the O4 absorption, while if you go beyond the peak, it gets harder for the polynomial to fit the features.

Thanks for pointing this out. We have further mentioned potential problem of O4 fitting.

In section 3.1.6 (P11 L23):

"In the case of DOAS retrievals over the cloudy background region and the volcanic plume, higher fitting errors were found in the wavelength intervals extending beyond 362 nm because of imperfect correction of the Ring effect and possibly also poorer fitting of O_4 related to the temperature dependency of the cross section."

In conclusion (P24 L20):

"At wavelength longer than 362 nm, poorer results were found in the cloudy scene and volcanic plume measurement scenario, presumably due to the wavelength dependency of the Ring effect and imperfect fitting of O_4."