Referee Report to "Joint spectral retrievals of ozone with Suomi NPP CrIS augmented by S5P/TROPOMI" by Edward Malina et al.

The manuscript presents an algorithm to retrieve vertical profiles of ozone by combining IR measurements from the CrIS instrument with UV measurements from TROPOMI. The idea has already been explored and results has been published by several authors. The authors highlight that their algorithm differs from the previous ones by a usage of a reduced spectral information, which makes the algorithm less demanding with respect to the calibration of the data reducing, however, substantially the information content in the stratosphere. As a result, the obtained improvement over the pure CrIS retrieval is often vanishing small. The authors must be honest and clearly state this in the abstract. Although the data from other combined CrIS-TROPOMI retrievals are available, the authors make no attempts to make comparisons. Instead many of quite useless comparisons are presented, e.g. with the total ozone product from GODFIT OFFL TROPOMI or ozone field from chemical reanalysis models. The paper is very lengthy and quite boring because of many similar plots (Figs. 10–14), which do not provide much information. In general, however, the obtained results might be useful for certain applications and the manuscript can be published after a major revision.

Major comments

- Please be objective in the abstract and conclusions. You completely ignore the fact that for some conditions the results from CrIS-TROPOMI are worse compared to those from CrIS retrieval. Considering only minor differences between the results from CrIS-TROPOMI and CrIS, as presented in the manuscript, the statement "These results demonstrate that CrIS/TROPOMI retrievals have the potential to substantially improve our understanding of ozone." is a clear over-rating.

- The message of the comparisons with a lot of different data products for one single day performed in Sect. 5 (Figs. 10–14) is unclear. As the declared objective of the paper is to present advantages of the CrIS-TROPOMI retrieval, 2-3 comparisons of this kind would be sufficient. Much more interesting would be an analysis of the time evolution, which is completely missing in this part of the paper. Please provide some time evolution plots or at least the plots for different seasons as it is done for ozonesonde comparisons.

- The goal of the comparisons with chemical reanalysis models is totally unclear. The authors state these comparisons cannot be treated as validation, so what exactly is the purpose of this comparisons? The purpose of the comparisons with the total ozone data from GODFIT OFFL TROPOMI is also unclear. Why do you compare the total ozone but not compare the profiles from other CrIS-TROPOMI retrievals? I think the manuscript would largely benefit if you make it shorter by removing Sects. 5.3 and 5.4.

- The comparison with ozonesondes is shown only up to 100 hPa although the ozonesondes deliver reliable ozone profiles up to about 20 hPa. Please present the comparison for the entire altitude range covered by ozonesondes.

- The information presented in Sect. 6.1 has already been published by Mettig et al. (2022) and does not need to be repeated. The fact that the information content increases when using UV bands is generally known. Furthermore, the reliability of an investigation with any assumed settings without having a working retrieval is questionable as adding new spectral ranges requires often an optimization of the settings to keep the retrieval stable, i.e. it is unknown if the retrieval of the real data using the assumed settings is possible. A pre-condition of having a perfect calibration accuracy, as assumed by authors, is never satisfied in the reality. For this reasons, I recommend to skip Sect. 6.

- Line 659: "Focusing on comparisons with MLS the stratospheric 'gold standard' on August 12th 2020, a linear slope of 1.029, intercept of -7.9 DU ($\sim$3%) and correlation coefficient of 0.952 are found, highlighting the quality of the retrievals." - A comparison for one day is definitely not enough to make any robust conclusion. Please extend the comparison with MLS to a longer period similar to the comparison with ozonesondes.

Minor issues

- Line 233: Please define the scaling matrix $\mathbf{D}$

- Line 234: " with large λ values prioritising the speed of the convergence, but making the steps more non-linear, while small values reduce the speed of the convergence, and is more linear in the iteration, similar to the conjugate gradient method." - In my opinion it is other way around. Small values of λ result in the Gauss-Newton method, which converges faster but is more non-linear, while large values of λ result in a gradient descent update.

- Figure 1: Please comment why single CrIS and TROPOMI retrievals are necessary before the joint retrieval.

- Line 255: "...as only one FoV per observation from the CrIS cross track position is used in the processing, the impact of clouds will be less pronounced."- please clarify why the impact of clouds is less in this case

- Table 6: Please explain how the cloud fraction is used in the retrieval. Please give some details how albedo is retrieved and what the orders mean, is it wavelength dependence?

- Line 277: "...pixels within a 20 minute time frame (where Suomi-NPP and S5P pass the same scene within 10 mins)" - Please clarify why passing the same scene within 10 minutes results in 20 minutes time frame.

- Line 279: "From the current sounding subset, select all pairs that are within < 50 km distance, and 4) select the pair that has the minimum distance." - what is the reason first to select all pairs within 50 km distance and then select the pair with a minimum distance? I expect that selecting a pair with minimum distance skipping the intermediate selection of all pairs within 50 km distance should have the same result. For the pair with minimum distance it can then be checked if the distance is within 50 km.

- Line 283: "... additional steps with respect to other target gases do occur in the pipeline, but are not highlighted here." - This text does not provide any information. Please skip it if the over gases are not related to the ozone retrieval or provide more details otherwise.

- Line 286: "Relevant cloud properties..." - Please clarify how the clouds are handled within the ozone retrieval. Is scattering within clouds considered?

- Figure 2: Please explain how NESR is defined.

- Line 316: "...suggesting that CrIS are subject to larger fit errors,..." - This is not a correct conclusion as mean RMSE for CrIS is the smallest.

- Figure 3 and all figures below having the pressure as the vertical axis: 1) upper limit of the pressure axis must be indicated 2) providing a second y-axis in km would help the interpretation of the results and facilitate the comparison with the results of previous publications;

- Line 321: "In general, longer wavelengths have greater sensitivity in the lower troposphere whereas shorter wavelengths are more sensitive to upper tropospheric ozone." - Please precise which wavelengths and pressure levels you are talking about.

- Equation 3: In the formulation chosen by authors the term "δ_{cs}" must not appear as it is implicitly contained in the $A[x_{true} - x_a]$ term. The citation to (Fu et al., 2018) is inappropriate as they just use the formula form (Worden et al., 2007) without any explanation and misinterpret the notations of (Worden et al., 2007), where the main term was written as $A_{xx}[x_{true} - x_a]$, i.e. included only a sub-matrix of A related to the main parameter. In accordance with Sect. 3.1 of the manuscript under review both A and x contain all retrieved parameters. Thus no additional cross-term must appear.

- Equation 6: As above, the last term is a natural part of the first term and must not appear here.

- Line 363: “For example, focusing on the Atlantic Ocean, there are regions with clearly improved DFS values from CrIS-TROPOMI, as opposed to CrIS.” - please explain how this sentence follows from Fig. C1.

- Line 364: “DFS values of 2 are achieved” - “values between 1.5 and 2” would be more correct.

- Line 365: “This suggests that CrIS-TROPOMI and CrIS are highly useful instruments for tropospheric ozone estimation.” - the notation “instrument” is incorrect if applied to CrIS-TROPOMI.

- Figure 6 caption: “the measurement or precision error” - do you mean “measurement noise error”?

- Line 382: “with the most reduction at the tropopause.” - at which pressure level is the tropopause?

- Line 383: “Given that the majority of the DFS are contained within the stratosphere for CrIS (Fig. 5), this is the expected result.” - why the reduction of the uncertainty is expected at the tropopause and not in the stratosphere, where the majority of the DFS are contained?

- Line 383: “except in the lower and upper stratosphere” - please specify pressure levels.

- Line 383: “... the variability of the total uncertainty is smaller than that of CrIS, suggesting that the inclusion of the TROPOMI radiances reduces the uncertainty of the CrIS retrievals.” - reducing the variability of the uncertainty does not necessary mean reducing the uncertainty itself.

- Line 424: “30° and 50°” - I guess you mean northern latitudes, please precise.

- Line 443: “For example, in Mongolia” I am not sure everybody can easily find Mongolia on your maps, please provide lat/lon.

- Sect. 5: when discussing stratospheric and tropospheric columns vertical ranges must be specified.

- Line 510: “ These results highlight the utility of the CrIS-TROPOMI retrievals in the stratosphere” - It is seen from Fig. B2a that the bias of CrIS-TROPOMI results is larger than for CrIS, i.e. the black solid line is shifted more upwards from the blue dashed line for CrIS-TROPOMI case. Is the intercept calculated correctly? If yes, is it representative for a bias? Also from Fig B2b the bias for CrIS-TROPOMI seems larger (green-blue color) as compared to that for CrIS (green-yellow colors). Please re-calculate the bias using the conventional definition as a mean difference and check your conclusions.

- Line 586: "...e.g., the Atlantic ocean." - please indicate the latitude region you are talking about.
- Figure 15: Please add a comparison for a priori.
- Figure 15: The range of x-axis is unnecessary wide. Please reduce to ±50%.
- Line 622: "as well as the other presented pressure levels across all seasons." - do you mean that RMS for CrIS-TROPOMI/CrIS-only is lower at all pressure levels in all seasons? Please reword this part of the sentence to make it more clear.
- Line 629: " ... as well as the stratosphere." - stratospheric results are not considered in Sect. 5.5. Thus, the statement about the stratosphere is not appropriate here.
- Table 9: Please provide statistics for the tropospheric column.
- Line 659: "In the stratosphere we find improved performance..." - the performance is improved only slightly, this must be said. An open question remains if the scatter plot really representative for the bias. This needs to be checked by calculating the mean difference over all data. If needed, conclusions have to be adjusted.
- Line 660: "Cross comparisons of CrIS-TROPOMI/CrIS/TROPOMI, with independent datasets from MLS, MUSES AIRS-OMI, JPL MOMO-Chem and CAMS, show in general CrIS-TROPOMI has the highest quality performance relative to the other instruments" - comparisons with JPL MOMO-Chem and CAMS cannot say anything about performance of one retrieval with respect to the other as it is unclear how to rate the agreement or disagreement with the model data.
- Line 664: "By contrast, despite being a TIR instrument CrIS shows close linear correlation with MLS, indicating the utility of CrIS by itself" - The vertical region should be mentioned, which this statement is applicable to.
- Line 686: "MUSES will immediately be able to take advantage of any improvements." - The statement is questionable as including additional spectral ranges often requires an adjustment of the retrieval parameters and subsequent tests and validation. The statement has to be removed.
- DFS for CrIS retrieval is significantly larger in Fig. 16 in comparison to Fig. 4 (4.12 vs. 3.62), although the scenario is expected to be the same. Please explain why it is the case?

Technical corrections

- Lines 126 and 128: Corrupted citations

- Line 279: remove "<", remove "4)"
- Lines 352 and 356: Corrupted citations
- Line 379: "is a maximum" - do you mean "is at maximum" or "is maximum"?
- Line 381: "Comparisons of the total uncertainty with the a priori uncertainty shows.." - the noun is in plural while the verb is in singular
- Line 402: "failures are by and large due to" $\longrightarrow$ "failures are by large due to"
- Line 403: "having too large a magnitude" $\longrightarrow$ "having too large magnitude"
- Line 443: "For example, in Mongolia" $\longrightarrow$ "For example, over Mongolia"
- Line 443: "retrievals had $\sim \times 2$ greater magnitude" - please replace "had" by "have" and write $\sim \times 2$ in words.
- Line 541: "Yet, considering Eq. 7, when using the TROPOMI AK, (which based on Fig. 4 is effectively 0 in the troposphere)." - incomplete sentence
- Line 543: "troposphere retrievals" $\longrightarrow$ "tropospheric retrievals"
- Line 543: "This was confirmed when we compared the TROPOMI-only tropospheric column results against CAMS, unmodified by the observational operator. Where the a priori and TROPOMI-only tropospheric column show almost identical comparisons." - Either it should be one sentence or the second sentence should be reworded to avoid a suboptimal beginning with "where".
- Line 582: "... spatial differences ..." - "spatial distribution of differences" would be more appropriate
- Line 594: "both CAMS have MOMO-Chem" - should it be "both CAMS and MOMO-Chem"?
- Line 622: "lower for CrIS-TROPOMI/CrIS-only than AIRS-OMI/AIRS/OMI" $\longrightarrow$ "lower for CrIS-TROPOMI/CrIS-only than for AIRS-OMI/AIRS/OMI"
- Line 637: Corrupted citation
- Line 647: Corrupted citation
- Line 685: Corrupted citations

References Fu, D., Kulawik, S. S., Miyazaki, K., Bowman, K. W., Worden, J. R., Eldering, A., Livesey, N. J., Teixeira, J., Irion, F. W., Herman, R. L., Osterman, G. B., Liu, X., Levelt, P. F., Thompson, A. M., and Luo, M.: Retrievals of tropospheric ozone

profiles from the synergism of AIRS and OMI: Methodology and validation, Atmospheric Measurement Techniques, 11, 5587-5605, https://doi.org/10.5194/amt-11-5587-2018, 2018.

Mettig, N., Weber, M., Rozanov, A., Burrows, J. P., Veefkind, P., Thompson, A. M., Stauffer, R. M., Leblanc, T., Ancellet, G., Newchurch, M. J., Kuang, S., Kivi, R., Tully, M. B., Van Malderen, R., Piters, A., Kois, B., Stübi, R., and Skrivankova, P.: Combined UV and IR ozone profile retrieval from TROPOMI and CrIS measurements, Atmos. Meas. Tech., 15, 2955-2978, https://doi.org/10.5194/amt-15-2955-2022, 2022.

Worden, H. M., Logan, J., Worden, J. R., Beer, R., Bowman, K., Clough, S. A., Eldering, A., Fisher, B., Gunson, M. R., Herman, R. L., Kulawik, S. S., Lampel, M. C., Luo, M., Megretskaia, I. A., Osterman, G. B., and Shephard, M. W.: Comparisons of Tropospheric Emission Spectrometer (TES) ozone profiles to ozonesodes: methods and initial results, J. Geophys. Res., 112, https://doi.org/10.1029/2006JD007258, 2007.