Comment on amt-2021-353
Joern Ungermann (Referee)


GENERAL COMMENTS

The paper describes a new level 2 data set for the CrIS instrument. The retrieval process is briefly described and an in depth comparison to in situ data gathered by the ATom campaign(s) is given. Due to the number of CrIS instruments in orbit and in planning, this is an important data product.

The paper identifies a strong bias of the provided data, which is larger than the assumed uncertainty. A H2O-VMR-based bias correction is suggested in the User Data documentation, but not discussed in the paper itself.

The paper should address the bias more explicitly and discuss causes and corrections. Ideally, the root causes for the bias should be identified and the data product improved.

I recommend publication after revising the paper to discuss these points in detail and answering the other comments below.

MAJOR COMMENTS

line 132

The paper identifies a bias of -100 pptv in the derived data, which is larger than the supplied uncertainty in the data (80 pptv) derived from the standard deviation computed from differences to in situ measurements.

This suggests that the bias is real and significant, particularly for non-polluted airmasses outside of plumes.
The employed spectral region is full of emission signatures of a wide range of trace gases. It seems as, e.g., CCl4 could still have an effect, but also other CFCs, or ClONO2 emit in this region. While the strong H2O emission line at 785 has been avoided, weaker lines are certainly present in the left window.

The User Guide for the data even provides a bias correction formula depending on water vapour.

I question the usefulness of the data set in the current state.
1) Why was the obvious and *astonishingly* stable bias not corrected in the data set?
2) Why was the retrieval not improved to the point, where no bias correction is necessary?
3) Why was the bias correction formula of the User guide not mentioned or applied for the comparison?

These points need to be addressed within the paper.

Figure 2
--------

High PAN VMRs occur often at higher tropospheric altitudes (particularly due to the longer lifetime at colder temperatures) close to the tropopause. The used aircraft data rarely go above 12km.
Biomass burning plumes reach higher than 12km, particularly in the tropics.
The given altitude range of 800hPa to 300hPa is key here, as 300hPa corresponds roughly to 10km.
How does this limited altitude range affect the accuracy of estimating total PAN in the UTLS?
Why is the instrument not sensitive (at all? enough?) to high PAN VMRs closer to the tropopause? Is this related to the low temperature at this altitudes?

SPECIFIC COMMENTS
==================

line 135
--------

Particularly in the face of the discovered systematic error, a discussion on the sensitivity of the retrieved PAN VMRs on the previously derived quantities (i.e. the 'b' vector) might be interesting. It is mentioned that the retrieval processor under-estimates the "observation error", without detailing what exactly this entails. Often this only contains - for practical reasons - an estimate of the noise induced error, not the systematic errors.

How does the identified systematic bias relate to the error diagnostics for systematic (b-related) errors?

line 165
--------
Please show a (representative set of) averaging kernels to show the region of sensitivity.

Figure 4
--------

The residual shows structure beyond the noise level (blue lines). The caption indicates that this spectrum was computed with a zero PAN profile. Please show both a spectrum with the derived PAN profile and with a zero profile to show the improvement and PAN signal as well as quality of fit of the used spectra (similar to Glatthor et al., 2007)

Figure 5
--------

The paper identifies a low bias of 100 pptv causing many VMRs to be negative as shown in Fig. 7. Figure 5 shows only positive VMRs. Please explain the discrepancy.

MINOR REMARKS
===============

line 108
--------

A big X with a hat was not in (1). Maybe big-hat-x -> hat-x and hat-x-a -> hat_x ?

line 113
--------

\Delta f should be 'bold'.

line 128
--------

an approximate solution?

line 134
--------

CCI_4 (small l)

line 139
--------

It is not clear from the context what the "forward stream" is. The given reference distinguishes a "reanalysis stream" without being clear on the difference. I suppose it has something to do with using (forward) extrapolation of calibration data in contrast to interpolation using (later) data. This is probably a very common term in certain scientific communities.

Maybe explain it in a brief sentence.