Reply to review #2 of our manuscript number amt-2021-158 'Glyoxal tropospheric column retrievals from TROPOMI, multi-satellite intercomparison and ground-based validation.'

We gratefully thank the reviewer for the careful reading of our manuscript and for the very constructive comments. Below, the reviewer's text is given in black while our replies and description on how the comments have been addressed in the manuscript are given in blue.

The manuscript "Glyoxal tropospheric column retrievals from TROPOMI, multi-satellite intercomparisons and ground-based validation" by Lerot et al., presents global glyoxal observations made by the TROPOspheric Monitoring Instrument (TROPOMI). The paper provides the description of the retrieval algorithm, an inter-comparison of glyoxal observations from TROPOMI and other low earth orbit (LEO) satellite retrievals, and validation leveraging a few available MAX-DOAS glyoxal observations. This new retrieval would enable new atmospheric chemistry studies given improved spatial resolution and retrieval noise levels in comparison with retrievals from prior LEO satellites. The paper well written and constitutes a good reference for future studies using TROPOMI glyoxal observations. Its publication its therefore more than justified.

There a few aspects of the retrieval description, the uncertainty calculation and the comparisons with other satellite could benefit from further descriptions and clarification. It would be great if the authors could address the following comments during the discussion before final publication of the manuscript in AMT.

My main concern regarding the different retrieval steps is the assumption of a constant 1×10^{14} molecules/cm^2 vertical column over the Pacific Ocean as reference for the background correction. This value is based on observations from one group (Sinreich et al., 2010) using an observation methodology similar to the satellite retrieval (DOAS fit) that could be affected by similar biases. At the same time, this results differ from other ocean glyoxal observations (for example Mahajan et al., 2014) reporting smaller columns over the oceans. It would be interesting to provide further discussion about the effect of the background correction in the final reported columns. How much would differ the final columns have the author's decided to use reference columns from chemical transport models or other sources?

We agree with this comment and the uncertainty related to the reference value within the Pacific sector. Below is first our detailed reply and then we describe how the manuscript has been modified to clarify this aspect.

The large uncertainty related to the reference value is included in the total error budget (section 3.4.3). We took an error associated to this reference value $\sigma_{N_{v,0,ref}}$ of 5×10^{13} molec/cm^2; which is propagated through the retrieval. Similarly, the errors on the AMFs and slant columns from the reference measurements are also taken into account to compute the total background correction error. As illustrated in the Figure C1 below for one S5p orbit passing over Africa, the background correction error (red plain curve) is significant and contributes as much as the DOAS (blue curve) and AMF (green curve) error components to the total glyoxal VCD systematic error (black curve). The figure also shows that the choice of the reference value $N_{v,0,ref}$ contributes to about half of the total background correction error.

The choice of the value for the reference glyoxal vertical column indeed directly impacts the overall level of the product, with some small modulation directly related to the ratio of

the AMFs over Pacific and in other regions (M_0/M in eq. (1)). This is illustrated in Figure C2, which shows the impact of using a lower Pacific reference value on the final glyoxal VCDs.

It has to be noted that all the intermediate variables for computing the glyoxal tropospheric vertical column are given in the product (SCD, AMF, corrected-SCD). An advanced user is then in position to recompute glyoxal VCDs using a different reference background value if needed.
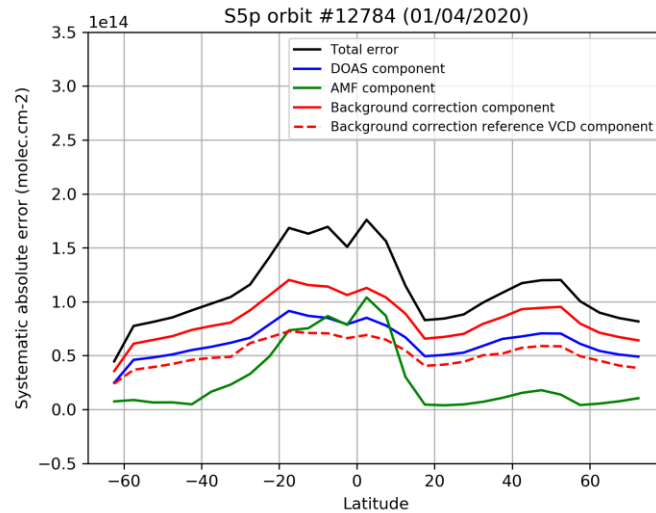

Figure C1 : Illustration of the zonal mean total systematic CHOCHO VCD error with its different components for one S5p orbit passing over Africa.
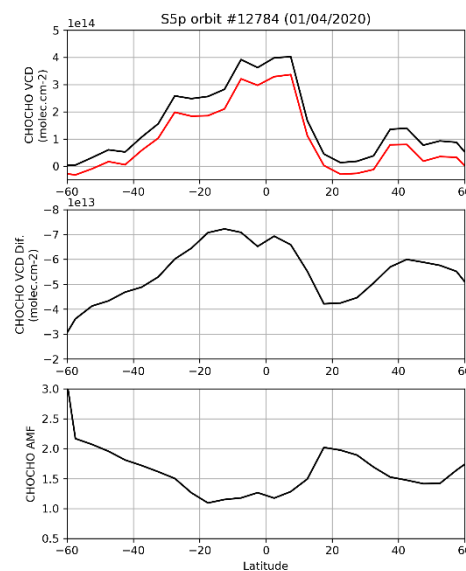

Figure C2 : Impact on the mean CHOCHO VCD for one S5p orbit of using a reference Pacific column of 0.5E14 instead of 1E14 molec/cm². The upper panel compares the corresponding lower columns (in red) with the original columns (in black). The absolute differences plotted in the middle panel are anti-correlated with the AMFs shown in the bottom panel.

Corrections in the manuscript:
1. After the first paragraph of section 3.3 the following sentences have been added:
   "There is nevertheless an uncertainty related to this reference value, which

impacts the overall level of the product. This error component is further discussed in section 3.4.3 and is taken into account to estimate the total glyoxal VCD error. As all intermediate variables (SCD, corrected-SCD, AMF) are provided in the product, a user could recompute glyoxal VCDs using a different reference Pacific value."

2. At the end of section 3.4.3, we have added: "*Using a different reference value would directly impact the overall level of glyoxal VCDs worldwide, with some small modulations related to the ratio of the AMFs over Pacific and in other regions following Eq. (2).*".

3. In section 3.4.4, we have added a simplified version of Fig. C1 and the corresponding description in the text: "*Figure 7 shows the zonally averaged total systematic error along with its different components for one S5p orbit passing over Africa. In general, the three components contribute similarly to the total error for emission conditions. On contrary, the AMF error becomes smaller in background conditions while the two other terms dominate.*"

Also, to understand the effect of each retrieval step in the final VCDs around the globe it would be beneficial to add a figure showing global values of dSCDs, VCDs, and background corrected VCDs so it is easier to interpret the amount of information present in the final VCDs brought in by each retrieval step.

This is a good suggestion and we have added a figure (Figure 1) in the general overview of the algorithm presenting the main output for one day of TROPOMI data of the different algorithmic steps, which are further described afterwards. The colorbar is slightly different than for other figures to better show possible structures in other ranges of values (see minor comment below).

Other comments and doubts:

The description about the calculation of pseudo-absorbers to account for scene heterogeneity leaves some questions un-answered: (1) what are the criteria defining the two additional cross-sections for scene heterogeneity? (2) What is the effect of using one vs. two extra pseudo cross sections? (3) how is defined the remote region over which the heterogeneity cross sections are calculated?

The heterogeneity factor is computed for every nominal ground pixel using the radiances available in the L1 data at higher spatial resolution for a limited number of wavelengths. This factor represents the scatter of the small pixel radiances within the nominal ground pixel. It generally ranges from -1 to +1 with values close to 0 indicating an homogeneous scene while higher absolute values indicate a larger level of brightness heterogeneity within the ground pixel. The heterogeneity cross-sections are constructed by comparing systematic residuals of scenes with a heterogeneity factor larger than +/- 0.08 with those from homogeneous scenes. Indeed the two cross-sections are significantly correlated. However, some remaining differences between them (likely related to the way the ISRF is perturbed depending on the radiance distribution within the ground pixel) lead to a further improvement of the fit quality when including the two.
The cross-section construction can be done using one single orbit over the Pacific. Within this process, we assume that the glyoxal fields vary smoothly along the orbit track.

In the manuscript, we have slight rephrased the section 3.1.1 to add those different pieces of information. It reads now as:

"*Those cross-sections are generated with a statistical analysis of the fit residuals for many observations in the Pacific Ocean as a function of the level of scene heterogeneity. The latter can be computed using radiance measurements at higher spatial resolution available in the TROPOMI level-1 data at a limited number of wavelengths. It ranges between -1 and +1 and is close to/deviates from 0 for homogenous/heterogeneous scenes, the sign indicating the part of the ground pixel that dominates the scene brightness. Following this approach, two additional cross-sections corresponding to the systematic residuals of scenes with an heterogeneity factor larger/smaller than +/- 0.08 have been added to the DOAS baseline and both the fit residuals and the identified glyoxal biases have been reduced as illustrated in the right panels (b) and (d) of Figure 2 . This effect is particularly visible along coasts and mountains but also over lands where some pseudo-noise caused by persistent broken clouds is also largely reduced. Although significantly correlated, including the two heterogeneity cross-sections leads to a further improvement of the fit quality, likely due to a slit function perturbation that depends on the radiance distribution within the nominal ground pixel.*"

How many Taylor expansion terms are considered in the derivation of the empirical correction associated with NO_2 slant columns?

We used a first order Taylor expansion of the NO2 optical depth around the wavelength and the vertical optical depth. Therefore, two additional cross-sections are used to derive the empirical correction.
The information has been added to section 3.1.2

Are the MAGRITTE a priori glyoxal vertical profiles computed daily at the satellite over pass time or are they compiled as a monthly climatology as done in most heritage glyoxal satellite retrievals?

Indeed, the a priori profiles are still based on climatological data. However, the model has been recently updated for the chemical and deposition mechanisms. In addition, the a priori profiles are provided by the model at the respective satellite overpass time. The information has been added in the text.

How is the interpolation of the background correction matrix done outside the 40°S to 40°N area?

To avoid meaningless correction values out of the 40°S-40°N area, no extrapolation is performed in the second step of the background correction but instead we simply use the nearest neighbour correction values. The final offset correction can be safely applied to all regions, even if only prescribed by the 40°S-40°N region. The information has been added in the text.

The classification of all AMF uncertainties as systematic is confusing. First, it is important to acknowledge how complicated it can be discriminate systematic and random uncertainties in the AMF calculation and the different sources of uncertainty. The authors should be thank for the efforts they have put in trying to quantify such

uncertainties. Said that, given the uncertainties inherent to chemical transport models and surface reflectance climatology, and the representation errors associated with different spatial and temporal resolutions some of the AMF errors have to be necessarily random. Given the mean biases between MAX-DOAS observations and TROPOMI retrievals reported in the manuscript (always < 0.6×10^{14} molecules/cm^2) should not the systematic uncertainties reflect this in panel c) of figure 5 with negative values?

We agree that this classification can significantly depend on the application and on the spatial and time resolution of interest.
Here, we considered as systematic the errors that would remain the same for a measurement at the exact same location and time, as well as same atmospheric conditions. As the reviewer says, when looking at extended temporal or spatial scales, part of those systematic errors may appear like noise and our approach to estimate the errors is therefore conservative. This has been discussed by Vigouroux et al. (2020) (https://doi.org/10.5194/amt-13-3751-2020), who attributed part of the scatter in HCHO satellite-MAXDOAS differences to a random component of the AMF errors.

The total systematic error is obtained by assuming that the different error components are uncorrelated and is derived as the root square of the quadratic sum of the different error components (eq (2)). Therefore, this estimate gives an indication of the possible range of the total systematic error without any indication of its sign. From this perspective, validation clearly provides crucial complementary information even if the identified absolute biases are consistent with the estimated errors.

We have added in the manuscript at the end of the 1st paragraph of section 3.4 the following statements:
"*It has however to be noted that the latter assumption may lead to conservative systematic error estimates and to an underestimation of the product scatter, depending on the time and spatial resolution of interest. In particular, uncertainties associated to the input parameters needed for the AMF calculation are directly related to the resolution of the used databases and may appear as random at coarser resolution. This has been discussed by Vigouroux et al. (2020) who attributed part of the scatter in formaldehyde vertical column TROPOMI/MAX-DOAS differences to a random component of the AMF errors.*"

During the discussion of uncertainties associated to a priori glyoxal profiles, an effective height uncertainty of 50 hPa is assumed. How is this value obtained?

The uncertainty of the effective profile height is determined by statistical analysis of the profile heights from one year of model data. 50hPa corresponds roughly to the standard deviation of the profile heights over polluted regions. The information has been added.

While the color scale used in figures 7, 8, 9, and 10 produce clean plots they fail to convey complete quantitative information. First, despite glyoxal VCDs ranging between 0 and 1×10^{14} molecules/cm^2 in most parts of the world the color scheme does not allow appreciating any structure for that given range. Second, what color is assigned for values below 0 and above 6×10^{14} molecules/cm^2?

Values smaller than 0 or larger than 6×10^{14} molecules/cm^2 are colored with the colorbar extremes (light blue or dark red). Indeed, the colorbar of Figs 7-10 has been chosen to highlight at best the more important emission regimes. Nevertheless, we think that the

main persistent weak CHOCHO signals are also visible. In the new Figure 1 that we added to illustrate the different algorithmic steps, the colorbar is slightly different and for example better shows that the amount of negative columns over oceans is limited.

Minor typos and language comments:

Line 56: I think "precursors is" should be "precursors are"

This has been corrected.

Line 207: "end hence" should most likely by "and hence"

This has been corrected.

Line 316: "Anthropogenic NMVOCs emissions of are" should be "emissions are" without the "of"

This has been corrected.

Line 438: "see above section 6.5.1" is meaning "see section 3.3"?

We intended to refer to section 3.4.1. This has been corrected.