Reply to Anonymous Referee #2

General comments

Although, the main finding are very well described, my main concern with the paper is the missing discussion on the reasons of main difference between TROPOMI and FTIR formaldehyde BIAS for some stations (large offsets) and also difference in seasonal cycle (e.g. Paramaribo, Paris, UNAM…) (See Figure 5.).

The topic of this work fits well within the scope of AMT. Although the paper is well structured, the text needs to be carefully revised in order to be more precise in some sections. I recommend acceptance to AMT after addressing the comments above and few minor comments below.

We thank the referee for their work and useful comments.

We first answer on the main remarks above and reply then following the minor comments below.

We have added some possible reasons for the observed TROPOMI bias in the revised version of the manuscript. To avoid repetition, we refer to our reply to referee#1 (last page) who had the same concern as referee#2 on missing discussion on the observed biases.
It should be noted that even if the offsets are large, they are within the accuracy requirements of the satellite (which were based on previous validation studies of HCHO satellite measurements), meaning that such large biases were expected.

The TROPOMI and FTIR seasonal cycles are usually in agreement. However, as pointed out by the referee, this is not the case for Paramaribo. But, as can be seen in Fig.5, the sampling (number of coincidences) is bad there with often only one coincidence per month. Then, if TROPOMI has a remaining outlier, it has a strong influence on the plotted seasonal cycle (e.g. June 2019 shows a negative TROPOMI value). With an improved QA value as expected for the next TROPOMI versions, the comparisons should also improve. For Paris, UNAM (Mexico City) and usually all polluted sites, the TROPOMI and FTIR seasonal cycles show similar features, but the amplitude is smaller with TROPOMI due to its proportional bias that leads to more under-estimation for high HCHO levels (so more under-estimation during the maximum of the FTIR seasonal cycle).

Page 2, line 5, confusing sentence, "accuracy is below the upper limit of the pre-launch requirements of 80%, and below the lower limit of 40% for 20 of the 25 stations", it does not make sense to write that HCHO TROPOMI retrievals are below lower and upper limits. Please clarify it.

We have clarified the text:
"The pre-launch requirements of the TROPOMI HCHO accuracy are 40-80%. We observe that these requirements are well reached, with the BIAS found below 80% at all the sites, and below 40% at 20 of the 25 sites."

Page 3, line 1, is there any study of validation of satellite HCHO observation with ship-based measurements?

Indeed. We have added two references as example of such studies (Peters et al., 2012; Tan et al., 2018).

Page 3, line 8, please define what is "TROPOMI Cal/Val"

Done.

Page 3, line 15, would you please mention what are the differences among versions from v.1.1.5 to v.1.1.7?

In the AMTD paper, we referred to the ReadMe file (https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Formaldehyde-Readme.pdf) for details on the differences in the versions because they have minor impacts on the HCHO TROPOMI time-series. However, as both referees ask that all is included in our paper, we have included a Table (Table 1 in the updated version) repeating the information about the different versions (dates and changes). (see also reply to referee#1)

Page 4, line 12, why to use OMI albedo climatology?

The OMI albedo climatology is the best product existing at 340 nm. The spatial resolution is indeed too coarse for TROPOMI. We are waiting for a climatology based directly on TROPOMI, but it is not yet available.

Page 4, line 13, "(Kleippol et al., 2008)".

Done (changed to Kleipool).

Page 4, line 20, please define all the quantities of the equation (e.g., M and M0)

All quantities have been defined in the text above the equation, except M0, which is an average of the air mass factors (M) of the slant columns selected in the reference sector, the Pacific Ocean (N(s,0)). We have added its definition in the new manuscript.

Page 6, line 6, what is the main difference between PROFITT9 and SFIT4.0.9.4?

Both codes are very similar. They are both line-by-line models for infrared solar transmittance spectra, including a radiative transfer model (FSCATM and KOPRA for SFIT4 and PROFITT9, respectively), and based on the optimal estimation method (Rodgers, 2000). They both allow for the Tikhonov regularization as well. Differences are minor, mainly lying in the different options that are available (but not used in the present work) in PROFFIT9: e.g. possibility to retrieve the temperature profiles,… The only relevant difference for the present work is the calculation of channeling error that is not yet included in SFIT4.
The use of different codes within the InfraRed Working Group (IRWG) of NDACC is historical. To certify a good homogenization in the delivered FTIR products in the NDACC database,

harmonization in the retrieved parameters (spectral micro-windows, a priori profiles, spectroscopic database,…) is required for all NDACC target species, and has also been done for the HCHO products presented here (Vigouroux et al., AMT, 2018). A comparison exercise of the two codes has been performed for four species (Hase, JQSRT, 2004) and an agreement within 1% has been found in the retrieved columns. Further details on both codes are available in this latter reference, which has been added in the next version of our manuscript.

Page 7, line 7, please be consistent between names used in the text "Maïdo" and used in the figure 1.

The figures are automatically generated using the name provided by the PIs in the geoms file. Indeed, it might confuse the reader to see two names, maybe not so much for Maïdo / LA.REUNION.MAIDO; but for Mexico City (used in Tables and text) and UNAM (used in automated figures). To help the reader, we have explicitly added the two possible names in Table 1 (Table 2 in the new version). We prefer to keep using both names because a geoms data user would find the "UNAM" name for the station, while a "Mexico City" name makes the information clearer for a simple reader that the station is in the city center of Mexico. Note that this situation is also there for other stations, but with clear signification (Izaña / IZANA; Mauna Loa / MAUNA.LOA.HI;…).
Because, the correspondence is less clear for Mexico City / UNAM, we have also repeated the two names in Table 2 (Table 3 in the new version).

Page 8, line 3,what are the reasons for the lowest smoothing systematic uncertainties in the 5 added sites.

The provided 3.4% number for median smoothing systematic uncertainty is the one given in Vigouroux et al., AMT, 2018. The 13% and 14% for the 5 added sites, are for the total systematic uncertainty (dominated by the spectroscopy), and not for the smoothing part only. To avoid the confusion, we have changed the sentence to:
"For the five added sites, the median total systematic uncertainty is 13% (Jungfraujoch, Tsukuba, Palau), or 14% (Rikubetsu, Xianghe), commensurate with the other sites."

Page 8, line 25, please remove "so"

Done.

Page 11, line 29, would be nice if you include one or two sentences describing the main differences between OFFL, RPRO and NRTI products. Are they different at all.

OFFL, RPRO and NRTI share the same algorithm (for the versions used in the paper). Changes of version numbers refer to changes in other components of the operational processor. However, slight differences come from auxiliary data. A priori profiles used for NRTI are from TM5 forecast model, while they are from TM5 analysis for OFFL/RPRO (this makes almost no difference since HCHO is not assimilated).
For the reprocessing (RPRO), data have been processed using 7-days parallelization (in order to speed up the reprocessing). It means that the slant columns used for the background correction are always at least 7 days older, while for OFFL and NRT, the gap is only 1 day. It results in stripes slightly more pronounced in the RPRO product than in the other versions.

We do not give these details on the OFFL / RPRO / NRTI data because they have negligible impact on the satellite data and validation results. But we have added a Table (Table 1 in the new version) with the date of the different versions, and all details can be found in the Readme file given as a reference in the manuscript.

Page 16, line 31, would you please clarify how the collocation plays a role in Maïdo? Fire emissions are included in the calculation of the a-priori profiles? Could fire emissions enhanced the HCHO amounts? What is the effect of changing the collocation radius in this station?

Maïdo is usually a clean site. During the biomass burning period, some plumes (mainly coming from Madagascar for the short lifetime species HCHO) can cross over Reunion Island. Since the overestimation of TROPOMI at Maïdo is larger during the biomass burning months, we suggest that this could be due to plumes that would be present in the 20km circle around the station covered by TROPOMI but not in the line of sight of the FTIR measurements for the collocated days. Unfortunately the collocation effect could not be confirmed at Maïdo: the 10km radius criterion lead to very few coincidences at Maïdo (see Fig. 1 of our Reply to referee#1), and none are during the biomass burning season.
This was only a suggestion from our side (we wrote “the collocation of the plumes *might* play a role there”), and could be investigated when more data are available in a future work.