Comment on amt-2021-39  
Anonymous Referee #2


This work integrates many up-to-date models into the TROPOMI NO2 retrieval system, which leads to a great deal of improvements in retrieval accuracy. The paper is well organized and written, and fits well to the scope of AMT. I recommend its publication after the following revisions.

- It is a long paper (43 pages) and involves so many models, making it pretty difficult to read thru the whole paper. So it will be very helpful if the authors could provide an error budget table which will highlight the biggest contributors to the overall accuracy improvement of the retrieval system compared to the reference models. That way readers can easily see the relative importance of one model/input parameter to others, even though Fig. 19 and section 7 (conclusions) provide some detailed numbers about uncertainty and improvement.
- Sensitivity analysis and case studies have been made for almost all important input parameters. However, readers may wonder when doing sensitivity analysis or case study of one parameter, what the values of other input parameters are. The authors can improve on this by listing the default values of all input parameters. That is, it does not change until it becomes the parameter of analysis.
- There are a few places which have inconsistency issues. For example, L157-163, p6 x-track striping issue is introduced and de-striping approach is discussed. However, it is not clear if de-striping correction has been applied to the slant column retrieval (not mentioned thereafter). There is discussion on how to remove VZA dependence of the stratospheric NO2 columns in DSTREAM as shown in Fig. 6 by dividing TROPOMI orbit swath into 3 segments (western, central and eastern), which, however, is not the approach discussed in page 6, L157-163 for de-striping.

Also, AMF is calculated using TROPOMI GE_LER data for surface albedo and OCRA/ROCINN_CAL model for cloud parameter as described in Table 1. But in SCD calculation (Eq.7), cloud pressure is from ROCINN_CRB model (see L302).

The box-AMFs in Eq.(3) tabled values are calculated using VLIDORT (what is the version number?), while radiance quantities for cloud radiance fraction computation (see Eq.5) and SCD calculation (see Eq.7) are simulated using LIDORT (again, which
version?). As we know, different versions of LIDORT and VLIDORT may produce inconsistent simulation results.

- **Some other minor issues**
  - Another cloud parameters retrieval algorithm – FRESCO is mentioned in L226-229, p9-10. But it seems that FRESCO is not used in this study, so why it is discussed here?
  - It is suggested to change “The IFS(CB05BASCOE) model” to “The IFS (Integrated Forecast System) model” in L262, p11, and remove “Integrated Forecast System” in the following line (L263). Also, in L267, change “using IFS(CB05BASCOE) data” to “using CB05BASCOE data”.
  - Fig.11, right panel, why does cloud fraction have negative values down to -0.2?
  - L371, how can ROCINN retrieve effective cloud pressure and cloud albedo values at cloud fraction of 0?
  - L395, it implies that the multiple scattering between the cloud bottom and the ground is not considered in the CRB cloud model. That simply is not true.
  - Fig.16, right panel, there is not much one can see. Suggested to change the color bar scale (reduce the up limit) to enhance the red color.
  - In section 6, page 25, why Uccle was selected as an example showing in Figs. 20-22. Readers may wonder why this site (not other site) was selected.