This paper reports on the capabilities of a thermal infrared FTS remote sensor to measure O3 and CO for air quality purposes, using the geostationary orbit. The study uses the instrumental specifications of a planned instrument (MTG-IRS, here called GEO-TIR2) to compare it to a concept of mission (GEO-TIR) and shows that when both the signal to noise (SNR) and the spectral resolution are improved, one can gained sensitivity towards the lower layers of the atmosphere. Some model simulations are also provided to demonstrate the usefulness of geo-based observation, at least when thermal contrast is significant.

I find the paper clear and well written. But I also find the scientific content weak and some parts of the paper are very misleading. I would not recommend the paper to be published as it is.

Major comments:

1/ About the GEO-TIR versus GEO-TIR2 comparison

The goal of the paper is to push a concept of an improved TIR mission, which (and I fully agree on that) should be useful for AQ purposes. This is done by comparing a potential instrument (GEO-TIR) to a planned mission, here called GEO-TIR2, and that is in fact MTG-IRS. But the study is limited to the comparison of the instrumental specifications only, whereas other criteria such as temporal and horizontal samplings are equally important. The pixel size is also crucial, and as the complexity/cost of an instrument depends on all this factors together, a smart compromise between all these criteria has to be found to define a useful mission. The study provides obvious conclusions, such as p3503: "*For both CO and O3, Fig. 5 shows that the GEO-TIR2 AVKs are broader than GEO-TIR ones, especially in the LmT, which confirms that GEO-TIR has a better vertical resolution".* People involved in the field all know that a better SNR and better spectral resolution helps to lower the errors and increase the DOFS. The paper spends a lot of time to re-demonstrate that, without properly compare the 2 instrument concepts (ie: a better spectral resolution has a cost in terms of acquisition time for each measurement, so e.g. how does it impact on the horizontal resolution, also a very important variable for AQ). IASI and AIRS are examples of such compromises, what you loose in terms of vertical information (eg as compare to TES), you gain in terms of horizontal resolution.
Another example of misleading sentence: Page 3502 at the end: *We simulate the observations from an instrument configuration close to MTG-IRS, which is dedicated to NWP (temperature and humidity), referred as GEOTIR2, to determine the added value of GEO-TIR.* Here again only the SNR and spectral resolution are used in the comparison (is it why there is "close to"?).

2/ About the use of DOFs

Much of the discussion is based on DOFs values. These figures in absolute values do not mean much, unless they are obtained using similar a priori and retrieval assumptions (which is the case here for "GEO-TIR" and "GEO-TIR2"). There is too much emphasis put on the values, that could be changed easily by eg allowing for more variability in the a priori.
P3501:
Considering current IR instruments, technical feasibility and cost (Astrium-EADS, personal communication) a DOF of ~1.5 for O3 and of ~2 for CO seems to be a good compromise to have vertical information in the troposphere.

It would have been more useful to make the discussion on the errors in the PBL (which is also done, but more importance is given to DOFs)

3/ About the focus of the interest on Lmt :
All the intro focuses on the fact that data are needed for AQ purposes, but several plots are presenting results for 0-15 km. Then the discussion should be limited to 0-3 km, or at least to the low troposphere. I doubt that what's happening above 8-10 km is influencing the PBL.
Eg the title says "to monitor the lowermost troposphere", but:
Figure 1 (2) is for 0-15 km;
Figure 3 (4) is for 0-20 km, as well as Fig 5
Figure 7 and 9 are useful
Figure 8 (10) a, b, c, d are for 0-15 km
Figure 8 (10) e,f,g,h are useful
Figure 11 is useful

Last but not least: to answer the question if these high resolution measurements from a geo platform would provide added value as compared to the current ground-based and polar orbit satellite data, one need to perform specific OSSEs. In the conclusion section the authors announce that these studies will be available. I think these results are really needed here, otherwise the whole context of AQ description and discussion with the improved GEO-TIR instrument is meaningless.

Other comments:
General:
- Emissivity is assumed to be 1 in the simulations? It is worth noting than if not one it changes the conclusions on the value of thermal contrast more favourable to see lower in the atmosphere.
- MTG-IRS should be called MTG-IRS, not GEO-TIR2, everywhere. A reader not aware of the Eumetsat mission program will be very confused.
- Section 3.2, Geostationary observation system: the model simulation is performed at the model grid resolution. But an instrument with a pixel of eg 4 km (as IRS-MTG) will provide more horizontal structures than an instrument with a larger pixel. Is GEO-TIR also 4 km? The model simulation is done with a larger grid, so it can not show the horizontal detailed structure.

Page by page:
- Abstract: I find it a bit too long + see my comments before on LmT versus 0-15 km
- In the introduction, p3492, it is reported that O3, NOx and PM are of great concern. But the paper focuses on CO and O3, it should be better justified why.
- Intro, page 3494, lines 5-12: True for the advantages of a geo platform but it should be mentioned that the observation is limited to 1/3 of the Earth
- Intro, Page 3494 the second para is very misleading: GeoTrope- GeoFIS, GeoCape, and MAGEAQ are mission concepts, whereas the MTG satellite is in phase B. This mission is not well described here (it is better described after): the launch dates should be checked on the Eumetsat web site (not 2017 for MTG-S), Sentinel 4 is the UVN instrument, not a full platform,

and the MTG-IRS instrument is not even mentioned here (whereas the info is provided on page 3503).

- Page 3497: retrieval is performed from 0 to 39 km: what is assumed above 39 km for O3?
- Page 3500: *Note that the DOFs depend on the instrument configuration but also on the retrieval method* : it also depends on the a priori (covariance matrix) used in the retrievals
- Page 3500, lines 26-30: the information content at eg. 2 or 3 km would be more useful that DOFs 0-15 km
- Page 3503: model grid used not provided
- Page 3505: I don't understand why the model fields are not smoothed by the avgk.
- Page 3506, line 3-5: this depends on the (vertical) correlation length for the cov matrix
- Page 3508 : I like Figure 11 and this plot is more useful than the previous plots to evaluate the impact on AQ
- P3510 Suggestion to remove or modify: *We confirm that the shape of the averaging kernels of TIR instruments is highly dependent on the thermal contrast,* as this was shown on real measurements in several publications, it is expected that simulations do not show the contrary.
- Page 3510: *O3 and CO distributions over Europe as measured by GEO-TIR and the future GEOTIR2 are simulated*: as said earlier I don't think the simulation takes into account all the specifications of the MTG-IRS mission.
- Similarly, on page 3511: *We have shown that such a configuration (GEO-TIR) is capable of bringing added value in the LmT compared to a configuration optimized for numerical weather prediction (GEO-TIR2).* Same comment, the demonstration should be done using proper pixel size, spatial resolution, temporal resolution. Eg for MTG the frequency of obs is 1/2h, here it is assumed to be one hour. By accumulating the data on one hour you will improve the SNR…
- Fig 1: in the text the authors use SSI and here spectral resolution
- Fig 1: I don't understand why the spectral res value for GEO-TIR2 is not 0.625 as in Table 1.
- Fig 3: The text (page 3501 l23) says thermal contrast=2K (not zero as here)
- Fig 7 and 9: avgk should be used to compare the obs. with the model.

Typos:

- Page 3504 : … theseS radiances (remove S)
- Page 3500 *(-2K Fig 2a and b)* >> 2K Fig 1b and 2b

So, in summary my recommendation would be to add realistic model simulations that take into account all variables, ie SNR, spectral resolution, pixel size, horizontal sampling, and frequency of the observation to demonstrate:

- **1/ the benefit of GEO-TIR on MTG-IRS;**
- **2/ the usefulness of such measurement to improve AQ forecast**

And that would be a nice and useful paper.