Review of the ACPD manuscript: *'Characterizing Sampling and Quality Screening Biases in Infrared and Microwave Limb Sounding' by Millan et al.*

The paper by Millan et al. describes effects of sampling and quality screening for two limb sounding instruments (MLS and MIPAS). Based on CCM data the analysis shows significant biases in the zonally averaging of various trace gases and temperature distribution, especially for a MIPAS-like instruments with its cloud-induced quality screening. Biases in the zonal means as well as for trend analyses in the upper troposphere can be quite large, which is an important fact and needs consideration in any trend analysis.

General comments:

There are two major concerns regarding the manuscript from Millan et al.:

a) The study highlights a very specific caveat of sampling and screening issues of space borne instruments. The whole approach is very technical, investigating a very obvious problem of the IR limb measurements due to cloud screening, which can have a severe effect on the results especially if you like to analyse trends in the very cloudy upper tropical troposphere. Although this effect is of interest for ACP community, these kind of analysis is not really suited for ACP. The whole manuscript looks more like a technical study or a technical note (for example as prepared for NASA or ESA studies for future satellite sensors) and is not sufficient for a full publication in a journal like ACP. A more in-depth study of the MW/IR limb caveats on trend analyses cloud be publishable in a more technical and instrument-orientated journal like Atmospheric Measurement Techniques.

b) To my mind, this paper does not contribute a substantial add-on to a former study of the authors. The paper by Millan et al. (2016) in ACP describes most of the general effects of sampling pattern, which can be quite similar effects to the quality screening, in the context of a much more detailed study on atmospheric trend analyses. My strong impression is that the actual manuscript misses new concepts and methods and just repeats parts of the Millan et al. (2016) analyses with another combination of satellite instruments.

Major comments:

1) P4, line 29: Although the screening is a fundamental issue of the study, there are no details presented for MLS (but some for MIPAS). Please, at least summarize the main topics of Livesey et al. (2006). In addition, more details on the cloud screening of MIPAS would be helpful as well (see next comment).
2) P5, line 1: Please note somewhere here that the MIPAS cloud-screening applied in the IMK product is a very conservative approach with respect to cloud effects in the measured spectra. This might have effects on the presented analysis. Especially, in the tropical upper troposphere high water vapour abundance will result in false-positive cloud detections (e.g. Spang et al. 2012) in the IMK-product and consequently an artificial overestimation in the biases, where other MIPAS level 2 processors might do a better job. This objective should be addressed in the manuscript.
3) The disadvantages of the cloud screening in the tropical UT for the IR measurements are highlighted frequently in the manuscript and are presented by different types of analyses, but with only very limited further information content. In contrast, the caveat of MW sounders to retrieve trace gases below ~300 hPa at any latitude is only briefly mentioned.
   Studies for future ESA missions, like the PREMIER report for mission selection (http://projects.knmi.nl/capacity/PREMIER/SP1324-3_PREMIERr.pdf), show nicely the excellent synergy by a combination of IR and MW (s. Fig. 4.2 of the report), where IR-limb still allows measurements in the mid-troposphere, where MW-limb fails to measure.
4) To my mind the trend analysis section needs to be improved. Although MIPAS tackles to deliver good trend analyses in tropical UT, this is a shortened result. Have you tested if the biases are better, if the trend analyses are restricted to longitude regions where convection is less pronounced (e.g. mid-Pacific, Atlantic regions, excluding the warm pool region, central Africa and South America, where deep convection and cloud-coverage is most pronounced). Local trends in the UT/MT in tropical and subtropical areas would be valuable in principle.
5) I am also wondering why the authors argue that most of the biases are not caused by a lack of observations in convective (cloudy) areas in the tropics but are caused by a reduction in the overall sampling number. I think this is linked with each other, especially for water vapour with its extreme

vertical and horizontal gradients. This objective would need further investigations for a final publication and is not sufficiently discussed by the only two sentences in section 3 (p5, line 25).

6) Further, the quantification of the number of years needed to compensate for the biases introduced by the intensive cloud-clearing for IR limb sounder, which is only briefly mentioned in the Summary and Conclusion section, would be a very interesting topic for more detailed analyses and would improve the quality of the manuscript.

Minor comments:

1) P2, line 23: Please, add a reference.
2) P3, line 3: Have you introduced CMAM30-SD already?
3) P4, line 7: Please, explain in more detail, why the horizontal gradients are that important and why a different retrieval scheme might fail.
4) P4, line 12: Are the ice cloud properties retrieved by a tomographic approach as well? Is a 2D or 3D tomographic approach, this should be mentioned.
5) This paper highlights technical aspects of the two sensors, so it might be helpful to include a minimum of instrumental details like field of view (FOV), vertical sampling (step size), retrieval resolution, and horizontal sampling. Otherwise the reader cannot judge on statements like 'the vertical resolution is good enough to resolve these model fields' without any reference.
6) Figure 1: please give some details why MIPAS fails to retrieve the most northern profiles. This looks like an artefact of the data processing. I would not expect strong horizontal gradients or clouds at these locations, like argued in the manuscript. Why fails MLS retrieval at some high northern latitudes as well? You should explain in more detail in the manuscript why and where retrieval my fail for both instruments.
7) Figure 3: Any explanation for the cold bias in the area of the south-pole for MIPAS? This is a bit confusing, because especially in the cold areas PSC are frequent and hamper proper retrievals for this part of the season, consequently I would expect a warm bias?!

Technical comments:

1) Figure 2, 3, and 4: A superimposed mean tropopause location will substantially improve the information content of the figures around the tropopause.
2) Figure 2: please, give a few more details in the figure caption. E.g. Za-Zr or Zr-Za?
3) Figure 5: The numbers on the x-axis are not well readable and should be sparser and separated.
4) Figure 6+7: Temperature biases should be presented with a scaling of T (e.g. K x 10). Then the significant temperature bias in the tropics will become more obvious. In addition, a tropopause location might help here as well.

References:

Millán, L. F., Livesey, N. J., Santee, M. L., Neu, J. L., Manney, G. L., and Fuller, R. A.: Case studies of the impact of orbital sampling on stratospheric trend detection and derivation of tropical vertical velocities: solar occultation vs. limb emission sounding, Atmos. Chem. Phys., 16, 11521-11534, https://doi.org/10.5194/acp-16-11521-2016, 2016.

Spang, R., Arndt, K., Dudhia, A., Höpfner, M., Hoffmann, L., Hurley, J., Grainger, R. G., Griessbach, S., Poulsen, C., Remedios, J. J., Riese, M., Sembhi, H., Siddans, R., Waterfall, A., and Zehner, C.: Fast cloud parameter retrievals of MIPAS/Envisat, Atmos. Chem. Phys., 12, 7135-7164, https://doi.org/10.5194/acp-12-7135-2012, 2012.