Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-371-RC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Intercomparison of low and high resolution infrared spectrometers for ground-based solar remote sensing measurements of total column concentrations of CO_2, CH_4 and CO" *by* Mahesh Kumar Sha et al.

Ruediger Lang (Referee)

ruediger.lang@eumetsat.int

Received and published: 16 December 2019

Accurate ground based observations of carbon dioxide and methane are becoming increasingly important in the context of continuous satellite remote sensing validation. While ground based in-situ and remote sensing networks measuring CO2 and CH4 concentrations in the atmosphere exist, and are instrumental in measuring the continuous increase in greenhouse gases over the past decades, it has been recognised that the existing network of stations is still lacking with respect to its expected role in future monitoring and verification system-of-systems (MVS) of greenhouse gas emissions.

Ground based remote sensing instruments like the Total Carbon Column Observing Network (TCONN) FTS instruments measuring total column dry air mole fractions of carbon dioxide and methane (XCO2 and XCH4) are key for the validation and calibration of future operational satellite based observing systems, targeting greenhouse gas emissions. However, the current distribution of stations and their representability for the validation of global satellite based measurements, as well as their maintenance and operability for providing a continuous flow of quality-monitored data are providing big challenges ahead for the network to become the much aimed for fiducial greenhouse-gas reference measurements within an operational MVS context. One of the key-challenges is achieving and maintaining the very high accuracies needed – significantly below 0.5 ppm for XCO2 – in order to become useful for the MVS and the Cal/Val of its satellite components.

The paper by Sha et al. is an important and significant contribution towards the latter aspect, by addressing the question on how to secure high accuracies across the network, e.g. through travelling instrument standards, while identifying, correcting and potentially even reducing instrument and measurement biases. The key finding of the paper is, from my point of view, the potential of at least one of the instruments taking part in the campaign exercise at Sodankyla, Finnland, functioning as a "travelling standard" in a potential future operational network of reference FTS instruments as currently operated under the TCCON umbrella. The paper also addresses important open questions with respect to remaining systematic biases and points at, maybe even more significant, remaining issues in measurement bias and precision. The paper is well written, although I think the paper could benefit from some restructuring of the results sections. I therefore recommend the paper for publishing in ACP, but would like to highlight in the following a couple of observations (apart from some additional minor comments), which the authors, from my point of view, should address.

1) Ari-core comparisons and non-linearity effects

The discussion of the non-linearity effect found in the measurements of the Bruker

IFS HR reference instruments (referred to as TCCON "reference", 125HR), as well as the comparison to the AirCore measurements, as taken during the campaign at Sodankyla - which are considered to represent the "true atmospheric" state - are both presented only at the end of the major results section 5. This is first of all confusing, since the section on non-linearity effects implies that the non-linearity correction "has been applied to TCCON" data as a whole. At this stage, the reader wonders if this therefore now applies to all previous results, but the answer is probably not, since the results and different labelling then implies that the main comparisons results are not (yet) non-linearity corrected.

Second, the performance of the "reference instruments" with respect to what is considered the "true state of the atmosphere" represented by the AirCore measurements is an important result against which also the results of the measurements from the other systems have to be evaluated and interpreted.

So both aspects have to be taken together. I would therefore recommend to present the results on the 125HR instrument non-linearity and the performance of the reference instruments against the "truth" (5.8 and 5.9) at the beginning of Section 5 (and ideally then present only the non-linearity corrected results for the TCONN reference – if this is considered a stable result – in the comparison against the other systems). This would help to interpret the results of the other systems when compared to the "reference" 125HR better with respect to the AirCore "truth".

In addition, establishing the TCCON 125HR as a "reference" and therefore then talking about a "bias" with respect to the reference for the other instruments (and not in terms of "differences"), requires, in my view, presenting the AirCore results first anyhow.

2) EM27 systematic biases with respect to the two references

The EM27 (COCCON) instruments show a convincing performance with respect to the "reference" TCCON (125HR), both with respect to the systematic biases and precision (for the latter see below). This is alas true especially with respect to the non-

corrected TCCON results. It seems peculiar that this bias is even lower than -0.2 ppm for the non-corrected comparisons, but then gets worse for the TCCON non-linearity corrected ones (-0.73 ppm). At the same time, the low-resolution "reference" measurements (125LR) show a high bias to EM27 of roughly the same order of magnitude than the effect of the non-linearity correction. Considering that the low resolution 125HR are more comparable in terms of information content (and probably AKs, although those have unfortunately not been presented –see below), one wonders if the non-linearity correction should not also be applied to the 125LR measurements (or has this been done?), which then may lead to a consistency between EM27 and the 125LR. This would then also physically make a lot of sense considering the information content of both measurements. Also it would be important to rule out any link (e.g. in retrieval processing) between the standard 125HR measurements (not corrected) and the EM27 results, which potentially make them similar to the standard reference by default (and therefore more biased with respect to the corrected 125HR results and the AirCore "truth"). In this context, the processing algorithm ProFIT vs GFIT performances and their potentially relevant "peculiarities" are not much discussed at all. ProFIT is used both for the TCCON 125LR and the COCCON EM27s but not for the uncorrected TCCON "reference" it seems. Is the bias correction scheme used in ProFit maybe somehow related to TCCON measurements (or associated climatologies?).

I think it would be important to add a discussion of the (potential) relationship between the observed systematic biases between "reference" HR, LR, and non-linearity correction on one side, and the EM27 measurement results on the other side, which could shine some light on the underlying mechanisms.

3) Averaging kernels, low res versus high res

The differences in performance and measurement information content for the low-resolution measurements (125LR and EM27) are of high importance, especially with respect to the fact that some of the forthcoming remote sensing satellite based systems may be operated at lower spectral resolutions, i.e. being more similar to the LR measurements in terms of spectroscopy. Also some of the differences observed between AirCore truth, high resolution and low resolution measurements, as discussed under 2), may be interpreted in this respect. The performance of these lower resolution systems are of importance also with respect to the knowledge of spectroscopy and potential future research needs there. But mainly the differences in performance between LR and HR (also with respect to AirCore “truth”) may be better interpreted if the differences in their respective AK would be discussed and addressed - at least to some extend (and potentially also differences in GFIT and ProFIT a priori profiles if any – see also 2)). This could be done e.g., and ideally, in a corresponding figure to Fig. 17 (or adding such results to the latter).

4) Precision of TCCON versus COCCON

The striking difference in measurement noise (Fig. 18) between TCCON and COCCON EM27 is not much discussed it seems. Is this feature a question of retrieval algorithm performance (or/and applied constraints therein), spectral resolution (like the more stable LR with respect to HR measurements), or is it solely related to instrument noise?

Additional minor comments:

p. 2, l.36ff: “Carbon dioxide (CO2) and methane (CH4) are the two main components of the carbon cycle of the earth’s atmosphere. They absorb and retain the heat in the atmosphere causing the greenhouse effect and global warming” – I would remove “greenhouse effect” here. Otherwise, water vapour would have to be mentioned in this context.

Section 5.3: I think it would be would be very helpful to introduce here the logic of the following sections (titles content) to guide the reader through what is coming, apart from potentially shifting the AirCore results and non-linearity sections first (see above).

Figure 1 to 4: It would be helpful to use the same axis limits in the scatter plots.

Figure 16: A horizontal dashed line or similar at zero would really help to interpret the results. Especially since the range of the y-axis has to be quite large.

Section 5.4.4. I think here it should be highlighted and maybe discussed in context of the discussion on biases between TCCON, TCCONmod, LR and EM27, why the EM27 is so much closer to one than TCCON (see also point 2) above).

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-371, 2019.