

General comments:

The manuscript “Recent advances in measurement techniques for atmospheric carbon monoxide and nitrous oxide observations” by Zellweger et al. presents comparison data collected through the WMO GAW WCC station audits for CO and N₂O. The WMO GAW program has set strict network compatibility goals to ensure data from various monitoring sites and programs are consistent and that biases in the data will not unduly influence scientific interpretation of the data. Network compatibility is a difficult parameter to assess. One tool for assessing it is the station audits conducted by the WCC’s. While these are snapshots in time, they are a very rigorous comparison and provide valuable information on the consistency of data from various sites and networks. The data presented in this manuscript is valuable information for determining the significance of spatial gradients observed when data from various sites and monitoring programs are combined. In addition, the authors have used the comparison data from the audits to clearly show the advantages of newer analytical techniques and to point out areas that need further improvements. This should be of high interest to many involved in atmospheric monitoring. I recommend publication of this manuscript with a few minor suggestions to improve the document.

Specific comments:

- 1) Throughout the document – the WMO network compatibility goals should be referred to as “network compatibility” rather than just “compatibility” to distinguish the WMO usage from a strict metrological definition of compatibility. Or making the distinction upon first usage if that is preferred.
- 2) Page 2, line 11: Define Empa on first usage.
- 3) Page 2, line 14: Define who is the WCC-N₂O on first usage.
- 4) Page 4, line 10: “WCC-N₂O uses a set of TS traceable to a set of secondary standards ...” Are these standards truly secondary (compared directly to the primary standards) or are they actually the normal tertiary standards distributed by the CCL? Or does this “secondary” label relate to the hierarchy internal to the WCC-N₂O?
- 5) Page 4, line 11-14: Have all of the comparisons been reprocessed onto the current CO_X2014A and N₂O_X2006A scales or are they presented on the current scale at the time of the comparison? If the latter is the case I would suggest making the scale explicit by adding a column to tables 1 and 2 showing which scale was actually used. As a follow up I would ask if the comparisons change significantly if reprocessed onto the current scale? This may not be possible within the scope of this paper but would be of interest to data users who would like to use this information to understand potential biases between data from various providers.

- 6) Page 5, line 1: When did the parallel measurement approach begin?
- 6) Page 6, line 28: WMO network compatibility goals are no longer listed as “±”.
- 7) Page 7, line 3-6: Have any of the comparisons been reprocessed after the stations have had working standards re-calibrated and drift corrected? It would be very interesting to see if some of these larger offsets are improved with better calibrated standards. This would also provide a more accurate assessment of the bias in the station data but again may be beyond the scope of this paper.
- 8) Page 8, line 21: Is this statement supported by meta data from the stations, i.e. is there a record of the number of standards used for those early audits that would support this conclusion?
- 9) Page 11, line 7: The figure plots the data as Station – TI, the text has the offset as TI – Station. I suggest changing either the sense of the comparison in this sentence or in the figure to be consistent.
- 10) Page 11, line 12: I think the description of the AMY offaxis_ICOS instrument should be “enhanced performance” off-axis integrated cavity output spectroscopy rather than the stated “cavity enhanced”.
- 11) Page 11, line 17-19: I suggest putting in the values for each third of the time period to show how different they are and how much of the variability is due to the calibrations.
- 12) Page 17, Table 1: I suggest listing the CO scale for each comparison if they are not all the same.
- 13) Page 19, Table 2: I suggest listing the N₂O scale for each comparison if they are not all the same.
- 14) Page 22, Figure 3: As mentioned in the text there are only 2 comparisons of the same FTIR instrument. It might be good to show this by listing n values for each category or at least for the FTIR. I might also suggest keeping the same categories as shown in figure 2 (combining NIR-CRDS and QCL) to be consistent between the two figures but leave this to the author’s discretion.
- 15) Page 24, Figure 7 caption (and other time series plots): The caption says “(1 h data)”. I take this to mean the data from both instruments was averaged to hourly averages. If true I suggest making this point clearer.

Minor Technical corrections:

The following are places where I feel the writing could be improved by slight wording changes, however this is subjective and I leave it to the author's discretion.

-) Page 1, line 26: ... clearly indicate that drying of the sample air ~~is leading~~ "leads" to an improved ...
-) Page 2, line 6: ~~To take full profit~~ "To make full use" of these observations, ...
-) Page 2, line 25-26: Carbon monoxide shows a high temporal and spatial variability, whilst the detection of very small changes is needed for N₂O observations "due to its low variability."
-) Page 2, line 31-32: ... and a variety of methods ~~is~~ "are" now in use at atmospheric monitoring sites.
-) Page 3, line 10: The ~~recent~~ "recently developed" optical techniques for ...
-) Page 3, line 13-14: Such comparisons of traditional and ~~upcoming~~ "new" techniques are crucial for ...
-) Page 3, line 24-26: ~~WCC-Empa~~ is the designated WCC for CO (since 1997), and since 2009 a collaboration between WCC-Empa and the WCC for N₂O hosted by the Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research (IMK-IFU), ~~exists since~~ "has allowed" WCC-Empa ~~started including~~ "to include" N₂O comparisons during station audits.
-) Page 4, line 31: Include the most recent version of the GGMT recommendations "WMO, 2018" in the list of references.
-) Page 6, line 9-10: To judge whether the combinations of the resulting slope and intercept meet the WMO/GAW compatibility, ~~respectively~~ "or" extended compatibility goals, the method ...
-) Page 6, line 32: ... which deviated with a mean value of 0.994 ± 0.068 (1σ) not significantly "different" from one ...
-) Page 8, line 20: WCC-N₂O comparisons "also" showed ~~also~~ no significant bias ...
-) Page 9, line 8: ... laboratories using ~~other~~ calibration scales "other" than WMO-X2006A.
-) Page 9, line 19: ... slopes of the linear regression gives information ~~of~~ "on" the linearity and ...
-) Page 9, line 18-19: ... which is in "the" case of N₂O ~~in~~ "of" the same order or even larger than ...
-) Page 10, line 19: As discussed in the section 2.2, the factory "water" vapour correction ...
-) Page 11, line 30: So far, this has not yet "been" implemented at all ...

-) Page 11, line 32: ... which ~~proved~~ “proved” to be ...
-) Page 12, line 1: ... which results in ~~relative~~ “relatively” high uncertainties due to ...
-) Page 12, line 2-3: ... which ~~is using~~ “uses” a Bronkhorst Vapour Delivery Module (VDM) ~~able~~ to humidify a gas stream from a tank, might ...
-) Page 12, line 6: ... bracketed by absorptions “lines” of CO₂ and H₂O.