

Interactive comment on “Validation of MOPITT carbon monoxide using ground-based Fourier transform infrared spectrometer data from NDACC” by Rebecca R. Buchholz et al.

Anonymous Referee #2

Received and published: 6 November 2016

In this paper the authors present a validation of the MOPITT v6 data (TIR, TIR+NIR and NIR) against ground-based measurement (FTS from NDACC). Mean bias are given. Differences between the four MOPITT detectors are presented. MOPITT drift is calculated at the different stations used in the study, the excellent quality of the MOPITT retrievals is demonstrated. Recommendations for data assimilation and trend analysis are given.

This study is a considerable work as it covers the whole MOPITT period and uses the ground-based data of 14 NDACC stations. The paper is well presented (Equations are given when needed) and illustrated but the methodology used raise some issues. 4 specific comments are detailed below and should be addressed before publication.

C1

The subject of the paper is appropriate to AMT. I recommend the paper to be published, after addressing the major issues I raise below.

Specific comments:

1) Averaging the MOPITT profiles and the MOPITT averaging kernels

The comparison methodology is detailed in Section 3. The authors average the MOPITT profiles (co-located with FTS measurement), inversely weighted by the square of relative retrieval measurement error. Error-weighted is also performed for averaging kernel matrices and a prior profiles. RegridDED FTS profiles are smoothed by these averaged AK.

I have concerns regarding these averages. Are we allowed to average averaging kernel matrices? Geophysical representation problems may arise. What would M. Clive Rodgers say?

When validating satellite data with measurements with higher vertical resolution (FTIR, airborne profiles), one should smooth the high resolved profile by the satellite averaging kernels. This should be done pair by pair. Kerzenmacher et al. (2012) validated IASI CO retrievals with NDACC FTIR measurements. A given FTIR measurement was compared to all IASI retrievals according to the chosen co-located criteria. No averages were performed on the IASI retrievals. Same thing in de Laat et al. (2014) when validating MOPITT V5 NIR data with MOZAIC/IAGOS measurements: each MOZAIC/IAGOS profile measurement was assigned to multiple MOPITT measurements, resulting in 8980 MOZAIC/IAGOS–MOPITT collocation pairs (2002–2010 period). Again, no MOPITT averaging prior to validation. In Martinez-Alonzo et al. (2014), comparisons between MOPITT data, ACE-FTS data and HIPPO airborne measurements are performed. When comparing MOPITT data and HIPPO measurements, the validation work is done by pair, there is no averaging of the MOPITT retrievals.

The authors argue that it reduces the number of comparisons, that is for sure. They

C2

argue that using this method of averaging profiles and AK has been tested, against comparisons by pairs. Page 6, lines 7-11, they quote results from the supplementary material of Martínez-Alonso et al. (2014): “A comparison between the validation of averaged and raw MOPITT values against HIPPO and ACE-FTS measurements is available in the supplementary material of Martínez-Alonso et al. (2014). These authors concluded that the two methods produced equivalent results, and found that averaged MOPITT values produced bias of <1.2% against HIPPO and <0.8% against ACE-FTS, over raw MOPITT values. Therefore, averaging MOPITT values here could overestimate MOPITT biases by approximately 1%.”

a) It is worth noting that Martínez-Alonso et al. (2014) conducted the study using pairs comparisons, not averaging. The averaging method (called “approximation method”) is compared to the “full method” (described in the article) in a supplementary material. This method is presented to be relevant for the MOPITT Level 3 data analysis: “A computationally less costly approximation method for bias calculation, relevant for MOPITT Level 3 data analysis, is discussed in the supporting information.”

b) I'm puzzled but the results presented in these supplementary material (Table S2). Difference between two percentage values are expressed in percentage, which can be quite misleading. One should use percentage point (pp).

c) Buchholz et al. quote “bias of <1.2% against HIPPO and <0.8% against ACE-FTS” and in Martínez-Alonso et al., I read “According to our ACE FTS-MOPITT analysis, bias obtained using this approximation method (Table S1) versus the full method (Table 2 in main text) differ by $\leq 0.3\%$ in average; absolute bias differences are <0.8% (Table S2). Similarly, our MOPITT versus HIPPO-QCLS analysis indicates that bias obtained using these methods differ by <1.2% in average; absolute bias differences are $\leq 1.8\%$ (Tables 2 and S2).” If I understand Table S2 well, one should read that the differences between the bias (in percentage) obtained by the two methods differ by 0.8 pp maximum for MOPITT/ACE-FTS comparison and by 4.2 pp maximum for MOPITT/HIPPO comparisons.

C3

The authors should correct this part of the paper, when quoting the supplementary material of Martínez-Alonso et al. The authors should use the proper method for this validation work, which is working by pair, i.e. smoothing each FTIR measurement by one collocated MOPITT AK, as many times as there are collocated MOPITT retrievals.

If the authors want to perform this “approximation method”, they should clearly say this is not the usual way to perform validation and they should prove that we can rely on such biases with test-results for different stations.

Such a validation paper is very useful and will serve as a reference for future work. These biases are crucial as they will be used by the scientific community, like modelers to constraint models.

2) Removing the lowest level of a MOPITT profile when different topography

Vertical re-gridding is detailed in Section 3.3 and processes (when FTS surface pressure is larger or smaller than MOPITT's) are illustrated in Fig. 2. When FTS surface pressure is smaller than the MOPITT surface pressure, one MOPITT truncated profile is used for the comparison.

Removing information of the MOPITT retrieved profile, as well as in the averaging kernel matrix raises a lot of questions. Are we allowed to remove one layer of information, i.e. to remove one column and one line of the averaging kernel matrix? It doesn't make any sens. Using a truncated matrix would introduce biases on every levels of the profile to be smoothed. The authors should extend the FTS profiles instead, like it is done in Kerzenmacher et al. (2012) and in de Laat et al. (2014). In both papers, satellite a priori information is used to extend the FTS/airborne profile.

3) Presenting averaged averaging kernels matrices (Fig. 3, 4, 12, 13)

One should be extremely careful when presenting an average averaging kernel matrix. It is rarely done. Some time it is but if one is sure that all the averaging kernel matrices look exactly the same. What if one or two matrices show completely different sensitiv-

C4

ities? I am not sure to have ever seen that for seasonal averages (Fig. 3 and 13) and I am not comfortable at all with this plots. It would be straightforward for the authors to pick one representative AK per season, e.g. the 15th of one month. Of course it is subjective, but we trust the authors to choose one representative AK. A least this AK matrix would be realistic, the result of one inversion.

Authors could also present DFS time series.

4) Last specific comment but very important: The relative bias values should be calculated with respect to the FTS, not MOPITT, as FTS measurements are supposed to describe the “truth”. This is how it is done in Kerzenmacher et al. (2012). Another way is to present only absolute biases, like in de Laat et al. (2014).

General comments:

The y labels of your plots should be “Pressure (hPa)” and not “Altitude (hPa)”. The x labels “Averaging kernel value” should be “Averaging kernels”.

Page 1 line 10 : Replace "greater" by "higher" or "larger"

Page 2, line 15: "Emmons et al., 2009, 2004)" > "Emmons et al., 2004, 2009)"

Page 4, line 19 : "The most recent retrievals are used for this validation study" > "The most recent retrievals available in 2016 are used for this validation study"

Page 6, line 30: "Fig. 2 (a)" Please mention “inspired by Fig. 2 of Kerzenmacher et al. (2012)”, in the text or in the caption.

Page 7, line 28: “had” > “as”?

Page 7, line 28: “described” > “describes”?

Page 8, line 11: “DFS are provided for each MOPITT retrieval and are calculated for FTS measurements from the AK matrices”. It sounds like MOPITT DFS are not calculated from the AK matrices. Consider “For each MOPITT retrieval and FTS measure-

C5

ment, DFS are calculated from the AK matrices”.

Page 8, lines 12-15: These two sentences should be removed. Unclear. Yes, theoretical DFS may be higher and also smaller depending on the a priori covariance matrix. We know that the authors don’t perform instrument comparison.

Page 9, line 1: “(apart from at Ny-Ålesund and Bremen)” > “(apart from at Ny-Ålesund, Bremen and Lauder)”

Page 9, line 2: “(except at Ny-Ålesund)” > “(except at Ny-Ålesund, Lauder and Arrival Heights)”

Page 9, line 8: How do you perform the normalization? I see that some values are above 1.

Page 9, lines 22-30: This paragraph is not very convincing. Above 30°N, there are low and high correlations for the 3 MOPITT products. I don’t understand the first sentence (line 21).

Page 13, lines 4-5: What would the mean drift be if we remove the stations above 60°N?

Page 18, line 15: “instrument.” > “instrument,” (remove dot)

Page 20, footnote: “125 HR” > “125HR” (remove space)

Page 21: “where not enough pixels”: how many is enough? Are these median DFS calculated from the MOPITT retrievals used for the validation (tables 3a, 3b, 3c)? The numbers of retrievals are those indicated in Tables 3a, 3b and 3c on the “# obs” line? It should be mentioned.

Page 25: “Values for land validation is” > “Values for land validation are”

Page 32: In the caption with the symbols: Replace “pixels 2-4” by “pixels 2 to 4” or “pixels 2-3-4” or “pixels 2→4”

C6

Page 33: “bias (second row)” > “bias ((MOPITT-FTS)/MOPITT, second row)”. But the method should be (MOPITT-FTS)/FTS, see specific comment #4.

Page 34: The bias (MOPITT-FTS)/MOPITT is not the same as the retrieval bias. Modify the caption.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-241, 2016.