Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-165-RC2, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Retrieval and validation of METOP/IASI methane" *by* Evelyn De Wachter et al.

Anonymous Referee #2

Received and published: 21 July 2017

SUMMARY

The paper describes the new CH4 product produced by BIRA retrieved from IASI satellite measurements, reconstructed from principal components. The retrieval set-up is described along diagnostics such as averaging kernels and DOFS. The retrievals are compared with NDACC ground-based sites using the standard transformations required to handle the the different vertical sampling.

GENERAL COMMENTS

My main concern is that, having read the paper in some detail, I am still unsure of the quality of the measurements, regarding both the influence of the a priori and the error analysis.

1) The authors state that the CH4 a priori comes from a WACCM climatology but does this a single global/time average, is it zonal/seasonal, or something else? Assuming it has some latitude/seasonal dependence then the question arises: to what extent are the latitude/seasonal cycles depicted in the maps and NDACC comparisons simply reproducing the a priori variations rather than the CH4 retrieval? One way of testing this would be to subtract the a priori from both the data and the NDACC sites and examine the statistics with which *variations* from the climatology are reproduced. This would be a more accurate measure of the added value of the IASI measurements compared to just assuming climatology. Another test, assuming the climatology has no trend, would be to compare time series with the annual cycles removed, but that would probably require more than 3 years of data. While there are plots showing the a priori (smoothing) error contribution to the individual profile levels, there is nothing equivalent for the a priori contribution to the 3-17km partial column which is presented for NDACC comparisons.

2) For the 'reference' paper for any new satellite dataset I feel there should really be a quantitative 'bottom up' error analysis, ie an formal assessment of the magnitudes of the various error terms based on internal tests, which can then be compared by the authors (or other users), to independent data for a 'top down' approach. That is really lacking in this paper. The errors would presumably include contributions from all of the following: instrument noise, errors in retrieved temperature, surface emissivity errors, residual cloud contamination, a priori biases, concentrations of interfering molecules (including HDO) PCC reconstruction error, spectroscopic errors - see specific comments below. These all seem to be handled in the retrieval as a diagonal covariance matrix of fixed size approx 5x noise, and in the NDACC comparisons not assessed at all. An assessment of random error, or precision, could simply be obtained from the SD of the 2x2 pixels or some other small area where it is assumed the CH4 concentration is relatively uniform. It would be useful to have some figures for all these terms, even if only upper limits, for the 3-17km partial column which seems to be the basic product.

And a couple more suggestions, which I leave to the authors' to include or not:

3) Since both day and night are processed separately, although one would not expect the CH4 to change significantly over a diurnal cycle, I would also like to know if the day and night zonal means are self-consistent within the error budgets.

4) A simpler error analysis could be to use the WACCM profiles to convert both NDACC and IASI to total CH4 column amounts.

SPECIFIC COMMENTS:

P1,L1: The product is described as 'global' but results are only shown for 60S-70N.

P1,L4: 'retrieval uncertainty ... less than 4%'. Be a bit more specific about whether this refers to precision (ie random error), or accuracy (total error, including systematic biases) - ideally quote both.

P1,L10: 'absolute differences ... less than 1%'. Again, not clear what this means, 'Absolute' usually means irrespective of +/- sign, and difference could be anything from single-profile match ups to mean bias throughout the whole dataset.

P2,L29: The other IASI CH4 products, currently cited only in the Conclusion, should also be mentioned here at the start to put this work into its proper context.

P3,L4: The final section also contains a significant description of proposed future work.

P3,L6: Pedantically, since there is mention of MetOp-B and MetOp-C, there should be some mention of MetOp-A. And presumably it's not just IASI that will provide a 15 year dataset but all, or at least most, of the other MetOp payload instruments as well.

P3,L14: 'four spectral bands' - I thought there were only three? (breaks at 1210 and 2000cm-1). Also, emphasise that these actually provide a continuous spectrum, without gaps (unlike several other FTIR instruments).

P4,L2: No reference given for Drummond et al.

P4,L8: The EUMETSAT L2 skin temperature is also used as an input (P4,L23).

P4,L11: Is this the dataset commonly referred to as the 'Wicsonsin' surface emissivity data? And what is used over the ocean?

P4,L18: Is there any evidence that the 10% cloud fraction does not contribute a significant error? One might hope that the skin temperature, or other non-CH4 elements of the state vector, will absorb any residual cloud, but that also depends on the tightness of the a priori constraints. I expect the EUMETSAT skin temperature is retrieved with a very small error, so may not allow for much cloud-compensation within the CH4 retrieval. A plot of bias and SD v CH4 a priori, or zonal mean, or NDACC, as a function of cloud percentage would answer this.

P4,L15: What molecule do these 'problem' features belong to? Are they the CH4 Q-branches? And does the forward model include CH4 line-mixing? Also, have the effects of the variation in HDO been considered?

P4,L16: Setting the radiometric noise to infinity (or very large) is the mathematical way to exclude spectral points from the retrieval, not setting the noise to zero. And if these points are excluded from the fit, in what sense is 'no information lost'?

P4,L22: Is the 23-level state vector an arbitrary choice or is it set by the EUMETSAT L2 or WACCM profiles used as a priori data?

P4,L28: Is there a reason for imposing a uniform a priori uncertainty for H2O rather than using the uncertainty associated with the EUMETSAT H2O product that is actually used for the a priori profile? Even if just for scaling the diagonal elements.

P4,L24: I assume this means that the climatology is some sort of average of the WACCM model output - global,annual mean? monthly zonal mean? - while the covariance represents the model statistical variability about this mean. If the WACCM output is on the same levels as your retrieval grid that's quite straightforward, but if it isn't then there are a few more steps involved.

P4, L29: The characterisation of forward model errors as a simple scaling of the nominal noise diagonal matrix is certainly convenient but requires a little more justification than just the plain statement presented here. The fact that PCC reconstructed radiances agree within the nominal radiometric noise (Fig 3) for a single spectrum is not in itself sufficient to demonstrate that the reconstruction error is negligible: unlike the radiometric noise the reconstruction error is likely to have significant and persistent correlations with the spectrum itself, so unlike the random noise, the impact will generally not be reduced as 1/sqrt(n), where n is the number of spectral points used.

P5,L1: The shape of the averaging kernels presumably depends significantly on the surface temperature contrast, but that information is not given with the figure. It's hardly surprising that, with DOFS$\sim$1, the profile uncertainty is dominated by the a priori error, or 'smoothing error'. A more useful figure would be the error in a quantity which more realistically represents the retrieval information, eg integrated total or partial column amount, and how this compares with the a priori uncertainty. It certainly makes for a more meaningful comparison with other CH4 retrievals which are on different profile levels.

P5,L2: Here the effective sampling range is defined as 2-16 km but elsewhere 4-17 km columns are used.

P5,L31: From Google maps I conclude that these 3 locations are all over the sea - but it would be helpful to state that in the text or the figure caption. If you have only three examples, I don't think it is useful to present both northern and southern mid-latitudes, which one would expect to be similar (particularly near the equinox). It would be more informative to have different land-air temperature contrasts instead, representing the max/min values shown in Fig 6.

P5,L10: 'thermal contrast' needs to be defined.

P5,L20: The correlation plot Fig 4 is used as evidence that the PC reconstruction error is negligible. However, this is a comparison of absolute CH4 values which, as already

demonstrated in Fig 2, are closely constrained by the a priori, so a good correlation may only represent the fact that the measurements have little influence on the a priori. I would be more convinced by a plot of the correlation of the *differences* with respect to the a priori profile.

P6,L20: 'one independent piece of information is retrieved with good sensitivity'. While DOFS $\sim$ 1, it would be more useful to have some idea of, for example, how this translates to a reduction in the a priori uncertainty for a 4-17km column.

P6,L23: What limits the latitude coverage? Here it says 60S-70N but two of the NDACC comparison sites are higher than 70N.

P6,L23: 'binned on a 1x1 deg grid'. So does this mean the plotted points represent not just an average of the four pixels but an average of all the pixels within the 1x1 box? Or is the binning some other process? And what happens if, say one of the 4 pixels is flagged as cloud-contaminated. Is the average then made of the remaining 3 or is this set of 4 pixels excluded?

P8,L24 gives IASI a priori systematic component as 2% of a priori value. Where does this come from? Just the error in WACCM? No systematic component of IASI retrieval uncertainty is considered.

References: not in alphabetical order, some missing publication year, and inconsistently formatted. Patra 2009 listed in references but not cited in text

Fig 1: Rather than just show a generic piece of spectrum it would be helpful if this figure was also used to show the indidividual contributions of different molecules to this spectral region (eg separate panel with same x-axis).

Fig 2: On the left panel it would be helpful to also have the a priori error bars plotted for comparison.

Table 1: This lists a priori information as WACCM, but that is only for certain elements of the state vector. 'IASI L2' should include the word 'EUMETSAT' for consistency with

the text. There should also be something about the a priori covariance information.

Table 2: I'm surprised at the spread in systematic errors in Table 2 for the various NDACC comparisons. Assuming this represents a combination of the NDACC systematic error budget and the 2% systematic error assumed for the retrieval a priori, this variation must mostly come from the NDACC data. Yet Sepulveda et al (2014) quotes a figure of 2.5% which is largely spectroscopic uncertainty (and therefore common to all sites). And the fact that these systematic errors are all much larger than the biases suggests something wrong.

TYPOGRAPHICAL ERRORS/SUGGESTIONS

P1,L3: 'usefulness'

P2, L2: 'greenhouse gas-intensive' - the hyphen here seems to change the natural coupling of words and might be better removed.

P2, L3 (and elsewhere): 'ppb' - suggest 'ppbv', to distinguish from 'ppbm'.

P2,L5 (and elsewhere): 'Bulletin , 2016' - remove space between reference name and comma

P2,L31: As an acronym, I think 'lidar' now has the same status as 'radar' and need not be capitalised or expanded (at least not for the benefit of AMT readers)

P3,L6: no comma required after 'MetOp'.

P3,L7: 'successive series' - a series is, by definition, successive.

P3,L15: 'prediction', singular, seems more usual in this context.

P3,L19: 'In the longer term...'

P3,L19: 'programme', unless you choose American English.

P3,L20: 'observations' (plural).

P3,L3: 'a modular software'. While 'the software' is acceptable, 'a software' doesn't sound right. I suggest just '.. is modular software ..'.

P4,L3: 'on-board' inconsistent hyphenation cf P3,L6 & L8.

P4,L8: 'range' mentioned twice in same line

P4,L12: 'and the residual'

P3,L14: 'certain spectral ranges ... spectral band'. Previously the 1210-1290 was described as the 'spectral range' but it seems that now becomes the 'spectral band' and 'spectral range' now refers to the problematic spectral features.

P4,L22: 'Table 1'

P4,L28: 'its a priori'. Firstly, unclear what 'its' refers to and, secondly, 'a priori' suggests a retrieved quantity. I suggest just replacing 'its a priori' by 'their' or 'with'.

P4,L32: 'vapor' here, but 'vapour' elsewhere, eg P4,L8.

P5,L3: 'diplays'

P4,L5: 'with of an'

P4,L11: 'su rface'

P5,L17: 'Tb/year' should be reduced to 'Tb' given the rest of the sentence.

P5,L19 & L20: 'm2 sr m-1' should be 'cm2 sr cm-1'

P5,L24: 'negligibles'

P6,L30: 'van' - should be 'Van' since starting a new sentence.

P7,L17: 'less' - should be 'fewer'

P8,L5: 'alitude'

P8,L30: 'slighlty'

P12,L32: 'Forc- ing'

Fig 1: insert space: 'sr cm'

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-165, 2017.