Referee comments Feb 2024

Comparison of the H2O, HDO and delta D stratospheric climatologies between the MIPAS-ESA v8, MIPAS-IMK v5 and ACE-FTS v4.1/4.2 satellite data sets

by K de los Rios et al.

Overview

The paper compares three different datasets for (primarily) stratospheric H2O, HDO and the derived delta D for the period 2004-2012 using latest processed version of data from the IMK and ESA retrievals of MIPAS data, and the ACE-FTS. This builds on previous work by other authors who used older and more limited datasets. In particular the extended time period illustrates how the H2O 'tape recorder' effect in the equatorial stratosphere is represented with very different levels of clarity.

Main Comments

Overall this seems to be more of a technical report rather than a scientific paper: the data are read, the recommended screening is applied, the results are plotted, analysed and discussed. There's nothing wrong with it, as such, but the authors miss some opportunities for providing new insights.

I list a number of suggestions, which the authors may wish to consider, which I think would improve the paper.

1) The descriptions of the algorithms and retrieval characteristics behind the different datasets read very much like extracts from the separate source papers, including many obscure technical details. I prefer to have seen a single, shorter and, most importantly, original description highlighting the similarities and differences where they might be relevant to the results presented, which would also show that the authors have applied some critical understanding of the technical details rather than simply relaying the information to the reader to evaluate.

2) The averaging kernels, in particular, seem key. There really should be a figure allowing these to be compared rather than verbal descriptions of the two MIPAS AKs and nothing at all regarding ACE-FTS. I couldn't find any mention of whether the ACE-FTS retrievals use any kind of regularisation and/or climatological a priori constraint, and I would expect the authors to have at least asked themselves the same question.

And, from the AKs of H2O and HDO, one point that could have been developed is how to determine the AK for delta D.

3) Given that MIPAS-ESA and MIPAS-IMK both use fundamentally the same set of observations, the comparisons would have been simpler if MIPAS profiles were *only* used when data from both processors were avaiable. There would be some loss of data from the UTLS-1 and AE modes, but negligible compared with the advantage of eliminating sampling bias.

4) While the systematic errors for both the MIPAS-IMK and MIPAS-ESA retrievals are dominated by spectroscropic uncertainties, and it is established that there are some difference in the H2O spectroscopic data used, it seems most unlikely that the spectroscopic data are so independent that they account for much of the difference between these two. A plot of the H2O and differences using the two spectroscopic databases, with microwindows marked, say for 20km altitude would have helped answer this.

5) Another point that wasn't addressed was whether there was any significant

difference between the day and night profiles for the two MIPAS datasets.
H2O, in particular, has a strong non-LTE signature in the stratosphere
and this could lead to spurious day-night differences in the results
(with, presumably, the night profiles being less affected).
This could simply be incorporated into Fig 3 and may explain some of the
difference between the MIPAS-ESA and MIPAS-IMK processors.

6) For the debiased SD (Fig 4) this should show some correspondence with the
sum of the random errors associated with the individual profiles, ie
sqrt(e_1^2 + e_2^2) where e_1 and e_2 are the reported random errors.
It would be useful to have these plotted on the same figure for comparison.
Another diagnostic would be to show the actual SD of each dataset about
its mean. A certain amount of this would be atmospheric variability - presumably
the same for all three instruments, but subtracting some variability due to the
regularisation while adding variation due to the instrument random noise.
One could go further: if it is assumed that the bias is constant, three sets
of comparisons between three datasets is enough information to assign as SD
to each dataset. Thus from the debiased SD results, one can empirically
determine the actual SD of each dataset about its mean bias.

7) Sections 3.1.1 and 3.2 should be an appendix. While useful for the
purpose of defining terms, this is just standard statistics.

Minor Comments

L34: This is not a particularly controversial or original statement, so I
would suggest "(e.g., Hegglin et al ... "

L42: This statement probably does need a supporting reference, instead of
just 'Scientists discovered it'. Rosenlof et al will probably suffice
but use that at the start of the paragraph.

L23: lower biases for HDO, but not delta D. But you should also mention that
the smallest bias for H2O is between MIPAS-IMK and MIPAS-ESA.

L24: I would interpret a 'meridional cross-sections' to be cross-sections
at a particular longitude, but here you really mean 'zonally averaged
cross-sections'

L61: "Atmospheric limb-sounding" would be better than "Limb Earth probing".

L71: "highly reliable" is unecessary and rather subjective. Perhaps just say
"made regular WV observations..." (and use past tense).

L75: Since you have already described Odin, I think it would be useful at
this point to briefly mention that MIPAS made continuous observations
of the infrared limb-emission, obtaining around 1000 profiles a day with
global coverage, while ACE-FTS used solar occultation which gave typically
28 profiles a day split into two narrow latitude bands (which varied
throughout the year).

L83: Suggest "e.g.," or "i.e.," instead of "like".

L85: references to web-pages should probably appear as usual citations rather
than directly within the body of the text (unless AMT has its own rule on
this). Also L156, L174,

L85: Suggest 'latest' rather than 'last' - they may want to produce another.

L85-L102: There is a rather confusing mass of detail over specific datasets
here, much of which is repeated in Section 2. For this part, the
introduction, the emphasis should be on clarity so try to remove some
of the obscuring details which are covered in Section 2.
(Even in section 2 I feel it would be more clearly represented in a table

listing dataset, date range, products compared and the reference).

L113: v4.2?

L136: What is the MIPAS FOV width? (and that of ACE-FTS?)

L148: A table listing the microwindows would have be useful

L149: I don't know why information on molecules other than H2O is provided here - are they expected to have a significant contribution to the results? OCS, for example, only has a significant absorption feature around 2100cm-1, well outside any of the spectral regions used for the H2O retrieval?

L154: Since these links refer to images you should include them directly in the manuscript (or else replot the data), otherwise this paper will be incomplete if the links ever disappear.

L157: 'is about 3 km'

L236: Since both MIPAS processors have used the same set of spectra, the differences in time/location are purely due to how these values are assigned to the resulting L2 profiles.

Fig1: Why would the MIPAS-ESA and MIPAS-IMK profile locations be any different?

L246: This is inconsistent. Is the grid from 0-70km or 1-70km? Is it 1km spacing up to 44km or up to 46km?

L260: Another approach you could have considered is averaging ln(H2O) and ln(HDO) (assuming the values are always constrained to be positive). Since there is a strong variation with height in the tropopause this avoids biasing towards large values in the average. This may explain some of the behaviour of the MIPAS-ESA HDO profile at low altitude in Fig 3.

L326: "along the stratosphere" - what does this mean? Along usually indicates a horizontal direction.

Fig 3: With >1000 profiles compared over most of the altitude range I think we can assume that the standard error will be negligible, so the error bars just clutter the plot.

Table 1: this would be clearer if the columns were lined up, eg split each into two columns, min and max, and use + signs for positive values. Also I don't think more than 1 significant figure is justified, certainly not 4 as used for the absolute bias of delta D.

L406-415: Table 1 already summarises the previous plots so I don't think yet more text summarising Table 1 is required.

Fig 5: "during I boreal"

Fig 5: "The climatology is based ..." - presumably you are referring to these plots as "the climatology" but the plots are introduced as "latitude-altitude cross-sections" not as a "climatology". Perhaps if you write "This climatology is based..." it establishes what you meant.

Fig 6: These plots might be clearer if presented as deviations from the mean profile. It's hard to distinguish the various shades of blue/green which contain the signal for H2O and HDO.

Typographcial inconsistencies
- Both upper and lower case for version, eg v8 in title, V8 in abstract

- Water vapor (eg L15) and water vapour (eg L16)

- Data set (eg title) and dataset (eg L69)

- Use '--' in LaTeX to indicate a range of numbers, not hyphens (eg L157).
