

## Interactive comment on "Intercomparison of stratospheric ozone profiles for the assessment of the upgraded GROMOS radiometer at Bern" by S. Studer et al.

## Anonymous Referee #1

Received and published: 5 September 2013

The paper is aimed at assessment of the upgraded GROMOS radiometer in Bern. GROMOS belongs to the Network for the Detection of Atmospheric Composition Change, thus the objective of the study is of high importance.

However, the presentation style should be improved, from my point of view. The paper contains several shortcomings. Many figures need revision, as well as analyses. My main and detailed comments are below.

MAIN COMMENTS

1) The new FFTS and the original FB backends measured in parallel for over 2 years,

C2394

for data harmonization/calibration. Keeping in mind the main objective for the study - to get a long-time ozone record - one would expect a detailed comparison of these perfectly collocated data. However, this part is missing (comparisons of FFTS and FB are scattered in sections and presented together with other datasets).

From my point of view, it would be very advantageous for the paper to present comparison from the old and new GROMOS detectors in a dedicated section/subsection. Focus should be on (i) overall agreement, (ii) temporal evolution of profile differences, seasonal cycles, (iii) representation of diurnal variations. This section should contain also discussion of differences, underlying reasons, and possible solutions for data homogenization.

2) The choice of satellite instruments. Please explain your choice of satellite instruments. As follows from section 3.2, both MIPAS and SABER have rather significant biases with respect to ozonesonde and lidar observations. Why you prefer using these data (are they conclusive for validation?), and not using the data from the instruments, which have small biases with respect to ground-based observations, like OSIRIS and GOMOS?

Related to ACE-FTS data, comparing monthly mean data with individual ozone profiles looks very strange. This cannot be conclusive – vice versa, can be misleading. I would recommend avoiding such kind of comparisons, and simply not including this dataset due to a very small number of collocated data. "No data - no show".

3) It is well known that the quality of ozone profiles in ERA-Interim is insufficient. The figures presented in the paper demonstrate this clearly. But again: the main objective is assessment of the new GROMOS instrument. Thus only reliable data should be used. (Validation of ERA-Interim is definitely outside the scope of the paper). I suggest removing all comparisons with ERA-Interim.

4) Averaging kernel smoothing (or its description) is incorrect. In order to get the profiles in the same vertical resolution, the 1st profile should be convolved with the averaging kernel of the 2nd profile AND the 2nd profile should be convolved with the averaging kernel of the 1st profile. Alternatively, if use Eq.(1) only, AVK should represent the DIFFERENTIAL smoothing (i.e., the difference in vertical resolution).

For ozonesonde data this second smoothing (low-resolution profiles with the averaging kernel of high-resolution profile) can be ignored, the difference in vertical resolution is too large. However, for satellite data it should be taken into account.

## DETAILED COMMENTS

1) Short title does not summarize the paper. "Assessment of GROMOS radiometer upgrade" would be better, from my point of view.

2) p. 6099, last line "measure ...ozone profiles with the same accuracy". Is there the instrument ageing? How does it affect the accuracy of ozone profiles?

3) P. 6100, I.23-26: it is repetition already 3rd time what is said above.

4) P.6109, I. 20-21. "only coincident GROMOS profiles between 5 and 9 p.m. in winter and between 7 and 11 p.m. in summer have been taken into account " - Why?

5) P.6111, I.6-7: "Two annual means are shown to see if there is a possible difference between the years." Why do you expect the differences between these years? Please explain. If no reasons – results for these two years are very similar - combine these two periods into one and reduce the number of figures/panels. In the paper, it will be then sufficient to notice once that you tested two year separately and found very similar results. This comment is also relevant for comparisons with other datasets.

6) P. 6111, I. 13-14. Please replace "slightly less ozone" with quantitative estimates. Furthermore, I would not rate a bias  ${\sim}5\%$  as "good agreement".

7) P.6112, I.10: "no seasonal biases":  $\rightarrow$  No seasonal dependence of negative bias?

8) P.6112. Relative standard deviations: what is illustrative purpose of this parameter? Its inclusion requires more discussions.

C2396

9) P.6112-6113, lines 25-26-1 -2. This should in the figure caption, not in the paper text.

10) P.6114, lines 24-25: "Error bars are standard deviation of the averaged data". The standard error of the mean is computed as the standard deviation divided by the square root of the number of measurements (for uncorrelated data).

11) P.6116, I.5-6: "The agreement ... improves with time ". I do not see this improvement.

12) P. 6116, I. 20-22. I would not rate R=0.64 as a good correlation.

13) Fig.2. I suggest combining 2 years into one period (2 subplots). In the panel with profiles, use contrast colors, clearly visible lines.

14) From Figures 3 and 4, I suggest extracting only relative difference plots, and put them all into one panel (colored lines for different seasons). So, instead of 8 panels, one panel only – without any loss of information.

15) Figure 5: see above comments on datasets and parameters. In addition, it would be advantageous to show bias curves (mean relative difference) with errorbars for the bias estimates (std/sqrt(N))

16) Fig.7 Use contrast colors, remove white spaces.

17) Fig. 9 : Above comments + remove white spaces (put subplots closer to each other)

Interactive comment on Atmos. Meas. Tech. Discuss., 6, 6097, 2013.