Response to anonymous referee #2

Lorena Moreira

August 26, 2015

We would like to thank the Referee #2 for the careful reading of our manuscript and for giving constructive comments which substantially helped to improve the quality of the paper. All proposed objections and suggestions have been taken into account and discussed. Below we try to answer every comment.

Specific comments

1. **Comments from the referee:** Section 1: Introduction

   A summary of the historical use of ground-based microwave measurements to detect ozone trends is missing. Reading the present paper, it seems that no study has been made about this except for the Dumitru et al. (2006) paper. But microwave measurements from several stations have been used in the WMO (2014) report. How the present study fits in this context? Is it the same methodology that is used for the Bern station in WMO (2014)? And what about the other stations used in WMO (2014)? Any other publications,...?

   Author's response:

   Regarding the historical use of ground-based microwave measurements to detect ozone trends what we have found on the literature are Steinbrecht *et al.* (2006 and 2009) and the WMO (2014). In fact, since Steinbrecht *et al.* (2006) we are not aware of any other ozone trend study based on ground-based microwave measurements, therefore, the purpose of this paper is to present a new trend estimation based on stratospheric ozone profiles measured by a ground-based microwave radiometer at northern midlatitudes.

   As mentioned in Section 6: Results and discussion, our trend result is in agreement with those reported in WMO (2014). The WMO (2014, Table 2.4) reported a statistically significant ozone increase of (3.9 ± 1.3) (% decade^{-1}) at 40 km in the upper stratosphere at northern mid-latitudes (35-60 °N) over the 2000-2013 period. This ozone trend value is based on observations from various space-based and ground-based measurement instruments. A slightly different ozone series of the GROMOS instrument contributed to the ozone trend estimation of WMO (2014). The ozone trend of our study cannot be directly compared to WMO (2014) since the method of trend determination, the time window and the measurement region differ. However, the ozone trend at Bern (about 3.0%/decade at 40 km) agrees well with the ozone trend at northern mid-latitudes (about 3.9%/decade at 40 km) as reported by WMO (2014).

Author's changes in the manuscript: ~~Ozone time series from the GROMOS microwave radiometer were used to assess the stratospheric ozone trend above Bern and for validation of satellite observations (Dumitru *et al.*, 2006~~ The concerns regarding anthropogenic depletion of stratospheric ozone increased the necessity for precise and accurate measurements to monitor long term trend in this specie. (Parrish *et al.*, 1992). Passive millimeter wave radiometry has been used to monitor the vertical distribution of atmospheric trace gases since the early 1970s (Parrish *et al.*, 1992). The need for continuous monitoring of the stratospheric response to anthropogenic trace gas releases, performed by a well defined set of instruments, led to the foundation of the Network for the Detection of Stratospheric Change (NDSC) (now Network for the Detection of Atmospheric Composition Change - NDACC) in 1991. The ozone radiometer GROMOS is part of the NDACC~~.~~, hence our entire dataset is available via http://ftp.cpc.ncep.noaa.gov/ndacc/station/bern/hdf/mwave/
Ozone time series from the GROMOS microwave radiometer were used for comparisons with lidar, ozonesondes and collocated satellite observations and for detection of long-term trends (Dumitru *et al.*, 2006; Steinbrecht *et al.*, 2006; Steinbrecht *et al.*, 2009; Keckhut *et al.*, 2010; van Gijsel *et al.*, 2010; Studer *et al.*, 2013; Delcloo and Kreher, 2013; WMO, 2014).

(Added at the end of the Introduction) ... with a representative station in central Europe. Trend studies of ozone profiles based on ground-based microwave measurements are rare. In fact, since Steinbrecht *et al.*, 2006 we are not aware of any other publication, therefore, the purpose of this paper is to present a new trend estimation based on stratospheric ozone profiles measured by a ground-based microwave at northern midlatitudes. The present study is organised as follows: ...

2. **Comments from the referee:** Section 2.3: Retrieval procedure (p.16376-16378)
I recommend to include more details on the characterization of the ozone profiles obtained with the OEM:
a) OEM requires the use of a priori information, which includes a priori profile (discussed in the paper) but also a priori covariance matrix which is not mentioned in the paper. Can the authors provide this missing information?
b) The authors provide in the text the vertical resolution in the stratosphere (8-12km) and lower mesosphere (20-25km), but the reader would have a more global picture with the knowledge of the numbers of degrees of freedom for signal and a plot of the averaging kernels as often provided when OEM is used. This would help to interpret Fig.8, underlying the fact that the actual knowledge of the vertical structure of the ozone trends is much less than what a reader could interpret from the trend profile in Fig. 8. Also, for information, how much of the a priori contributes outside of the 20
c) Errors: the smoothing error values (in function of altitude) should also be provided. The link between the errors discussed in this section (2.3) and the uncertainties in Section 5 is not clear. Do I understand well that the error budget provided by OEM (l.9, p.16377) is not used in the "uncertainty analysis" section?

Author's response:
a) As diagonal elements of the a priori covariance matrix we assume a relative error

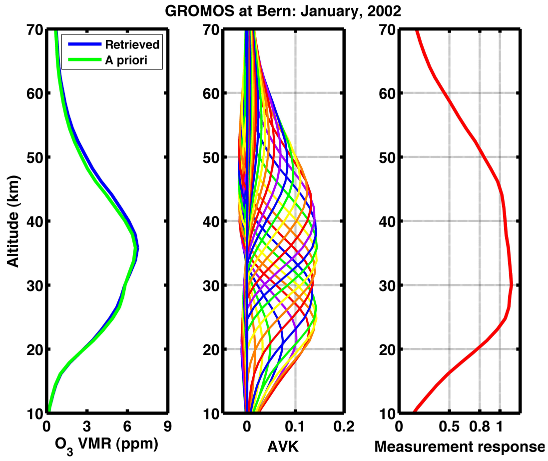

Figure 1: Example of an a priori profile and a retrieved ozone profile (left panel), averaging kernels (middle panel) and the measurement response (area of averaging kernels) (right panel) of the GROMOS retrieval for January, 2002

around 35% at 100 hPa. The error decreases in the lower stratosphere up to 28%. Then it increases linearly from 35% in the upper stratosphere to 70% in the lower mesosphere. The off-diagonal elements exponentially decrease with a correlation length of 3 km.

b) Figure 1 shows an example of an a priori profile (green line) and retrieved ozone profile (blue line) in the left panel, the averaging kernels (AVKs) (matrix columns in (ppmv/ppmv)) in the middle panel and in the right panel the measurement response (area of averaging kernels) of the GROMOS retrieval for January, 2002. An estimate of the a priori contribution to the retrieval can be obtained by the area of the averaging kernels (measurement response). Between 20 to 52 km (50 to 0.5 hPa) the measurement response (right panel in Figure 1) is higher than 0.8 what corresponds to an a priori contribution less than 20%. Consequently, the retrieved ozone values at these altitudes are predominantly based on the measured line spectrum.

c) As described in Subsection 2.3: Measurement principle and retrieval procedure, the total error provided by the OEM during the retrieval procedure includes systematic error, random error and the smoothing error. The systematic error originates from the tropospheric correction, calibration error due to systematic error in the load temperatures, errors due to baseline features, spectral parameter, etc. The smoothing error is caused by the low altitude resolution of the data retrieved from the measurements; hence smoothing error is almost time-independent in size and sign, i.e. is an almost constant error. We do not consider any kind of systematic error during the trend analysis since they cancel out, thus, it does not have any impact on the trend analysis. Accordingly, the only contribution from the OEM error budget to the uncertainty analysis in Section 5 is the random error. The random error is mainly due to the propagation of the thermal noise of the brightness temperature into the ozone profile. In Section 5 we recall this random error as observation uncertainty.

Author's changes in the manuscript: The vertical resolution depends upon altitude and can be estimated from the full width at half maximum (FWHM) of a kernel line. The averaging kernels (AVK) and the area of the averaging kernels (AoA) of GROMOS are shown on the middle and right panel of Figure 1. For GROMOS the vertical resolution lies generally within 8-12 km in the stratosphere and increases with altitude to 20-25 km in the lower mesosphere. ~~In the case of GROMOS, the ozone VMR profiles are retrieved with less than 20% of~~ An estimate of the a priori contribution ~~from 30 to 0.3 hPa (altitudes from about 25 to 57 km).~~ to the retrieval can be obtained by the area of the averaging kernels (measurement response). Between 20 to 52 km (50 to 0.5 hPa) the measurement response is higher than 0.8 what corresponds to an a priori contribution less than 20%. Consequently, the retrieved ozone values at these altitudes are predominantly based on the measured line spectrum.

As diagonal elements of the a priori covariance matrix we assume a relative error around 35% at 100 hPa. The error decreases in the lower stratosphere up to 28%. Then it increases linearly from 35% in the upper stratosphere to 70% in the lower mesosphere. The off-diagonal elements exponentially decrease with a correlation length of 3 km.

3. **Comments from the referee:** Section 3: Harmonization strategy for the ozone profiles and Fig.3
a) Fig.3 left panel: the authors show the mean of the profiles obtained from the 2 instruments during the same period. And the blue dashed lines are the error on the FFTS instrument. This error is "the total error from the retrieval along with the error from natural variability of FFTS for this time": is it relevant to include the error "from natural variability" (note that this error is not very clear in this section and only explained in Sect. 5), since both instruments are measuring at the same time and since therefore the natural variability should be the same in both cases? Secondly, is it the total error of an individual measurement or the total error on the mean profile (random errors divided by the square root of the number of measurements) which would be more relevant since this plot compares means of profiles?
b) Fig.3 middle panel: "...presents the mean relative difference profile between data of both spectrometers..." It is not clear if it is the difference between the two profiles shown in the left panel (means of the profiles), or if it is the mean of the individual differences between 2 measurements (collocated in time). Please, clarify. The latter is usually used for assessing bias between instruments. Using the former method could lead to wrong bias assessment (e.g. the difference on the 2 means could come from few outliers in one of the 2 data sets means , while the other individual collocated measurements would be in agreement. When the latter method is used, one can conclude if the bias is significant or not looking at the statistical standard deviation of the individual differences. From the figure (middle panel), it gives the impression that the bias is not significant because of the grey

dashed lines (error on one single FFTS measurements or on the mean profile?) being larger than the difference. If this is true, why correcting for a bias if it is not significant? How looks your Fig. 8 if you do not make the bias correction? I suggest to plot the mean of the relative differences between individual measurements and the error on that mean (using the standard deviation of the statistics) in Fig.3, middle panel (and right panel), and conclude about the significance of the bias between the 2 instruments.

Author's response:

In accordance with the referee wishes, we have changed Fig.3. In Figure 2 we have represented in the left panel the mean ozone profiles recorded by FB (red line) and by FFTS (blue line) for the Oct.2009-Aug.2011 time interval. The dashed lines are the standard deviations of the measurements. In the middle panel we have plotted the mean of the relative differences between individual measurements. The dashed line is the standard deviation of the differences. Finally in the right panel is presented the mean absolute difference profile used to harmonise the GROMOS data set.

Author's changes in the manuscript: In Figure ~~3~~2, we show in the left panel the mean ozone profiles recorded by FB (red line) and by FFTS (blue line), for this time interval. The ~~blue~~ dashed lines are the ~~total error from the retrieval along with the error from the natural variability of FFTS for this time interval. As we are using FFTS data as reference for the FB data and in order to make the figure more straightforward we only show the error profile from FFTS~~standard deviation of the measurements. The middle panel presents the mean relative difference profile between data of both spectrometers with the FFTS data as reference. The ~~grey dashed lines delimit the FFTS error centred at zero~~dashed line is the standard deviation of the differences.

4. **Comments from the referee:** Trend estimation method

a) Choice of the von Clarmann et al. (2010) method: can the authors explain more what is the benefit of using this method compared to the "traditional" one (i.e. without using a full error covariance matrix) which corrects the trend uncertainties for the autocorrelation in the residuals by using a Cochrane-Orcutt transformation to the model (e.g. in Brunner et al., ACP 2006; Chehade et al., ACP 2014; Coldewey-Egbers GRL, 2014; ...) or by simply applying a correction in the uncertainty (e.g. Nair et al., ACP, 2013; ...)? Does the von Clarmann et al. (2010) method impact also the trends themselves or only the uncertainties on the trends?
b) It seems that 7 harmonics is much more than what is usually found in the literature (e.g. from 2 to 4 harmonics are found in the references given above). Does it improve significantly the residuals and the coefficient of determination (R^2) to add harmonics? What was the criteria to choose 7 harmonics?
c) The ENSO signal in Fig. 6 seems very low: is the parameter "f" (in Eq. 1) significant? If not why not remove it from the model?
d) Other proxies (e.g. Northern Atlantic Oscillation; eddy heat flux,...) can also be used in ozone trend studies (Weber et al., ACP, 2011; Frossard et al., ACP, 2013;...). Any reason not to include them? Were they tested and found not significant?
e) The plot of the residuals (Fig.6) rises a question: do you have an explanation

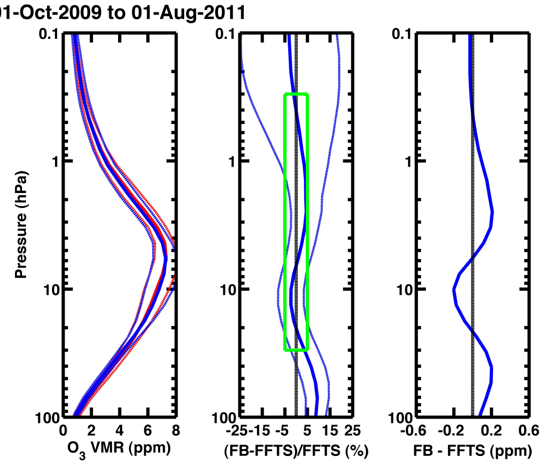

Figure 2: Harmonisation of ozone profiles retrieved from the FB (red line in the left panel) and FFT (blue line in the left panel) spectrometers. The bias between FB and FFTS is less than 5% (middle panel) as derived from the overlap measurement (2009 to 2011) of ozone profiles at pressure levels from 30 to ~~0.5~~ 0.3 hPa (valid altitude range of GROMOS, green box). The blue dashed lines and the grey dashed lines represent the error of FFTS

for the observed oscillation in the residuals?
f) Since the authors provide a trend profile (Fig.8), it could be useful to add the proxies contribution as a function of altitude as well, and to include the information about their significance (as a function of altitude).

Author's response:

a) The advantage is that a scheme which supports the use of the full covariance matrix can solve multiple problems having to do with covariances: (a) model deficiencies which go along with auto-correlated residuals (b) correlated measurement errors (c) biases between data subsets. For particular applications the methods mentioned may be equally well suited but the von Clarmann's approach offers a tool which serves multiple purposes. Indeed the trend itself can change. This is because the algorithm supports weighting of the data points according to their data errors. By adding the covariance term (which assigns the same model error to each data point), the weight of the data points will change. The contrast of the weights will become smaller.

b) Other publications using the von Clarmann *et al.* (ACP, 2010) method also employ 8 harmonics (Eckert *et al.* (ACP, 2014); Stiller *et al.* (ACP, 2012); Kellmann *et al.* (ACP, 2012)). As mentioned in Section 4: Trend estimation method, the selection of the harmonics was done based on the power spectra of GROMOS time series.

c) The ENSO signal, as you can see in Fig.6, is useful to take into account the effect of ENSO 1997.

d) Due to the von Clarmann et al. (2010) method is a novel trend estimation program it has not been tested for all the proxies yet.

e) The shape of the residuals at 10 hPa shows some short term anomalies, maybe due to some regional variation not included in the trend method. In addition, one cannot expect that the series of the climate indices are sufficient to represent the natural variability of the Earth system completely.

f) In Fig.6 is represented the ozone fitted signals of the proxies QBO, solar F10.7 cm flux and ENSO at 10 hPa. Nevertheless, we consider that the study of the contribution of the proxies to the ozone variability is beyond the work of this linear trend paper. Further, it is planned a separate and detailed study of the natural variations of ozone due to these oscillations.

Author's changes in the manuscript: No changes.

5. **Comments from the referee:** Uncertainty considerations

   a) I don't understand why in the second type of error ("observation error"), the authors just use the thermal noise error instead of the full random error budget from OEM (thermal noise + smoothing + maybe other random sources).

   b) The authors consider a systematic error using validations studies: I would guess that a (constant) systematic error should not impact the trends. Instead, I would investigate for a possible drift in your time-series by using the validation results. This would enlarge the trend uncertainties.

   Author's response:

   a) As indicated previously, in the answer of the referee's comment number 2 c), we do not consider systematic error contributions during the trend estimation.

   b) In order to avoid misunderstandings we should rename this systematic error as instrumental error. In this instrumental uncertainty we assume an uncorrelated monthly instrumental error of about 5 to 10% for GROMOS. The reason to enlarge the GROMOS uncertainty by adding this instrumental uncertainty is to be on the safe side during the trend uncertainty estimation. The large number of GROMOS measurements per month (an ozone profile each 30 minutes) allows a robust assessment of the uncertainty from natural variability. But in turn this large number of measurements lead to a small contribution due to the error of the natural variability. From past cross-validations of GROMOS with collocated instruments we know that the different instruments and measurement techniques can differ by 5-10%. Sometimes, these biases may change with month, season or year. This behaviour of the biases might be due to the sophisticated measurement and retrieval techniques.

   Author's changes in the manuscript: Finally, ~~owing the fact that we had not taken into account any kind of systematic uncertainty of the instrument, we have chosen to include as uncertainty source an estimation of systematic error profile~~we assume an uncorrelated monthly instrumental uncertainty. The aim is to take into account the bias between GROMOS and other instruments, and thereby to get a realistic uncertainty estimation.

6. **Comments from the referee:** Results and discussion

    Any explanation for the negative trend in the lower mesosphere? You checked your results with other experimental studies: are there available modelling studies that could help you to explain this trend?

    Author's response:

    In Section 6: Result and discussion it is mentioned that this lower mesospheric negative ozone trend was also reported in recent trend estimations (Tummon *et al.* (ACP, 2015); Remsberg *et al.* (ACP, 2014); Kyrölä *et al.* (ACP, 2013)) but what would be the origin of this effect would have to be investigated in a separate project.

    Author's changes in the manuscript: No changes.

Minor or technical comments

1. **Comments from the referee:** p.16372, l.4 (abstract): Provide longitude and latitude of the instrument. Also in p. 16374, l.13.

    Author's response:

    No comments.

    Author's changes in the manuscript: The ozone radiometer GROMOS (GROund-based Millimeterwave Ozone Spectrometer) performs continuous observations of stratospheric ozone profiles since 1994 above Bern, Switzerland (46.95°N, 7.44°E, 577 m).

    Among other advantageous technical features, the 20 years of continuous observations and the privileged location of the instrument offer us a pretty clear vision of the distribution of ozone in the northern mid-latitudes (46.95°N, 7.44°E, 577 m).

2. **Comments from the referee:** p.16372, l.20-22 (abstract): Provide the corresponding approximate altitudes for the trends at 4.36hPa and 02.hPa.

    Author's response:

    No comments.

    Author's changes in the manuscript: With our observed ozone profiles, we are able to support this statement by reporting a statistically significant trend of +3.14 % decade^{-1} at 4.36 hPa (37.76 km), covering the period from January 1997 to January 2015, above Bern. Additionally, we have estimated a negative trend over this period of -3.94 % decade^{-1} at 0.2 hPa (59 km).

3. **Comments from the referee:** p.16373, l.4-6: Not clear to me. "...ozone losses were 15-18

    Author's response:

    We delete this sentence.

    Author's changes in the manuscript: ~~Early predictions of ozone losses were 15 - 18% if CFCs added 5.5 - 7.0 ppbv of chlorine to the stratosphere at 1975 rates (Hudson and Reed, 1979). Thereafter, these were confirmed in many publications,~~

~~such as in the~~ In the last Scientific Assessment of Ozone Depletion: 2014 of the World Meteorological Organisation (WMO, 2014) ~~, where it~~ is stated that global ozone levels decreased through the 1980s and early 1990s while stratospheric abundances of ozone depleting substances (ODS) were increasing.

4. **Comments from the referee:** p.16375, l. 5: "And" instead of "An".

   Author's response:
   No comments.

   Author's changes in the manuscript: ~~An~~ And finally, Section 7 is a summary of our findings.

5. **Comments from the referee:** p. 16381, l. 8: the "." is missing at the end.

   Author's response:
   No comments.

   Author's changes in the manuscript: ~~Most of the ozone variations can be explained by the fitted proxies~~ , maybe due to some short term anomalies.

6. **Comments from the referee:** p.16381, l.12: "...we have considered 3 different ways to assess the uncertainties": this gives the impression that you will then choose between the 3, while at the end you sum them. So change to something like "We have considered 3 types of errors".

   Author's response:
   No comments.

   Author's changes in the manuscript: We have considered three ~~different ways to assess the~~ types of uncertainties.

7. **Comments from the referee:** p.16395, legend Fig3.: "...30 to 0.5 hPa": should be 0.3hPa, according to the text and the green box.

   Author's response:
   No comments.

   Author's changes in the manuscript: Caption Figure 3: Harmonisation of ozone profiles retrieved from the FB (red line in the left panel) and FFT (blue line in the left panel) spectrometers. The bias between FB and FFTS is less than 5% (middle panel) as derived from the overlap measurement (2009 to 2011) of ozone profiles at pressure levels from 30 to ~~0.5~~ 0.3 hPa (valid altitude range of GROMOS, green box). The blue dashed lines and the grey dashed lines represent the error of FFTS