Atmos. Chem. Phys. Discuss., 12, C13640–C13656, 2013
www.atmos-chem-phys-discuss.net/12/C13640/2013/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Comparison of three vertically resolved ozone data bases: climatology, trends and radiative forcings" *by* B. Hassler et al.

B. Hassler et al.

birgit.hassler@noaa.gov

Received and published: 28 March 2013

We would like to thank both reviewers for their helpful comments and suggestions to improve our manuscript. Our responses to the reviewers' comments are listed below. Original comments are printed in black, answers are printed in blue.

Anonymous Referee #1

The paper presents three global ozone data sets available for global climate models as input data or for validation purposes. In general the paper is well written and the results are presented in a clear fashion. A thorough comparison of the three data sets will provide a very useful tool for the modeling community. Such a comparison should if possible compare to independent measurement data sets, compare exactly

the same quantities and attribute analyzed differences to the different ways of compiling the data sets. I feel that the manuscript needs to be improved in all three points before publication in ACP.

We thank the reviewer for the detailed suggestions on how to improve our manuscript. We tried to follow the suggestions closely to make the manuscript a useful tool for the modeling community. Concrete changes are described in the answers to the general and specific comments, but here is a summary of our changes regarding the raised issues:

1.) Compare to independent measurement data sets. We used data from TOMS/SBUV, and SBUV/2 in the manuscript for comparisons with the three data sets. Following the advice of the reviewer we also added data from UARS MLS and Aura MLS.

2.) Compare exactly the same quantities. We agree that the climatology comparison of the integrated ozone of the data sets with TOMS/SBUV was not sufficient. We therefore removed this comparison from the manuscript and added a comparison of climatologies (data sets and SBUV/2 data) at 10 hPa. Additionally, we also added a climatology from BDBP that was calculated for the same period as is covered by the Fortuin & Kelder climatology to examine the effects of the length of the time period on the climatological values.

3.) Attribute the analyzed differences to the different ways of compiling the data sets. We agree that the attribution of differences in the data sets to the different ways of compiling them would help greatly to better understand the data set characteristics. Unfortunately, this is not possible to do in an unambiguous manner, given the information that is available to us about the RW07 and SPARC data sets, their input data and exact used methodology. The data sets differ in three different ways: (i) in their choice of input data, (ii) in their number of used basis functions, and (iii) in their applied regression method. Changing one of these three things to make it consistent between two data sets (for example the number and choice of basis functions) will not help to

understand all the remaining differences between the two data sets, but might only help understand at least a few. We followed the reviewer's advice and looked at the variability of the BDBP data set if less basis functions are used. However, since BDBP and, for example, RW07 differ also in their choice of input data and the applied regression method, the remaining differences in the variability of the data sets (after the reduction of the BDBP basis functions) could not unambiguously attributed to either one of these things.

General comments

The evaluation of integrated ozone shows clear differences between the data sets. However, since not the same quantities are compared it is not possible to deduce the origin of these differences. Among other things they could result from the differences between ozone integrated from 250 to 1 hPa and stratospheric ozone or from the different time periods on which the climatologies are based. The authors should aim at comparing the same quantities between the three data sets (e.g., ozone integrated between the tropopause and 1 hPa). Additionally the comparison would be more meaningful if it could be based on an independent observational data set (TOMS/SBUV) with the tropospheric ozone column removed. If this is not possible it should be shortly discussed in the text.

We agree that the comparison of the integrated ozone climatologies of the three data sets has some flaws. Since the RW07 data set does not include tropospheric ozone values, it is necessary to either remove the tropospheric values from the SPARC and BDBP data set, and the independent observations (e.g. TOMS/SBUV), or tropospheric values have to be added to the RW07 data set. In either case, assumptions about the tropospheric ozone distribution (in space and time) have to be made that might introduce biases to one or more of the data sets. We therefore decided to remove the comparison of the integrated ozone climatologies (Figure 1 and Section 3.1), and rather add a comparison of the ozone at a specific level (new Figure 1 and new Section 3.1).

Finally the impact of using different time periods for the climatological background value should be explored more in detail and possibly presented in a Figure by analyzing BDBP based on a time period comparable to the FK98 climatology (which has already been done as stated in section 3.2). If the time period length has a substantial impact on the absolute values this would be important information about the RW07 and SPARC climatologies.

We had not added more detail about the influence of the different time periods for the climatological background values since we did not want to add length to the already long manuscript. However, we agree with the reviewer that it is important to see the differences in the data sets that are caused by the use of different time periods for the climatologies. As suggested, we therefore added a BDBP climatology graph in the new Figure 1 that is based on the shorter FK98 time period, and some more discussion about the differences that occur between the BDBP climatologies based on the short and long time series.

The first sentence in section 3.2 (vertically resolved ozone) is very confusing. In which sense are the patterns similar? Is this comment referring to the global distribution or to the differences to FK98 discussed in the next sentence?

We agree that the first sentence of Section 3.2 was confusing, so we have rephrased it.

The whole section reads like a comparison to FK98 would be included but only three data sets are presented in Figure 2.

We agree that the section read like the FK98 climatology was part of the comparison. We rephrased that section.

To me it seems that differences change sign at 7 hPa (instead of 10 hPa) and that the peak of RW07 is at 5 hPa (instead of 7 hPa). BDBP has high values mostly between 30 and 7 hPa. It seems surprising that RW07 has its maximum at 5 instead of 10 hPa.

Can this be explained somehow?

We agree with the reviewer that the description of the climatological profiles shown in Figure 2 might have been confusing. We therefore rephrased parts of the second paragraph of Section 3.2 to clarify our description.

We do not know what causes the maximum of the RW07 climatological profile in the tropics to be near 5 hPa, clearly higher than BDBP and SPARC. In the manuscript we suggested that this could be caused by the different measurements on which the three data sets are based, or the different time periods on which the climatologies of the three data sets are based (BDBP: 1979-2005, RW07 and SPARC: 1980-1991).

One more possible cause could be the fact that in order to compare the three data sets, we had to convert the RW07 data from altitude onto pressure levels, and from "DU/km" to volume mixing ratio. We tried to convert RW07 in the best possible way, namely close to the conversion Bill Randel and Fei Wu used to be able to combine the FK98 climatology with the anomalies calculated with their regression model. But lacking the exact information on how their conversion had been done, we did some sensitivity tests to find out how big the differences in ozone can be, depending on the conversion used. We used a fixed scale height of 7km, a standard atmosphere, and a zonal mean, monthly resolved temperature climatology. The ozone differences in the tropics can be as high as 20% (for pressures lower than about 20 hPa). In the polar regions the differences can be as high as $\sim$30% at pressures around 3-4 hPa (in SH spring or fall), and as high as $\sim$60% at pressures around 4-5 hPa (in SH winter). Based on our comparison, the profiles obtained by a conversion with a fixed scale height were clearly lower than the profiles obtained by conversions with temperature profiles and the profiles from BDBP and SPARC. Therefore, we decided that a conversion using a fixed scale height would not be accurate enough, especially higher up in the atmosphere, despite the fact that Bill Randel provided the information that they used a fixed scale height of 7 km in combination with a zonal mean temperature climatology for their original conversion. Results for the RW07 conversion with a standard atmosphere

and a zonal mean temperature climatology are similar. Therefore, decided that the profiles shown in the original manuscript (converted with a standard atmosphere) can be used for the comparison.

Independent of the applied conversion, the peak in the tropical ozone profile was still clearly at lower pressures (around 5-6 hPa) in RW07 than in BDBP and SPARC. The conversions therefore might have a small effect on the placement of the maximum within the profile, but we think it cannot explain the bulk of the difference that can be seen between RW07, SPARC and BDBP. The exact reason why the RW07 maximum is located higher in the profile is not clear to us.

We added a paragraph with a discussion about different conversion for RW07 and their related difficulties to the RW07 data description section (Section 2.1), another small discussion about this topic to the discussion section (Section 8), and an explanatory figure to the Supplementary Material.

The discussion of the anomalies could be strengthened by referring to what the BDBP anomaly time series looks like if one (or more) basis function (e.g., volcanoes, ENSO) is omitted from the regression model. This way one could distinguish between the impact of the different measurements used as input for the three data sets and the impact of the various basis functions. It is hard to follow the discussion in 4.2 without a figure.

The inclusion or omission of one or more basis functions to a regression model does have a clear impact on the amount of variability that the resulting time series will show. For example, if the volcano basis function is omitted from the BDBP regression model, the anomalies in the years following the eruption of Mt. Pinatubo are clearly more positive, especially in the tropics, the mid-latitudes, and the SH polar regions in spring, than when the volcano basis function is included in the regression model. Therefore generally removing basis functions from the BDBP regression reduces the variability in the BDBP data set, and will result in variability that is more similar to RW07.

However, the three data sets do not only differ in the measurements used as input for the regression, and the number of basis functions used, but also in the regression method itself. While we assume that for RW07 and SPARC the ozone measurements of each latitude band and pressure/altitude level were fitted separately (this would be the traditional way to apply a regression model, and there are no additional specifications given in the respective publications where the creation of RW07 and SPARC is discussed), the regression model for BDBP uses all measurements of one pressure level simultaneously for a fit and therefore fits all latitude bands at the same time. The omission of the volcanoes and ENSO basis function from the BDBP regression model does therefore not necessarily help in shedding light as to where the differences in variability between RW07 and BDBP originate from, the choice of basis functions or choice of input measurements. If differences still exist, they could be due to the different input measurements or the differences in the regression method.

We did not add further discussion about this to the manuscript (in addition to the brief description of the creation of the three data sets) since we thought it would not help with the discussion about the variability in the different data sets, but rather cause more confusion about the origin of the differences in variability. However, we rephrased and reworded the discussion in Section 4.2 to make it easier to understand it without a figure.

The comparison of the time series seems to be mostly based on observational data used for the compilation of the BDBP data base. Since BDBP includes a large part of the vertically resolved ozone measurements this is only a natural consequence, however, at least for later parts of the time period additional independent data sets are available.

It is true that all shown data sources in the comparison of the time series are included in the compilation of the BDBP data set, except for SBUV. However, since RW07 and SPARC are based on SAGE I/II and ozonesonde measurements (from Resolute and Syowa), all the data that was used to compile these data sets are also included in the

comparison of the time series. However, we appreciate the reviewer's point and have added monthly means for the different pressure levels from UARS MLS and AURA MLS, where available, as additional independent data source.

Why is the raw BDBP data shown? It is not discussed in the text and somehow is a bit misleading when looking at the plots.

We took the raw BDBP data out of the plots, and updated all respective figures.

Specific comments:

Page 26564, Line 24: Give full name for SAGE here (instead of doing so on the following page).

We changed the manuscript as suggested.

Page 26565, Line 21: SAGE data where, in the polar regions or globally?

SAGE II data and ozonesonde data in the polar regions are only fitted with an EESC basis function. We clarified this in the manuscript.

Page 26565, Line 22: What about data between 55° and 65°?

Trends derived from ozonesonde data for the latitudes pole ward of 65°, and trends derived from SAGE data equator ward of 55° are merged between 65° and 55° by interpolation. We changed the manuscript to clarify this.

Page 26565, Line 22: So this means that the regression model is only applied to anomalies? Are those deseasonalized anomalies?

Yes, for the creation of the RW07 data set, anomalies were calculated and then a multiple linear regression model was applied to the anomalies. The FK98 climatology was then added to the regression model results to create the global, gap-free data set. The anomalies were calculated by deseasonalizing the different data sets ("Each data set is first deseasonalized using a harmonic regression of monthly binned data",

Randel and Wu, 2007). We clarified this in the manuscript.

Page 26568, Line 3-4: TOMS and SBUV have been defined before.

We removed the definition of TOMS and SBUV at this point of the manuscript.

Page 26574, Line 6: The measurement systems are not independent from the three data sets.

We agree that the wording of this sentence was not accurate enough. We changed the manuscript accordingly.

Page 26574, Line 23: Why BDBP Tier 4? What does it stand for and is it different from the version used earlier in the paper?

The BDBP data set comes in four different varieties, based on the combination of basis functions used to create the dataset (Bodeker et al., 2013). BDBP Tier 1.4 stands for the data set that is used in this manuscript for the comparison with RW07 and SPARC, and that was described in Section 2.3. We changed the description of the data set in Section 2.3 to avoid confusion about the "Tier 1.4" nomenclature.

Page 26575, Line 1: The term "higher pressure levels" used here could be interpreted as a pressure level which is higher up in the atmosphere or as a level of higher pressure. Note that the term "higher pressure" was used before in the text to indicate regions lower in the atmosphere.

The term "higher pressure levels" is used here, as it was earlier in the manuscript, to indicate regions lower in the atmosphere. We rephrased the sentence to "levels lower in the atmosphere" to clarify the intended meaning.

Page 26580, Line 5: The differences are not small in my opinion.

We adjusted the wording in the manuscript to ". . .although some differences exist."

Page 26585, Line 12-14: shouldn't this be the other way around (i.e., RW07 larger at

higher altitudes . . .)?

We agree that the description on page 26585, Line 12-14, was not specific enough, and could therefore easily be confusing. We rephrased the sentence in the manuscript.

Page 26585, Line 25-27: this statement would benefit from improved comparisons in section 3 (see general comments).

The sentence on page 26585, Line 25-27, does not refer to comparisons of the climatologies (Section 3), but it is a statement concerning the time series comparisons shown in Figure 4 to Figure 6, and the Figures in the Supplementary Material. We think the statement therefore does not have to be changed and we kept it as it was.

Anonymous Referee #2

This manuscript compares a new ozone climatology from Bodeker et al. (2012) to two somewhat older ozone climatologies by Randel and Wu (2007) and Cionni et al. (2011), which have been hitherto used to evaluate past ozone trends and prescribe ozone fields in global transport and tropospheric chemistry-climate models (used in the IPCC AR5) without full representation of stratospheric chemistry, respectively. Given that changes in stratospheric ozone are known to strongly impact surface climate, representation of a realistic ozone field in these models is crucial to understanding past climate change and variability. However, the reality of the available ozone climatologies is not yet known. In order to address this issue, the study presented uses multiple diagnostics such as climatology and variability of both total column and vertically resolved ozone in order to evaluate differences in the three available ozone data sets, derives long-term ozone trends and assesses the impact of the differences among those trends on radiative forcing. It is important to keep in mind for the interpretation of the results that the three regression-based climatologies were constructed using different sets of basis functions in order to serve different needs of the community and hence cannot be expected to reproduce all aspects of variability. The presented detailed comparison of the three available data sets and the guidance provided for their use constitutes a

very valuable contribution to ozone research and the study hence fits well within the scope of the ACP journal. The manuscript is generally well written. However, I got the impression that some evaluations have not been done as thoroughly as they should have been done. I hence recommend minor revisions before publication addressing the comments as outlined below.

Major comments

(1) Evaluation 3.1 'Integrated ozone': This evaluation and discussion thereof need to be revisited. While the authors may be right with their conclusions, their evaluations are not accurate enough to convince the reader. You cannot just add some rather arbitrary tropospheric ozone column value that is strictly valid only for the period between 2004-2010, since there are tropospheric trends between 1979 and 2010 that may make significant differences to your main conclusion (claim) that RW07 total ozone column is biased high. A first simple remedy is to compare total columns for all data sets in a consistent way, namely integrated from the monthly mean tropopause height up to 1 hPa, instead of simply integrating between 250 and 1 hPa. Secondly, as shown in your figure 3, BDBP shows too high values in the early period and too low values in the later period, so that the average of these values may just by pure coincidence agree better with TOMS. Using shorter time periods of averaging (e.g. 1979-1985, or 2000-2005) would avoid this issue and a 2000-2005 based climatology may be better comparable to a combined TOMS/Ziemke et al. (2011) climatology.

We agree that the comparison of the integrated ozone climatologies of the three data sets has some flaws. As already mentioned in a response to reviewer 1, since the RW07 data set does not include tropospheric ozone values, it is necessary to either remove the tropospheric values from the SPARC and BDBP data set, and the independent observations (e.g. TOMS/SBUV), or tropospheric values have to be added to the RW07 data set. In either case, assumptions about the tropospheric ozone distribution (in space and time) have to be made that might introduce biases to one or more of the data sets. We therefore decided to remove the comparison of the integrated ozone

climatologies (Figure 1 and Section 3.1), and rather add a comparison of the ozone at a specific level (new Figure 1 and new Section 3.1).

(2) Evaluation 3.2 'Vertically resolved ozone': The first paragraph should be based on a figure the reader has access to. Second paragraph onwards: your discussion on Figure 2 seems to talk about 4 data sets (the three data sets plus FK98), but I can only see three! Please revise figures or explain on which basis you discuss, but it cannot remain as it is.

We agree that the first and the second paragraph of Section 3.2 were confusing. We rephrased the first paragraph and removed inconsistencies between the description of Figure 2 and the actual Figure 2 from the manuscript.

(3) The study emphasizes most strongly the differences in the polar regions, where we know that RW07 has difficulties due to the limited data coverage of SAGE. I like to see similar evaluations as in Figs. 2 or 4/5/6 also for midlatitude regions to learn more about the validity of the different data sets in these regions.

Examples for mid-latitude ozone time series are included in the supplementary material, but since readers might not look at this material in detail, we agree that it would be helpful to show examples of the mid-latitudes in the manuscript. We therefore added two panels to Figure 2 which show climatological profiles of all three data sets in the Northern and Southern Hemisphere mid-latitudes, and one more subsection to the manuscript (Section 5.2) where time series of Northern mid-latitudes are compared. We also added more references to the supplementary material in the manuscript.

(4) I suggest moving presentation and discussion of Figs. 9 and 10 into the Result section. The problem I have with showing them in the Conclusion section is that they are using the Hassler et al. (2008) gridded data climatologies for comparisons, which are also main input into the BDBP data set and hence can be expected to agree better with BDBP than with the other data sets. It would also allow more space for a more balanced discussion of the strength and weaknesses of the different data sets in the

conclusion section, e.g., that the BDBP shows too strong Arctic trends and some artifacts in the SH upper stratosphere. Also, you may want to explain again why RW07 has problems in the polar regions, and that the SPARC data sets basically only provides information on the chlorine-related part of ozone changes.

We would like to keep Figs. 9 and 10 in the Discussion and Summary Section (Section 8). We put the figures there present to the reader summary-like graphs that quantify the comparison of the three data sets with measurements systems.

For this comparison we did not use, however, gridded data from SAGE II and ozonesondes. The database described in Hassler et al. (2008) only assembles individual measurements of different systems on a common vertical grid. Spatial and temporal characteristics of the measurements were kept as they originally were. RW07, SPARC and BDBP values were used where they were interpolated on common pressure levels (the levels on which SPARC is provided since it has the coarsest resolution). Individual SAGE II and ozonesonde measurements were then combined in 5° latitude bands, and interpolated onto the pressure levels on which the data sets were provided, and the difference between the data sets and the measurements calculated. The main reason we used the SAGE II and ozonesonde data from the BDBP database (from which the zonal mean, vertically resolved BDBP data set discussed and compared in this manuscript is derived) is the already-performed quality control and screening of the SAGE II and ozonesonde data. The quality screening of the data should only have a small effect on the comparisons.

However, we understand that the description of the data source (in this case the BDBP database) might be confusing. We therefore changed the manuscript to make the description clearer. We also added some more discussion to the discussion section about the strength and weaknesses of the different data sets as suggested.

(5) Evaluation 'Annual mean trends': P26581 L 11-13 you say that SAGE I could be the problem for the large trends seen in BDBP. However, RW07 includes these data as

well and do not get a similar result? Also, given the rather too high/low anomalies at the beginning/end of the BDBP data set, I find it a strange argument on P26581 L19 to say that the trends derived by Randel and Thompson (2011) and Forster et al. (2007) cannot really be used to confirm the results, i.e. a better agreement with the trends of RW07 and SPARC and disagreement with BDBP due to the different time periods over which the trends are calculated. You have most of the data needed, so could adopt the periods over which the trends are calculated as used in these studies for comparison to confirm your argument.

The reviewer is correct that RW07 also includes the SAGE I data in their data set. However, their method of including the data is different to the method used for the creation of the BDBP. According to Randel Wu, JGR, 2007, they "use a regression analysis to generate a continuous interannual anomaly data set from the combined SAGE I and II and polar ozonesonde data sets. Each data set is first deseasonalized using a harmonic regression of monthly binned data, . . .", whereas the approach for the BDBP was to not deseasonalize the data before, but use the absolute values to describe the climatology and the variabilities with the regression model. By deseasonalizing the data sets separately, the climatologies of the different data sets will look different (because they are based on a different time period), which can then have an effect on the trend.

We recalculated the trends for all three data sets, as suggested, for the period 1984 to 2005, as it was done by Forster et al. (2007), where only a linear trend and no EESC basis function was used to describe the chemically caused long-term ozone changes. The trends derived from this analysis are weaker than the trends shown in the manuscript: in the tropical lower stratosphere, where a comparison with the results of Forster et al. (2007) is possible, they are about -3% per decade for RW07, about -2% per decade for SPARC, and about -7.5% per decade for BDBP. This brings the BDBP trends close to trends reported in Forster et al. (2007), whereas trends for both RW07 and SPARC are clearly weaker. We added a few sentences about this additional

analysis to the manuscript. Trends described by Randel and Thompson (2011) could not be reproduced since RW07 and BDBP do not cover the whole time period that was analyzed in this study (RW07: 1979-2205; BDBP: 1979-2006; Randel Thompson: 1984-2009).

Minor comments

Abstract, L22: I do not agree that the uncertainty in our knowledge of ozone trends is large, since there are multiple other ways used to derive those trends (see latest ozone assessment WMO, 2011) than just using the here discussed climatologies. The problem to point out lies rather in how the different ozone climatologies are constructed, i.e., which particular basis functions are used for the respective regression models and how these influence what trends the climatologies pick up.

We agree with the reviewer that the uncertainty in our knowledge about ozone trends is not as large as it might have been suggested by reading the abstract. We rephrased the sentence to specify that the uncertainties in the data set trends are large, and changed the sentence to "However, the differences among the three suggest that there are large uncertainties in their respective ozone trends".

P26563 L15-18: Please rephrase, it seems an awkward sentence starting with 'Although...' Also, see previous comment, you say yourself that one should not use the climatologies to derive trends!?

We assume that the reviewer meant P26564, L15-18. We kept the "Although" at the beginning of the sentence, however, we changed the sentence structure later in the sentence, so that the meaning is clearer. We think that with the implemented changes, it becomes clear that we do not perform the trend analyses with the three data sets to derive better knowledge about past ozone trends, but to get an estimation of how well the data sets capture the trends that are derived from observations.

P26574 L24: Why do you say BDBP Tier 1.4? Is this a different data set all of a sudden

then used before?

No, the data sets described in Section 2.3 and in the sentence on P26574 are the same. We changed the description of the data set in Section 2.3 to remove the confusing term "Tier 1.4".

Technical comments

- Introduction P26563 L2: Did you want to say 'in a changing climate'?

We did want to say "in a changing climate" and changed the manuscript as suggested.

- P26564 L1: Please rephrase to 'that (i) show high vertical resolution, (ii) cover...'

We changed the manuscript as suggested.

- P26567 L16: Suggest to delete 'a manuscript by'

We deleted the suggested phrase.

- P26568 L14: Use the abbreviation you defined earlier for this publication 'FK98'.

We changed the manuscript and use now the abbreviation defined earlier in the manuscript.

- P26572 L3: Please refrain from using 'significantly' unless you mean statistical significance.

We changed the manuscript from "variability of the anomalies is significant" to "variability of the anomalies is considerable"

- P26580 L17: Suggest using 'higher into the middle stratosphere' instead of 'atmosphere'

We changed the manuscript as suggested.

- P26582 L26: What are 'purely statistical uncertainties'?

We agree that the term we used might be confusing. We therefore changed the manuscript to ”statistical uncertainties”.

- Figure 7: It is confusing to use different contour levels (and colors?) in these plots. I also find it somewhat confusing that the y-axis for the plots on the right is on the right instead of on the left. You don’t gain space or comparability by doing so.

We changed the contour levels of the plots in the right column to match the contour levels of the plots in the left column. We also show the y-axis of the plots in the right column on the left side now, instead of the right side.

- Table 2: Please add latitude band of the different regions considered.

We added the definition of the regions in Table 2 to the table caption.

- Table 1: Please add highest pressure level also in km for the SPARC data set for consistency.

We agree that the presentation of the highest pressure level was not consistent. Since the comparison of the databases is based on pressure levels, we removed the altitude information (highest level) for RW07 and BDBP from Table 1.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 26561, 2012.