Atmospheric
Measurement
Techniques
Discussions

*Interactive comment on* "Comparison of GTO-ECV and Adjusted-MERRA total ozone columns from the last two decades and assessment of interannual variability" *by* Melanie Coldewey-Egbers et al.

**Anonymous Referee #4**

Received and published: 27 November 2019

General comments

The authors present the newest version of the GTO-ECV dataset, containing long-term harmonized total-column ozone measurements from the GOME, SCIAMACHY, OMI, and GOME-2A satellites, and compare it with the Adjusted-MERRA dataset which assimilates measurements from the SBUV and SBUV-2 satellite instruments, as well as MLS, IASI, CRIS, ATMS, as well as OMI.

First, a comparison of the two datasets in terms of zonal mean ozone and its trends

and seasonal cycles is presented. Then the two datasets are analyzed and compared seasonally at 5 x 5 degree spatial resolution. Similarly, the ozone anomalies for selected 5x5 grid cells are shown and a comparison of the two datasets is made. The inter-annual variability of the anomalies is calculated as a standard deviation, and once again the two datasets are compared.

Finally, the authors perform an empirical orthogonal function (EOF) analysis on the total column ozone anomalies, within 25 degrees of the equator. The first four EOFs are found to explain 92% of the variance, and these are related to climatic indicators: quasi-biennial oscillation (QBO), solar flux, and the multivariate ENSO index (MEI).

This analysis and presentation of large public datasets is important, since users of the data will probably not do this analysis themselves before using it in their specific application. Thus, users will not be aware of potential problems which hide in the details. However, the present paper does not draw any new scientific conclusions, despite the huge effort in gathering, processing and analysis of the data.

The article is generally well written and the figures are of a suitable quality, if often quite small on the page (however, see the section on technical corrections).

Specific comments

When initially looking at Figure 1, and investigating the genealogy of the two datasets that are being compared, one's initial conclusion is that the clear change in October 2004 is because of the OMI data that is being assimilated into both of the datasets.

Therefore, it is easy to conclude that after this date, the comparison is simply OMI-to-OMI. The adjustment of the MERRA using SBUV is described in one sentence in page 5, and this subtlety is easy to miss. The start to section 3.1 on page 6, which is not very clear, further adds to the confusion.

It is almost impossible to make sense of the present paper and the data presented without a thorough reading of Coldewey-Egbers et al (2015) and Garane et al (2018), with

particular attention to the subtle changes that have been made. A genealogy/timeline of the evolution of the datasets will assist readers to understand what they are working with.

Technical corrections

Pg 1

line 11: "second period", "later period": without the context of reading further into the paper, it's not clear what this means. The abstract needs to stand alone.

Pg 3

line 33: why is the analysis limited to the low and middle latitudes? Surely total-column ozone is important at high latitudes? In your 2015 AMT paper (which I call CE2015), (pg 3924 second column, second paragraph) you state that GODFIT is robust at high SZA.

Pg 4

Line 1: "seperate" -> "separate"

Line 4: "remarkable long-term stability": where is this remarkable stability demonstrated? Perhaps you should point out here that this is different to CE2015, where GOME is used as the long-term reference. Also, in the 2015 edition of your dataset, you use a "soft-calibration" procedure, which has since been discontinued. Indeed a genealogy of the various data products (GTO-ECV and SBUV/OMI-MERRA-derived) and their versions, and the changes from version to version, would be helpful to make sense of them all.

Line 22: the url takes me to news about the ocean-colour dataset, rather than ozone.

Pg 5

Line 17: A glance at your Figure 1 shows that the discontinuities have not been re-

moved (or perhaps they exist in the GTO-ECV dataset?). Here you speak or "renormalizing", while lower down in the page it is referred to as simply normalizing. Use consistent naming.

Pg 6

line 4: "which is completely independent of GTO-ECV": at first reading this might be taken to mean that Adjusted-MERRA is independent of GTO-ECV. Perhaps "which is itself completely..." is clearer. This paragraph as it is currently written adds to the confusion regarding the different datasets and their history.

line 17: "difference in zonal mean total ozone column is ..." here you quote a single number while talking about the zonal means. This is shown later on in the paper (i.e. Fig 5) however here it is somewhat surprising. Does this refer to the global mean?

line 27: "high latitudes and before 2002: that is probably caused by sparse data coverage..." this is in stark contrast to your detailed explanation in CE2015 where you show that the poles receive better spatial coverage.

Pg 7

Fig 1: When you say A-MERRA vs GTO-ECV, what does this mean? There is no agreed-upon meaning for 'vs' in this context. In C-E2015 you spell it out, e.g. (AM - GE) / AM. Later on in the present paper (pg 13) it becomes clear that this "vs" is not calculated how one might assume it is. Fig 1: In Garane et al 2018, their Fig 9, the lifespans of each satellite are shown as horizontal bars on a figure similar to Fig 1. It be helpful to show these, for both GTO-ECV and A-MERRA. For example, one might unkindly split the GTO-ECV into the Gome-SCIA-ECV and the OMI-ECV, such is the heavy influence of OMI measurements on the dataset; however, one might for example ponder if there has been a change since the launch of GOME-2. Might this also give a clue as to the subtle change towards the end of the time-series? Fig 1: The upper figure shows differences as a percentage, the lower figure shows differences in stdev

as Dobson units. Is this intentional?

Line 10: "differences are found the north": insert "in"

Pg 8

Table 2: Here you describe a global dataset. However, you describe DJF as being "Winter", which is only true in the northern hemisphere. Similarly for the other seasons. This is repeated several times throughout the figures and text. Line 2: "5^{\circ}latitude band separately": space missing

Line 11: "introduction of OMI data into GTO-ECV data record..." OMI data is also introduced into the adjusted-MERRA dataset. This omission is made several times in the text.

Line 17 & 18: "trough" is spelled "through"

Page 9

Lines 7 & 8: the seasons for a global dataset are described in terms of northern hemisphere seasons.

Page 10:

Line 1: "Atlantic Ocean, in particular in autumn" is this the North Atlantic, or the South Atlantic? Is this the boreal autumn or the austral autumn? This is extremely confusing.

Line 7 & 8: "southern hemisphere minimum ozone columns in autumn" is this the boreal or austral autumn? This is particularly confusing after reading the beginning of the paragraph, and looking at the figure.

Page 11:

Figure 4: This figure describes the seasons in terms of the northern hemisphere, for a global dataset. Also, the small title above each global map indicates the northern hemisphere seasons. Perhaps you could put the southern hemisphere season below

the map? This might make the figure too busy: you decide. This figure is very small. It could easily be split across 2 full pages.

Page 12:

Figure 5: This figure is very small. The bars are too close together in these histograms, e.g. in the left-hand pair of figures, I can't see if the blue or the orange is taller for a given total ozone amount. Are adjacent orange and blue bars meant to be for the same interval or for consecutive intervals?

Line 11: Northern-hemisphere seasons are described for a global dataset

Page 13:

Figure 6: northern hemisphere seasons in the figure and caption

Line 1: "i.e. Adj-MERRA standard deviations are higher..." if you simply give the formula for what you mean by "vs" on your graph titles (such as in CE2015), then this sort of clarification is not necessary. Indeed, this clarification makes me go back and question how I have interpreted all of your figures, since this is the opposite of my intuition.

Line 10: "corresponding seasonal cycle": is this the seasonal cycle presented in Figure 3?

Line 11: on what basis do you select your seven grid cells, or rather, the longitude at which you have selected them?

Page 14:

Line 3: "variability is dominated by the QBO". While there is clearly a biennial cycle in the data, the QBO itself is a climatological phenomenon, and you present no mechanism or evidence linking ozone column anomalies to the QBO.

Line 10: "coffcient" spelling

Page 15:

Line 4: "ozone anomalies" are these the same as the ones presented in Fig 8?

Page 16

Fig 9: does [-] denote the units?

Line 5: "variability maximizes ..." this is not idiomatic English. "variability reaches a maximum..." would be better.

Line 7: "also linked to wave activity." Have you demonstrated this connection somewhere?

Line 7: "Fig 10 indicates an excellent agreement..." by using an eyeball to examine the differences on a very small plot perhaps. A plot showing the difference between (i.e. GTO-ECV - A-MERRA) the two datasets would show the agreement more clearly.

Line 11: "lower correlations between GTO-ECV and A-MERRA ozone anomalies." Are these the correlations shown in Figure 9?

Line 13: "According to Press et al..." perhaps "Following Press et al..." would be better.

Line 21: "To a large extend..." -> "extent"

Page 18

Title and Line 1: Is it PCA or EOF? Please use consistent naming.

Line 7: "to a lesser extend..." -> "extent"

Line 8: Perhaps "The EOF analysis is performed on the detrended and deseasonalized 5x5 monthly mean ozone columns presented earlier" is better?

Line 9: You give a reference for the EOF analysis in line 2, do you have one describing the Savitzky-Golay filter? Why did you choose 13 months?

Page 19

Fig 12: The units in on the colour-scale in the first column of plots don't match the

C7

second column.

Page 20

Line 15: "extend" -> "extent"

LIne 20: "also a positive correlation..." delete "also"

Page 21

Line 8: "included in GTO-ECV." and also in A-MERRA.

LIne 13: "seaonsal"

Line 21-22: "more than 97% of the grid cells..." mention here the size of the grid cells.