

## ***Interactive comment on “A vertically resolved, global, gap-free ozone database for assessing or constraining global climate model simulations” by G. E. Bodeker et al.***

**G. E. Bodeker et al.**

greg@bodekerscientific.com

Received and published: 3 January 2013

In the text below the reviewer comments are in blue and our responses are in black.

Response to Anonymous Referee 1

We would like to begin by thanking the reviewer for taking the time to review our paper.

I found this paper well written and clear with a definite application/target audience for the end product of the analysis. I have a few minor comments below. In general, I do wonder sometimes if the expectation that things with annual cycles should be fit with a sine wave or such isn't a simplification of much more complex processes. Its not like I have an actual solution for that but I wonder if some residuals in the fits are the

C321

result of fitting complex and complexly varying processes with quite simple functions and whether there is more information hiding in the data than these sorts of fits usually yield. I don't expect a response to this more just thinking with my fingers...

But let me respond anyway. First note that it is not just a sine wave that we fit for the seasonality, but rather a linear combination of sines and cosines with varying periodicities. Certainly the processes that generate seasonality in the fits are far more complex than a simply linear superposition of sine and cosine terms with varying periodicity. But since we are not aiming for an attribution of variance to processes, it doesn't matter that the underlying processes are complex. All we require is that the fit is robust (avoids over-fitting) and explains as much of the variance as possible. Yes, some of the regression model residuals may result from the character of the seasonality varying from year-to-year in some stochastic fashion, but that just becomes part of the variance not explained by the regression model.

Page 5, line 9-15

Some screening of the source data was performed before the monthly means were calculated, specifically SAGE data below 18 km, SAGE II data below 10 km, and LIMS data below 25 km were excluded since they were found to include occasional anomalous values which biased the monthly means. For data from ozonesonde flights, only data from flights with normalization factors (integrated ozonesonde profile divided by independent total column ozone measurement) between 0.9 and 1.1 were used.

This is correct.

1. Did the nature of the anomalies require completely excluding this data?

It would have been possible to conduct a more careful screening of the data based on a separate validation study so as to include more data. But since we were not constrained by insufficient data, we felt it best to be safe and rather exclude more data instead of trying, perhaps subjectively, to select potentially good data from a mixed set of good and bad data.

C322

The current wording suggests that some filtering process may have allowed more data to be used. Some additional information would be helpful.

This has now been clarified in the text as 'Some screening of the source data was performed before the monthly means were calculated, specifically all SAGE data below 18 km, all SAGE II data below 10 km, and all LIMS data below 25 km were excluded since they were found to include occasional anomalous values which biased the monthly means. Rather than attempting to separate the reliable measurements from the unreliable measurements in these altitude ranges, and because the analysis is not data limited, all data in the altitude ranges for these data sources were excluded'.

2. Were the normalization factors used in the analysis or were factors used only for filtering purposes (but not applied)?

As stated in the manuscript 'The normalization factors were applied to correct the ozonesonde data'.

I am not a sonde expert but I don't think that the usual practice (currently) is to apply the normalization factors.

It depends very much on the intended use of the ozonesonde data. Applying normalization factors very different from 1.0 as a correction can distort the shape of the ozonesonde profile. But since we are only using data whose normalization factors are between 0.9 and 1.1, this is less of a concern. To get the best quality data possible, the normalization factors were applied as corrections.

Page 6, line 4-5

For each month and 5° latitude zone, at least 6 measurements were required to calculate a valid monthly mean, 1. Any concerns regarding the timing of events in a month (all early or late in a month rather than evening spaced)?

This is a good and valid point raised by the reviewer. This potential sampling bias was corrected for in the analysis. As stated in the manuscript 'While each measurement

C323

passing this initial screening contributes to the calculation of a monthly mean zonal mean value, uneven sampling of the individual ozone measurements in time, latitude, or longitude could introduce biases. To correct for these biases, each ozone measurement was scaled by the monthly mean zonal mean total column ozone divided by the daily total column ozone at the latitude and longitude of that measurement'.

Page 9, line 1-2

This transformation also prevents the regression model from producing negative values when applied globally, which can occur otherwise. 1. Not wishing to sound snide or anything but positive values are not necessarily better than negative ones even if esthetically more pleasing.

True, but I think that when it comes to ozone concentrations, the reviewer would hopefully agree with us that positive values are better (in the sense of being more physically realistic) than negative values.

Response to Anonymous Referee 2

The article shows a well developed dataset ready for the community to use. The article is well written, detailed enough to follow the essence of the work, yet not overly tedious. I see no significant holes in the discussion or presentation.

We thank the reviewer for taking the time to review our paper.

Technical corrections: Page 1006 line 18 should be: "no basis function is included"

Thanks and fixed.

Page 1008 line 4 should be "fit coefficients derived by applying"

Thanks and fixed.

---

Interactive comment on Earth Syst. Sci. Data Discuss., 5, 999, 2012.

C324