**Atmospheric
Chemistry
and Physics**
Discussions

Open Access

Interactive comment on "Intercomparison of vertically resolved merged satellite ozone data sets: interannual variability and long-term trends" *by* F. Tummon et al.

Anonymous Referee #1

Received and published: 23 October 2014

This paper summarizes the most recent works regarding the creation of vertically resolved composite ozone data sets, each a result of merging data from two or more instruments. The authors describe the individual instruments used and merging methodologies for each of seven merged data sets. The paper goes on to assess the different annual cycles of the different data sets and discusses potential biases. A pre-established regression technique is applied to derive long-term trends in ozone in each of the merged data sets. Overall this analysis is simple and straightforward and, while the paper does not draw definitive conclusions regarding the overall quality and usefulness of each merged data set, it does present a suitable summary paper for the SI2N

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

special issue. However, I do have a few questions/concerns regarding this paper.

1 Major Comments

Section 2.2 talks about how the monthly mean profiles of each of the seven data sets are all converted to a common grid. Is this a common pressure and latitude grid or just a common pressure grid for each? What is the resolution of the pressure grid that is used? If a spatial interpolation is performed, how exactly is that done?

Equation 1 shows the regression model, where each coefficient A-H is actually some number of coefficients as part of Fourier pairs multiplied by some predictor parameter. The text on page 25072 states that the number in the subscript is the number of Fourier pairs. Does this mean that if NB=2 there are 2 pairs and thus 4 terms (12 month and 6 month, sine and cosine for each) or are there only 2 terms (12 month, sine and cosine) and thus 1 pair? Additionally, shouldn't each coefficient also have a constant term to represent the mean value (not seasonally varying) of each predictor?

At what vertical and spatial resolution is the regression model applied to each data set? Perhaps this ties in to a previous question about whether there is a standard spatial grid.

Section 3.1 discusses the annual cycles of each data set, computed simply as the mean of a particular month over the entire record and not from coefficient A in the regression. Do the results of coefficient A of the regression agree reasonably with the results of section 3.1? If not, their removal to derive trends from anomalies could introduce biases.

The SAGE-OSIRIS anomalies from the MDM in Fig. 5b look odd around 2001. The data approach and then are identically zero in the time period in late 2001 when no data exist there. This behavior is likely an anomaly of smoothing involving non-existent

data and needs to be corrected. Possibly similar results are seen near 1986. I would think the code used to generate the data sets in each of these figures needs to be corrected to properly handle data gaps and rerun on each data set to ensure any potential anomalies are corrected. I cannot tell, given the overlapping data, if a similar problem exists for Fig. 6.

2 Minor Comments

Table 1c states that SWOOSH uses a number of instrument data sets, including SAGE III v7.0. I am aware that there is a new version of SAGE II data, but was a new version of SAGE III data also released and used for that work?

Table 1e: Just to clarify, the data screening mentioned refers to the method of screening used to create the merged data sets, correct? It does not refer to any kind of screening the authors used for this work in relation to the merged data sets provided to them.

Table 1c states that SWOOSH uses Aura MLS v2.2 but Table 1e states that data filtering uses guidelines from v3.3. The paragraph about SWOOSH on page 25699 also states that Aura MLS v3.3 was used.

Equation 2 shows the autoregressive model used to account for autocorrelation in the regression. I believe there is a typo here where ϵ_2 is not supposed to be multiplied by time but rather is a function of time (like ϵ_1). I just wanted to clarify the authors' intent.

Pg. 25704, Line 1: "Averages and standard deviations were only calculated for months that had data for more than 20 of the 28 years available for analysis." Figure 3 states that data was only used if there was over half (14) of the 28 years. I was curious which criterion was used.

3 Grammatical/Typographical Corrections

Pg. 25692, Line 22: "they do no have" should be "they do not have"

Pg. 25700, Line 06: "Each . . . use" should be "Each . . . uses"

Interactive comment on Atmos. Chem. Phys. Discuss., 14, 25687, 2014.