Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2018-56-RC2, 2018

Earth System Science Data
Open Access Discussions

Interactive comment on "An updated version of a gap-free monthly mean zonal mean ozone database" *by* Birgit Hassler et al.

Anonymous Referee #2

Received and published: 7 June 2018

This manuscript describes the next generation of the Bodeker Scientific data set BDBP v1.1.0.6, known as BSVertOzone v1.0. Most notably this version includes AURA MLS data, a means of accounting for offsets and drifts between source data sets, and an explicit treatment of uncertainties. In addition the filling of gaps in the data set is now done independently of the full regression fit. The manuscript is well written and well resourced. The data set is a unique product, and access information for the final data product and all inputs are given in the manuscript. I have mostly minor comments/questions, and recommend publication after these issues have been addressed or clarified.

Minor Comments: Are both MLS ascending and descending profiles included? To that end, is there any attempt to account for diurnal ozone variations when combining the

target data sources? Does the CTM sufficiently capture day/night variations? Update: I see mention of this at the end of the article. However, does the CTM 12-hour resolution provide any diurnal information?

On the vertical coordinate transformations, do the various data sources provide the same pressure/temperature information? Would it be more consistent to use the same temperature/pressure data to do all the conversions rather than use the sources provided with each satellite data set, which may vary? Has the sensitivity to pressure/temperature fields been tested or considered as part of the measurement uncertainty (I may have missed this in previous papers)?

P9 On the description of the homogenization technique, I did get a bit confused. After reading it a couple of times, I think one problem is the word merging. I think of merging as going from multiple measurements to one in some fashion, but if I understand correctly you are adjusting individual measurements and then accumulating more and more measurements into the standard (as opposed to averaging monthly zonal mean fields at this stage).

If this is correct, I suggest the following wording tweaks be considered in lines 3-8 " . . . is a sequential process where each measurement from a selected satellite instrument is adjusted with respect to the standard, hereafter referred to . . . After the measurements have been adjusted to the standard, they are incorporated as part of the standard, and the new set of standard measurements is used to determine the adjustment for the next set of target measurements. This process is repeated until all satellite-based and ozonesonde measurements have been homoginized.

P9 Step 2: I suggest using a bit more precise wording here. I first read it as area weighting measurements within a 1 degree band. What about "Calculate an error weighted monthly mean zonal mean of the differences at 1 degree zonal resolution, then scale each bin average by the cosine of the bin's central latitude. Is that what the authors meant? What is the error weighting, 1/sigma or 1/sigmaˆ2?

P10 Step 6: suggest Incorporate the adjusted measurements. . .

P11 In the bootstrap steps 3 and 4, I'm wondering if the order matters. It seems measurement uncertainty should be locked in time. That is, first the residuals should be randomly rearranged in time to represent a different noise structure (Step 4), then the measurement uncertainty should be added (Step 3). If a given instrument has a period of time where the measurement uncertainties are higher, those higher measurement uncertainties should occur at the same time, the time of the known issue. If the larger uncertainties occur over say a month, but then are randomly rearranged to occur at 30 random days over the domain, the integrity of the measurement uncertainty is lost.

I also got a bit confused as to what was being bootstrapped. The fit, and thus the residual, is to the monthly mean zonal mean difference field. Is the measurement uncertainty in bootstrap Step 3 that of an individual profile, or of a particular monthly mean zonal mean bin? If the latter (which I was thinking it should be), how is this uncertainty computed? At step 4 it seems to go back to the individual profile, but still the residual uncertainties being added are for monthly mean zonal mean values. If it truly is going back to the original data, the new data with the added uncertainties are then re-averaged into monthly mean zonal means, thus beating down some of that variability. Or is it just that each monthly mean zonal mean value is given a random residual to construct the new difference field, and then this new field is fit with the regression model? That approach makes more sense to me, to work with the monthly mean zonal mean values, rather than revert all the way back to individual profiles. It seems the month to month noise from the residual is being used to replicate the day to day noise within the month of the individual measurements.

P11 L20: In reference to Figure 3, it would be helpful to state what the original standard was at this level. You can see from the plot that sonde data didn't change, but adding it to the text would be good. Also, why are the sonde data so limited before 1995? How were the SAGE adjusted in this case, is the CTM used over a multi-year period to make matches between SAGE and sonde? Finally, in reference to Figure 3, is there

published literature to compare the stated offsets for MLS and HALOE, or do direct comparisons between HALOE and sonde/MLS and sonde back up these values? Even with diminished match-ups, a difference approaching 40% should be detectable.

P15, Eqn 9: Did the authors do any sensitivity testing with the regression fits to two levels, in particular in cases where the ozone at the two levels is highly correlated. I would think that often using fits to both levels is not needed, are there specific situations were the fit to two successive layers is particularly useful?

P18: The details in the differences in Tier 1.x are difficult to see. Maybe a particular year or event period can be highlighted, showing the full time series for Tier 0 and Tier 0.5, but a shorter sample period for the other data sets to point out specific features.

Did the authors look at the consistency between the integrated Tier 0 and Tier 0.5 and the total ozone? I think such a comparison would be useful as a validation point and to show the influence of using the total ozone to fill the data set. With a full profile data set and independent total ozone merged data set from the same group, users might naturally work with both simultaneously, so analyzing their consistency would be useful. This is a validation that the other merged data sets cannot easily do because they do not have full vertical resolution.

Technical Corrections: P2 L20 monhtly → monthly P3 L4 beside → besides (or in addition too) P3 L7 suggest wording change, maybe still have limited coverage in the troposphere P3 L11-12 suggest wording change, maybe "This is particularly important when seeking to detect the small but expected signal of ozone recovery due to reductions in ozone depleting substances." P3 L24 (Sect 2) P4 L30 where → were P5 L2 remove "for measurements" P7 L14 remove "," after Dhomse references P7 L17 heavy side function → Heaviside Function P8 L3 It would be helpful to include the approximate pressure corresponding to 15km here. P8 L6 I'm not sure of the meaning of the sentence starting "As each ozonesonde is individually prepared..." I would argue that the inherent physics of the satellite measurements vs sonde is the primary reason the

ozonesonde measurements are more reliable in the troposphere (not better calibration/validation). Or are you saying that we know sonde data are more reliable because of the very precise calibration/validation? In any case I think you could easily remove this sentence. It seems to repeat the sentence before, which stands on its own. P8 L12 '. . . ,or temporally and spatially highly-resolved output from a CTM, ...' P9L21 were corrected to the extent possible (remove "for") P11 L2: suggest "This step represents the influence of measurement uncertainty on the residual." P15 L24 remove "used" P19 L1 and Figure 8 Caption: green -> red P20 L3 (85deg S to 90 deg S, 58 hPa) P22 L1 remove "different" or "selected" P22 L12 different data sources P22 L16 remove "of" P23 L4 though some discrepancies P23 L23 each individual measurement P23 L29 and therefore reduce P24 L12 remove "planned" (used later in sentence) P24 L14 remove "," after measurements

---

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2018-56, 2018.