Earth Syst. Sci. Data Discuss.,
doi:10.5194/essd-2016-16-RC1, 2016

Open Access Earth System Science Data Discussions

Interactive comment on "The Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) database: A long-term database for climate studies" *by* Sean M. Davis et al.

A. Dessler (Referee)

adessler@tamu.edu

Received and published: 15 June 2016

This paper describes a combined water vapor and ozone data set. This could be a widely used data set and the paper describes it well. I recommend publication after the authors consider my comments below.

My most important comment has to do with the construction of the combined anomaly data. Looking at earlier versions of SWOOSH, I remember that the different satellites had different shape seasonal cycles (at least in certain places), and this caused problems in the combined anomaly data. So I ended up doing my own merge by taking the anomaly time series for each individual satellite and merging those, which required subtracting the offset during their overlap periods. I presume this problem with the sea-

sonal cycles is still an issue. I'm not recommending they change anything in how the data set is presently constructed, but I would suggest that the authors mention something about this as well as the fact that users of the data might want to create their own merged data set.

page 7, line 8: The conserved quantity in HALOE is between 1.7 and 1.9 ppmv H2O/ppmv CH4, as described by Dessler and Kim [Determination of the amount of water vapor entering the stratosphere based on HALOE data, J. Geophys. Res., 104, 30,605-30,607, 1999; see Table 1]. I don't know if that'll make a difference, but the authors should use the correct value.

Also, does "discard the profile below 15 hPa" mean altitudes below 15 hPa or pressures below 15 hPa?

page 10, line 5: In the calculation of the uncertainty, do the authors take autocorrelation into account? Because of temporal autocorrelation, the number of independent pieces of information is less than the number of months, and this needs to be accounted for. See, e.g., Santer et al. (2000), Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, J. Geophys. Res., 105, 7337-7356, doi: 10.1029/1999jd901105.

page 10, line 26: Something is wrong with this sentence; the word "seven" doesn't make sense.

An aside: the authors spend a lot of effort matching satellite measurements with other satellites and sondes. These approaches have been done for decades, and I find them pretty unconvincing. The "match" criteria is set to get a large number of matches rather than because the criteria actually designate similar air masses. While I don't doubt the answers that the authors get are right, I think they'd get pretty much the same answer if they just compared zonal average (perhaps using equivalent latitude). I'm not suggesting they change anything, just letting them know what I think.

page 14, line 22: I know what equivalent latitude is and I still couldn't understand what the authors were trying to say. Please clean up this definition and make it understandable. Also, is the equivalent latitude calculated independently at each altitude? (in other words, could measurements at different altitudes from a single profile have different eq. latitudes?)

Fig. 10: I don't see a big difference between geographic latitude and equivalent latitude in Fig. 10. The real advantage of equivalent latitude is not in the zone average, where excursions tend to cancel, but the fact that there is much lower scatter among measurements made at a particular equivalent latitude then at the same geographic latitude. The authors might want to find a better way to demonstrate the benefits of eq. latitude.

page 19, line 5-6: 82 hPa is not in the lowermost stratosphere.

Interactive comment on Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2016-16, 2016.