Earth Syst. Sci. Data Discuss., 8, C258–C262, 2015
www.earth-syst-sci-data-discuss.net/8/C258/2015/

Open Access Earth System Science Data Discussions

ESSDD
8, C258–C262, 2015

Interactive Comment

Interactive comment on "Gridded global surface ozone metrics for atmospheric chemistry model evaluation" *by* E. D. Sofen et al.

M.G. Schultz (Referee)

m.schultz@fz-juelich.de

Received and published: 21 August 2015

The article by Sofen et al. describes a major undertaking which aims at pulling together all available surface ozone datasets from publically accessible network data centers across the world in order to make them available for (global) model evaluation in a harmonized, gridded format. This dataset constitutes a valuable asset to the global atmospheric chemistry modelling community and the authors need to be applauded for their rigorous efforts to collect the data, identify metadata and data issues, and harmonize the datasets into a single collection.

As is well described in the article, the processing of thousands of datasets from different networks constitutes a great challenge, and not all existing data centers have yet

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

managed to provide all their datasets in fully quality-controlled and harmonized form. It can only be hoped that this article, alongside with other efforts within GAW, TOAR, and other communities, will lead to a better "global view" on these data which are often sampled in order to adhere to national or regional legal requirements.

The article is generally well written and concise, but I do have some issues with the Introduction section which are outlines in the detailed comments below. I also noticed one or two minor flaws with the applied methods or their description. Specifically, there is no explanation of any "data capture" criterion, which makes it difficult to understand section 3.6 on seasonal AQS sites. Also, I am somewhat concerned about the gridding procedure if high and low altitude sites are merged together in one grid box. Ozone mixing ratios can vary by up to 20 ppbv within 2000 m altitude, so the gridded product might be useless in mountain regions. At least some assessment of this effect should be given.

All-in-all this article describes a unique dataset which will be of great usefulness for model evaluation. Judging on the completeness of the data is difficult, because (i) as the authors themselves note it is limited to data from publically accessible archives, and (ii) the content of the original data archives keeps changing, which means that this dataset represents a snapshot at a given (download) time. In any case this dataset provides a much more complete view on global surface ozone than what has been available in previous model evaluation exercises.

I recommend publication of this article after the following minor issues have been taken into account.

Detailed comments:

Abstract: spell out metrics SOMO35 and AOT40

Introduction:

p. 606 line 5 – remove sentence "Overall ..." – this is an unnecessary motherhood

ESSDD

8, C258–C262, 2015

statement

Line 8 – I don't fully agree with the statement "they have focused heavily . . ."; there are different communities who employ CTMs for other purposes such as aerosol modelling, mercury, POPs or other substances as well. You may say that there is a strong scientific community who have focused. . .

Line 13 – "against". Perhaps re-phrase that ozone is a central species in such evaluations?

Line 15 – "limited by the availability or type of observations". This mixes two very different aspects in one rather general statement. Better to separate these two issues.

Line 18 – The discussion of the issues with satellite data is too vague ("makes understanding the vertical distribution difficult"). It is simply physically impossible to obtain a good vertical resolution in the troposphere from spectral irradiances measured by the satellites. Even though there are now products emerging which claim to have at least some resolution in the troposphere, one can at most distinguish between upper and lower troposphere. Please state more explicitly that satellite retrievals of ozone do not allow for an adequate vertical resolution in the troposphere.

Line 24 – I don't like the implication here that surface observations are only useful because there is nothing else available. There are good technical (and programmatic) reasons why people measure at ground-level (for example it is much easier to maintain stable instrument characteristics and calibrations on the ground compared to aircraft measurements). Furthermore, the boundary layer, where the stations reside, is often de-coupled from the free troposphere above, and aircraft usually can perform only very little sampling in the boundary layer. Hence, surface observations sample a different part of the atmosphere which is particularly relevant, because this is the air we (and the plants) breeze.

p. 608 line 10 – "For convenience"; I don't think that this is the major reason for (re-

ESSDD
8, C258–C262, 2015

gional) model studies to have limited their evaluation to one region. Most of these models were built to support regional air pollution abatement policies, and so it is natural that they would chose data from their region as the only battlefield. I think it would be worthwhile if the authors included a somewhat more meaningful discussion about the differences between regional air quality and global model evaluation. Section Datasets:

p. 610 line 3 – the authors should add a sentence here when these data were downloaded or last updated. The content of the archives changes frequently and it is important to have a reference date for the dataset described in this manuscript.

Line 26 – maybe worthwhile mentioning that Airbase also contains data from some oversees representations (for example in the Carribean).

p. 610 line 10 – "are removed." Please add "in this analysis."

p. 611 line 24 – "unlikely to adhere to quality standards". Strong statement. How would the authors know if they cannot access these data? At least replace "and" by "or" in line 25?

p. 612 line 1 – for your information: at a recent meeting on GAW data management strategies the conclusion was almost opposite to your recommendation. Rather than one central archive one wants to work towards a federated, distributed archive. There are very strong programmatic and funding reasons why several data archives exist, and this will not change. Thus, rather than trying to bring everything together in one place the strategy is to harmonize data access and formats and build a central portal from where the individual archives can be accessed, ideally entirely transparent for the user.

p. 614 line 2 – AQS sites do come with a station identifier "AQS site id" in the metadata file (format "NN-NNN-NNNN")

p. 617 line 9 – Units correct? "ppmv ppbv-1"?

p. 618/619 – there is no mentioning of data quality flags in the original datasets. In our

ESSDD
8, C258–C262, 2015

experience, many outliers disappear if these flags are correctly taken in into account. It is a problem, though, that in some networks the definition of the flags is not very clear or changes between stations. Specifically with respect to "missing values" we also found that there are cases when these are not flagged as invalid data (and the code used to mark a value as missing, i.e. the number reported in lieu of an ozone value, can also vary – even within the time series of one site.

p. 620 lines 1-14 – it is not clear to me what is the intention of this analysis? Do the authors want to exclude sites with incomplete annual series? Or do they want to detect these and include them only in statistics over the summer months, for example? I believe the problem here is that a "9 months minimum" data capture criterion is referred to without introducing this first. In general, it is a bit difficult to follow the text in section 3 where stations are removed in various steps of the analysis, without having said at the beginning what the criteria are for a "good" data record.

p. 621 line 23 – grammar "reported hour"?

p. 625 – this might be the only concern I have about the robustness of the methods applied: when combining information from mountain and "valley" stations in one grid box, the information in these grid boxes might be much less meaningful than in flatter terrain. It might be useful to create an additional dataset which includes only stations below or above a given threshold altitude. At least the impact of this "altitude merging" should be assessed.

p. 638 Table 3 – I am somewhat doubtful about the usefulness of the air quality metrics DaysMDA8_ge_60ppb_Gridded_Ozone and FourthHighestMDA8_Gridded_Ozone in a gridded product. These metrics analyze the high end of the distributions, and these are highly variable across sites from a single region. Have the authors demonstrated that a model can develop any skill with respect to these metrics?

Interactive comment on Earth Syst. Sci. Data Discuss., 8, 603, 2015.