

# Review of the paper by M. Coldewey-Egbers et al.

---

The paper is dedicated to the very important issue: creating the essential climate variable for ozone. The GTO-ECV is based on merged data from GOME, SCIAMACHY and GOME-2 total ozone column measurements. The paper is interesting and, in general, well-written. Please find below my comments and suggestions related to the analyses presented in the paper.

## MAIN COMMENTS

1. Please explain and discuss the selection of the horizontal resolution  $1^\circ \times 1^\circ$ . Is it based on the characteristic structure of the ozone field? Will sampling uncertainty be reduced if you select a coarser resolution (e.g.,  $2^\circ \times 2^\circ$ )? Please discuss also the influence of different horizontal resolution of GOME, SCIAMACHY and GOME-2 Level 2 data on actual horizontal resolution of daily Level 3 maps.

## 2. Merging strategy

a) Please explain the selection of GOME-2 as a reference instrument (the most stable? If yes, provide the corresponding references or analyses)

b) Please explain the rationale for using only one instrument at a time. Would the alternative approach, using the data from all instruments in their overlapping period, reduce the sampling uncertainty?

## 3. Validation of Level 3 data.

From my point of view, the analysis presented in Section 3 is somewhat misleading. First, you explain that the representativeness of the ground-based network is insufficient in many places, and therefore cannot be used for monthly mean comparisons. Despite this, you then compared the monthly mean satellite and ground-based data and - as expected! - found deviations in some cases. Such comparison cannot provide information about the real quality of the GTO-ECV data.

From my point of view, it would be much more informative to compare daily Level 3 data with coincident ground-based measurements in  $1^\circ \times 1^\circ$  bins. The validation results will be, probably, close to those of Level 2.

Furthermore, it is important to demonstrate that the applied correction improves the long-term stability of the dataset, compared to original datasets. This important analysis is missing in the paper.

If you want to compare monthly mean data, then only the bins with uniform coverage by both satellite and ground-based data should be used. This would avoid speculations on the influence of non-uniform sampling.

In short, I suggest revision of the analyses presented in Section 3, and corresponding text and figures (and including only the results that illuminate the quality of GTO-ECV).

4. Terminology and related issues.

a) “The standard error” (Sect. 2.3) is not a good term, because it is estimated as a sampling bias, and thus should be named properly.

It should be also clarified: what parameter should be used as uncertainty of the monthly mean data?

b) “Seasonal variability” (Sect. 3.2.2) is also not a proper term for seasonal and latitudinal dependence of biases with respect to the Dobson network data.

c) Section 3.2.4. As explained above, I think that any conclusions about the data quality should be based on data with removed contribution of sampling uncertainties.

It seem that SD (page 4622, line 8) has a different meaning than the parameter with the same name introduced in Sect. 2.3

I have not understood the parameter “monthly mean variability”. Is this based on biases? Variance of a parameter  $x$  is  $\text{var}(x) = \langle (x - \bar{x})^2 \rangle$ , but your unit is %, not %<sup>2</sup>. Perhaps, a formula would help.

It would be advantageous to rename “seasonality” into “seasonal variations of biases” and “latitude” into “latitudinal variations of biases”.

5. The 2d histogram presented in Fig.14 indicates relatively large differences between GTO-GDP and GTO-CCI in some cases. Please discuss this in more detail and indicate the reasons for the differences.

6. P.4625, A discussion on “a small but significant time difference” (lines 15-20) looks not good without an illustration. Can this difference be seen in Fig.16 (bottom)?

#### **DETAILED COMMENTS (small clarification, rephrasing, technical corrections)**

P.4610, L. 28. “the third article” Please indicate directly what are first two papers.

P.4616, L.11: explain the abbreviation SD

P.4616, L. 24-26: “left... right panel” -> “top... bottom”, respectively

P.4618, L.1 “IFS-MOZART” explain the acronym

Figure 7 is illustrative. However, it should be specified which datasets are included to “ground-based measurements”. Since the comparison is performed with the each type of

ground-based instruments separately, it would be reasonable to perform such analysis for each type of ground-based measurements.

P.4618, L.21: "Any level 2 comparisons" - >"All level 2 comparisons"?

P.4622, L. 5-6: "(left)...(right)" . Please indicate also number(s) of Figure(s).

P.4624, L. 14 "Both data records", please indicate them directly here.

Figure 16 (bottom). Please put a gray background for separation of small values and areas with no data.

P.4627, L.10-11. Perhaps, the sentence with explanation of CCMs is not needed.