

Review for Atmos. Chem. Phys.

“Harmonisation and trends of 20-years tropical tropospheric ozone data”

Authors: Leventidou, Weber, Eichmann, Burrows, Heue, Thompson, and Johnson

Review submitted: 18 March 2018

General Comments:

This review is of the revised paper. While the authors attempted to address many of the concerns raised by the other reviewer and me, I’m afraid I still have some important concerns about this paper. It is not at all clear that Scenario 1 is the best approach (or better than any of the other 5 scenarios the authors test). As a result, the tropical trend calculations are more uncertain than the paper communicates. Perhaps a Monte Carlo approach is necessary to better characterize the uncertainties in the trend analysis? I’m afraid this paper still needs another round of revisions and should not yet be published in ACP.

Comments on Response to Reviewers:

“...and Hilo is strongly affected by volcanic outgassing resulting in negligible ozone concentrations in the boundary layer.” That’s not quite true. More accurate would be that the tropospheric ozone data in Hilo often shows interference from SO₂. It’s not that the ozone isn’t there; it’s that the ECC measurement doesn’t work well in regions with significant SO₂ (see Komhyr, 1969 and Morris et al., 2010).

“Tropical troposphere ozone trends critically depend on the merging/harmonization approach.” Agreed. But a problem for the authors is that they have not shown any one approach is better than another!

Page 1

Lines 13-14: “...was applied for GOME, and mean biases...were calculated and applied...”

Lines 21 – 21: Since you’re citing “decreases,” I believe you do not need negative signs in front of the magnitudes.

Page 1 – second one

Line 10: “...during the monsoon period...”

Line 15: “...on the order of...”

Line 16: You repeat “year-1.”

Line 34ff: This is the first point at which we find out what THIS paper is going to do. It's a lot of introductory text. Might be good to get to this earlier and better integrate how previous studies shape/motivate the need for this one.

Page 2 – second one

Line 13: "...integrated (up to 200 hPa)..."

Line 14 – 15: Rework. Something is not right...

Line 16: Since you've got two sources listed, delete the word "One," and change to "Large sources of uncertainty are..."

Line 28: "...because SCIAMACHY is the only instrument..."

Page 3

Line 5: Seems like the sentence should end, "...GOME and GOME-2)." Just delete the rest of the sentence.

Line 21: I thought in the response to the reviewers that the authors were going to limit their analysis to 15S – 15N, yet here they reference results in the 17.5 – 20N latitude band. See also Figure 1 on Page 4.. Why not limit every part of this analysis to the more restrictive latitude band?

Page 5

Line 4: "...with respect to SCIAMACHY is added..."

Page 6

Line 16: As I remarked in my comment on the response to the reviewers section, this statement on "high SO₂ emissions, resulting in negligible ozone concentrations..." is not accurate. See my earlier explanation.

Lines 3 and 4: Is it "ozone sondes" or "ozonesondes." I'd pick the latter.

Line 14: "For these reasons, scenario 1 has been selected..." I can't figure out a good justification here – lots of issues. The next page contains a table of the ways I've sliced and diced the data you provided and upon which you based your decision. Your analysis looked at the mean bias and took the one closest to 0, which led to your choice of Scenario 1. I've added the standard deviation calculation to the calculation of the mean. As you can see, in every scenario, the standard deviations of the data are greater than the mean biases. Thus, I would argue that the differences in the means are statistically insignificant. This approach is not a good one for selection of the best scenario. Furthermore, if we just eliminate American Samoa, the conclusion is not robust: now Scenario 6 is the clear winner by the smallest mean (although again, the standard deviations exceed the mean biases in every case). Better, I think, is looking at

the root-mean-square bias, because you're really interested in which approach produces the smallest magnitude bias on average rather than the smallest mean of the biases (e.g., biases of -10 and + 10 would have a mean bias of 0 but a rms bias of 10). If you use this approach (which I believe marginally better), you conclude that Scenario 2 has the lowest mean bias, but again not statistically significantly different from any of the other scenario means. Finally, if you eliminate American Samoa, Scenarios 1, 2, 5 and 6 are all pretty close.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Am Samoa	-0.89	-0.92	-1.99	-0.61	-0.93	4.59
Ascension	0.03	-0.14	-0.77	-0.42	-0.6	0.03
Java	-0.11	-0.12	-1.12	-0.54	-0.55	-0.11
Kual Lumpur	-1.81	-2.12	-2.12	-2.14	-2.48	-1.78
Nairobi	1.81	1.1	1.8	1.48	0.74	1.84
Natal	0.56	0.63	-0.21	0.22	0.28	0.57
Paramaribo	-2.98	-2.95	-3.02	-4.11	-4.34	-0.11
Mean bias	-0.48	-0.65	-1.06	-0.87	-1.13	0.72
Std. Dev.	1.58	1.46	1.57	1.79	1.74	2.01
Eliminate Samoa						
Mean bias	-0.42	-0.60	-0.91	-0.92	-1.16	0.07
Std. dev	1.72	1.59	1.66	1.96	1.91	1.17
Root-mean-square						
Mean bias	0.87	0.84	1.34	0.90	0.93	1.49
Std. Dev.	0.79	0.74	0.76	0.75	0.79	1.71
Eliminate Samoa						
Mean bias	0.86	0.82	1.20	0.96	0.93	0.87
Std. Dev.	0.89	0.83	0.77	0.82	0.88	0.89

Page 7

Line 2 – 3: “However, the biases of each scenario with ozone sondes are very close to each other for every station.” I don’t see the data in Table 1 supporting this statement. There’s great variability in both rows and columns.

Line 8: “...the scenarios that can be confidently rejected are...” I see nothing in Table 1 upon which to base any rejection of one scenario over another.

Line 13: "...has the smallest mean bias with the ozone sondes (-0.4 DU)." As you can see above, this mean bias is not statistically significantly different from any of the other scenarios.

Page 10

Figure 3: Why does panel f show more area as statistically significant in %/year trends than panel e? I think it's because the criteria in panel e is stricter (exceed the range of all harmonization scenarios), but it's confusing to have these next to one another. Why not use the same criteria? I think f makes a more interesting map than e. But based on my analysis above, I don't see one scenario as preferable to another.

Page 12

Table 2: I take it these 2 sigma uncertainties are determined by the trend analysis itself and to not include the additional uncertainty resulting from the harmonization choice itself? If that's right, these results look better (and more significant) than they are. Perhaps a better approach would be to use a Monte Carlo analysis that mixes between the scenarios and reflects the uncertainty in the bias of the scenario to figure out the total uncertainty in the trend. At this point, I have little confidence in the quoted uncertainty in this Table and as a result, the associated discussion.