Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-866-RC2, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Missing OH Reactivity in the Global Marine Boundary Layer" *by* Alexander B. Thames et al.

Anonymous Referee #2

Received and published: 26 November 2019

==========

Summary

==========

Thames and co-authors present OH reactivity measurements over the remote oceans from the ATom campaigns. They use a 0D box model constrained to other ATom observations to interpret the OHR data, specifically focusing on 'missing' OHR in the marine boundary layer.

The dataset is the first of its kind, spanning the globe, and is collected under analytically challenging conditions (very clean air). The team is to be commended for the work that went into collecting this dataset. It represents a great contribution to the community

Printer-friendly version

Discussion paper

that I expect will be used by many future researchers.

My main comments on the manuscript have to do with its treatment of uncertainties and statistics. Since we are looking at very low OHR conditions, the missing OHR values are also low and pushing the uncertainty limits. I feel that the paper needs a more sophisticated and robust treatment of uncertainty statistics (including in the modeled OHR, which is itself constrained to measurements) in order to provide a convincing case that the missing OHR values are indeed robust. Once that is done the paper should be published in ACP.

==========

General comments

==========

Along with uncertainties in the OHR measurement itself, the "modeled" OHR also has uncertainties. It is constrained to trace gas measurements, which have their own uncertainties. It predicts unmeasured species using rate coefficients that have uncertainties. It seems to me that in order to judge whether the missing OHR is statistically robust, these uncertainties need to be fully propagated through the modeled OHR derivation. Then one could do a proper statistical comparison of the measured and modeled OHR values.

164: "these variations were tracked with measurements of the OHR instrument background in the laboratory". I am confused by this because an earlier statement (line 141) appears to indicate that the background was measured every measurement cycle. Please clarify.

180-184: how can we be confident that the pressure dependence of the calibration is prescribed accurately enough to define (for example) a mean 0.5 1/s discrepancy? The relatively large amount of scatter in Figure 2 (e.g., for the ATom-1/4 calibration) by itself does not by itself inspire confidence in this respect. Given the small OHR discrep-

Printer-friendly version

Discussion paper

ancies that are discussed later, I feel the paper needs a more rigorous discussion of the background and calibration uncertainties, along with a quantitative analysis of how these propagate onto the end results, for us to have confidence in the findings.

It doesn't appear that a pressure-dependent calibration curve was performed at the time of ATom-1. How are we confident that the 2018 curve fits the ATom-1 data given the 2-year separation in time?

190-194: A pressure-invariant offset is being applied to the measurements based on model output for the upper free troposphere. Please indicate the magnitude of this offset that is being applied (e.g., compared to the inferred missing OHR magnitudes that are discussed later). Is there reason to believe that this offset is in fact pressure-independent as assumed?

207-208: "Therefore, in each ATom phase, the total uncertainty in the OH reactivity is dominated by the instrument background uncertainty." My interpretation of this is that we should be considering the errors as primarily systematic rather than random. I.e., the campaign-specific background at any given pressure is a single constant quantity that we can define to 1-sigma of 0.4 1/s. And therefore that uncertainty is not reduced by temporal averaging of the campaign measurements: the background uncertainty is the same (0.4 1/s) whether we are considering 1 measurement or thousands. Is my interpretation correct? If so then I don't believe 1-sigma is an appropriate metric, since sigma is a measure of variability rather than of certainty about the central value. A more appropriate metric would be the 95% confidence interval about the pressure-dependent backgrounds – for example, obtained via boostrap analysis of the data in Figure 2.

299: "The median measured OH reactivity equals the median model-calculated OH reactivity to within $\pm 1\sigma$ statistical uncertainty", see above, sigma is a measure of scatter rather than uncertainty

Figure 2, "Darker grey points indicate OH reactivity values greater than the 1-sigma

uncertainty in the MBL." Wording is unclear here. At first I thought it meant "missing OH reactivity values greater than the 1-sigma. . .", but from the plot it looks like the colored values are just those where the actual OH reactivity is 1-sigma above the median value. Please clarify.

305-311 and Figure 5: There is some conflation of spread and uncertainty here. First, the "For missing OH reactivity to be meaningful, some missing OH reactivity points must be much greater than the statistical spread of the OH reactivity measurements." A bit oddly worded, rather one should say that to be meaningful, the missing OH reactivity should exceed the statistical uncertainty of the OH reactivity measurements. Spread and uncertainty are not the same thing. Similarly, in the Figure 5 caption: "Dotted black lines represent +/-2-sigma uncertainty derived from a median of the missing OH reactivity values greater than 4 km." If the lines are just twice the SD they are showing the spread, not the uncertainty. And wording-wise it is not clear what "2-sigma uncertainty derived from a median" means. Finally, "About 95% of all points above 4 km are within that phase's 2σ uncertainty bands, which is consistent with a statistically normal distribution." – again mixing up variability with uncertainty.

313-327 and Figure 6: This (qq-plots and t-test) is a nice demonstration that the missing OHR data above and below 4km follow differing statistical distributions. Please discuss the robustness of this finding in view of i) the statistical uncertainty of the pressure-dependent background corrections in Figure 2, ii) the propagated uncertainty in the modeled OHR, and iii) the assumption of a pressure-invariant offset (line 190-194). Second, the figure is only showing ATom-2 data but the text (by not mentioning this) implies that all of the data from ATom-1, 2, and 3 have this feature. Is that the case?

333-334: "The latitudinal dependence implies that air or sea temperature or other latitude-dependent factors contribute to missing OH reactivity." Also, the highest missing OHR values fall in the NH, implying that the generally higher abundance of trace gases in the NH plays a role . . . right?

342: “the main correlations that stand out are. . .” please be more precise in your language here, are these the 4 variables with the highest correlations?

342-350 and Figure 8: please discuss whether these correlations persist when the campaigns are considered individually.

363: “the 1-sigma confidence level”, please see earlier comments about confidence intervals. What is needed here is a statement of whether the slopes agree to within (say) 95% confidence based on a bootstrap / monte carlo test. In the same way, please also indicate whether the slope is significantly different than zero.

372-373: “become substantially less than observed” and “become greater than observed”, please be quantitative

My understanding is that the NO2 measurements during ATom have high uncertainty. Is that right? Are you using the measured NO2 or is this being predicted by the model from other species?

==========

Minor / technical / language corrections

==========

There are some minor grammatical errors throughout; please do a careful proofreading.

29: “which IS 0.5 s-1 larger”

35-36: “for much of the free troposphere”, awkward, suggest “throughout much . . .”

45: suggest “with THE hydroxyl RADICAL”

46-47: “is lost by the sum of the reaction frequencies”, wording is not quite right b/c the loss is via the chemical reactions themselves, the frequencies just determine how fast that occurs. Suggest “is lost at a rate determined by the sum. . .”

47: suggest “This sum of loss frequencies is called. . .”

68: “exceeded the calculated AMOUNT by”

69: VOC not defined

71: “in A northern Michigan forest”

75: “20%, which is approximately the uncertainty”. But doesn’t this percentage depend on the absolute OHR amount?

103: as stated later this 0.4 1/s LOD is for 1-minute averages, consider specifying that here

149: “in high NO environments”, please specify the approximate NO level at which this effect becomes relevant

190: “1-minute sums”, perhaps this should be “1-minute averages”

Figure 1: I don’t know that it is helpful to include ATom-4 in this Figure given that the data is not ultimately used in the analyses that follow.

230: “and other measurements were used to fill gaps in the primary measurement”. Can you please add a few words to be more specific here? E.g., “linear regression to other measurements”?

250: Need to specify assumed OH level giving this 1-hour lifetime

Figure 2: If I understand Figure 2 correctly, the blue fit is being used for both ATom-1 and ATom-4, is that correct? If so, the legend should be relabeled to make this more clear.

Fig 2: “The median OH reactivity of 500m altitude bins is shown formeasured OH reactivity (blue line, with 1σ error bars)”, in fact the error bars are only shown at 2km increments, suggest clarifying in caption

270: “These legs. . .” awkward wording

270-271: suggest stating range of MBL heights during ATom.

290: some representative OHR ranges would be helpful here.

318: “The missing OH reactivity values measured below 4 km altitude lie along the red dashed line” I think you mean “above 4 km” here.

Figures 1, 4, 5, 7: I recognize that this information is also in Table 1, but it would be helpful to your reader if you indicated the time-frame of each ATom deployment somewhere on these figures.

Table 1: a single season is given for each ATom deployment, but ATom covered both hemispheres.

336-340: do you suspect instrumental factors here?

340: “While present on some figures”, please be specific

Figure 8, “at the per-flight time resolution” is unclear, I think you mean that each point is an average over all the data for a given flight?

356-357: wording is awkward here

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-866, 2019.

Printer-friendly version

Discussion paper