Atmos. Chem. Phys. Discuss., 11, C15577–C15583, 2012
www.atmos-chem-phys-discuss.net/11/C15577/2012/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Impact of sampling frequency in the analysis of tropospheric ozone observations" *by* M. Saunois et al.

M. Saunois et al.

marielle.saunois@lsce.ipsl.fr

Received and published: 22 February 2012

Response to the Anonymous Referee #1:

We thank the referee for his useful comments on the study, which have helped us to greatly improve and clarify the manuscript. We acknowledge him for his corrections and suggestions. Below are the responses to his comments that have been quoted [. . .] before each response.

[- I would strongly suggest that the authors add a new table (or even better a detailed appendix) to very clearly summarise all the metrics used. There they could define the means (X), standard deviations (sigma), number of elements considered (N, n), and the uncertainty, separately for the following distributions: (a) all the morning profiles in

a season, (b) a subsample i, and (c) the distribution of {Xi,samp – overline X}. The appropriate subscripts and superscripts should be used in those definitions and be always consistent everywhere in the text. It would also be clarifying to indicate why the uncertainty is sometimes given by "3 x sigma" but others by "3 x sigma divided by the square root of N-1" in the text (I refer to this below).]

We thank the referee for this comment and the following suggestions. As suggested we added a Table (new Tab.1) in which the metrics are clearly defined and summarized. This helped us to use more consistent writings (sub and super scripts) among the metrics. Also an Appendix has been added in which the metrics are introduced and the method is described in greater details than in the former section 3.1 "Definition of the metric called sampling uncertainty". Moreover, following Reviewer#2's comments, we dedicated Section 3 to an overview of the methodology used. We encourage the reviewers to have a look at the Sect.3, Appendix and Tab.1, as well as the new notation of the metrics, which should fulfill their expectations. Also a careful attention has been paid to the vocabulary used throughout the manuscript (text and legends of the figures). We now name 3*sigma/sqrt(N) the confidence limit of the standard error of the mean; for clarity we write "confidence limit on the mean" instead of uncertainty. The word "uncertainty" is kept to designate the uncertainty due to low frequency sampling, defined as 3*sigma, where sigma is the sample standard deviation of the distribution {mean(xis) –mean(x)}.

[- Page 27114: In line 9 of that page you define the standard deviation of {Xi,samp - overline X} as sigma (subscript = yr, superscript = samp). You refer to this variable again as sigma (subscript = yr, superscript = samp) in line 24 of the same page and in caption of Figure 4. However, I think you refer to it as sigma (no subscript, superscript = samp) in lines 14-16 of this page and nearly everywhere else in the paper. Please use the same terminology everywhere to avoid confusion. This is particularly important because you also use other definitions of sigma in the paper.]

We understand the confusion that could exist between sigma (subscript = yr, super-

script = samp) and sigma (nosubscript , superscript = samp). There was no mistake in the text. The sigma with the subscript "yr" refers to the distribution calculated using only one year of data. That is why this is the sigma used in Fig.4, which shows time series. However in order to obtain one value per season (and per altitude level), we aggregated the 14 yearly distributions into a single one and derived the sigma (nosubscript, superscript = samp). The section 3.4, the Appendix and Tab.1 of the revised manuscript clearly specify the difference between these two metrics.

[- Page 27114, lines 13-16 and 20-22: You should clarify if you have checked that the {Xi,samp - overline X} values and the set of all the MOZAIC morning profiles (with mean given by overline X, and standard deviation given by sigma) are normally distributed. I believe that should be the case (at least in the first distribution due to the large number of data points). If so then it is true that about 99% of the values will fall within 3 standard deviations (3-sigma) around the mean. Please also clarify why you calculate the uncertainty as "3 x sigma" for the first distribution (lines 15-16) and as "3 x sigma / sqrt (N-1)" for the mean of the whole morning data set (lines 20-22).]

We thank the referee for pointing out this missing part of the discussion. The MOZAIC morning profiles measurements are generally normally distributed except in the lowest levels due to ozone titration. For clarity purpose, we keep the values of 3-sigma as the 99% confidence limit for all pressure levels. However, we should keep in mind that near the surface the distributions deviate from a normal distribution. Given the large number of subsamples created the distributions {Xi,samp - overline X} are also normally distributed. We added this discussion in the Appendix.

[- Lines 22-26 of page 27116 and caption of Figure 5. You do not define sigma (superindex=s) and show directly overline sigma (superindex=s), which makes everything very confusing. Please clarify the differences between them and make any definition on this consistent with similar definitions on lines 20-22 of page 27114. Is sigma (superindex=s) the standard deviation derived from one specific subsample? Is overline sigma (superindex=s) the standard deviation or uncertainty derived from all subsam-

ples in a season? And how is it derived? I have trouble to understand whether sometimes you refer to a specific subsample or to an average seasonal subsample: note that in caption of Figure 5 you say that overline sigma (superindex=s) is "the standard deviation of a subsample", while on line 24 of page 27116 you say that it is "the standard deviation of the seasonal subsample mean". Any correction that you may introduce will affect the whole Section 3.4 (pages 27116 and 27117). Based on the results from Figure 5, in the last lines (19-27) of page 27117, one can read "intra-seasonal variability detected by a subsample" while I do not believe you really refer to a specific subsample.]

Following previous comments, we have paid careful attention to the vocabulary used for the metrics. As for now: sigma (superindex=s, subscript=i) is the sample standard deviation of subsample i (as it is seasonal it includes either 12 or 36 profiles); and overline sigma (superindex=s, subscript=i) is the standard error of the mean of one subsample (overline x (superindex=s, subscript=i)). We hope that now the metrics are well defined, enough to avoid misunderstanding. As a result, these two metrics are calculated for each subsample. However for Fig.5 we consider the average values of these metrics. The text and the legend of Fig.5 have been modified in the revised version to clarify this point.

[- Section 3.5 (sampling effect on ozone trends, pages 27118-27119) is hard to follow due to both the small size of Figure 6 and also the difficulties to understand the methodology. There is confusion from the very beginning (lines 2-4 of page 27118) because it is very unclear how you calculate linear trends with "points weighted by overline sigma". By the way, overline sigma is defined in section 3.2, not in section 3.1 as you indicate. Are the "points" the subsampled seasonal means (yellow and red in the figure) and the real seasonal means (black diamond) in the case of Frankfurt? And when you mention "overline sigma", do you refer to sigma (subscript=yr, superscript=samp) (see definition in section 3.1) or to overline sigma (superscript=s) (defined in Section 3.4) in the case of the subsampled seasonal means? And do you refer to overline sigma (defined in

section 3.2) in the case of the real seasonal means?]

This part has been reworded in the revised manuscript. In the linear trend the seasonal means are weighted by their standard error as explained in the manuscript. The revised version should clarify this point.

[- Caption of Figure 8: Please make sure that the definitions of overline sigma (subindex=s), overline sigma (no subindex), 3 x overline sigma (subindex=s) and 3 x overline sigma (no subindex) are consistent with any other definitions given before in the paper.]

We paid careful attention to that in the revised manuscript. Mismatched definitions between text and figures have been corrected throughout the revised version of the manuscript.

Specific comments:

[- The "regular" sampling methodology is introduced in the first paragraph of Section 2.3 (page 27112). If I understand well that method does not allow using the same profile in more than one subsample. Considering that on average there are around 70 profiles per month above Frankfurt, then the average number of subsamples that can be created with 12 profiles should be of around 5–6. When you mention that "there are USUALLY less than 10 subsamples created with 12 profiles for each month" (same page, lines 26-28), I understand that you do not refer to an "average month". Are you referring to some specific months with a higher density of profiles? This should be clarified in the text.]

Indeed, if we consider an average month with around 60-70 profiles, we can create around 5-6 subsamples with 12 profiles a month. As a result there are less than 10 subsamples created with 12 profiles for each month. This has been clarified in the text.

[- Page 27118, line 16: You indicate that the trends result from the "decrease of nitrogen oxide emissions". I would better to say that "They most probably result from the

decrease in ozone precursor emissions". I agree that the positive winter trends are very probably related to reduced titration (as a consequence of decreasing NOx), while reduced ozone production in summer could be related to both decreasing NOx and VOCs.]

This has been corrected.

[- Page 27122: When introducing Windhoek it would be convenient to indicate that it is an elevated site (around 1650 m a.s.l). That will help understand why you do not show data below the 800 hPa level in Figure 8. Is Windhoek on a plateau or is it elevated with respect to its surroundings? That might need to be considered together with your comments on the low pollution levels there (lines 14-15 of the same page; by the way, is that supported by observations?) to explain why the vertical profiles of uncertainty due to sampling do not present a C-shape.]

We thank the referee for pointing this out. Indeed Windhoek is located in the Khomas Highland plateau area (around 1,700 meters above sea level). We have added this piece of information in the revised manuscript. This has to be considered to discuss the ozone levels and variability along with the lower populated area in the vicinity of Windhoek airport compared to the Northern Hemisphere airport vicinity. More details are given on the Windhoek sources of pollution and ozone variability in the revised manuscript.

[- In section 3.2, together with Table 1 and Figures 4-5, you show a very relevant result: the uncertainty due to the low time resolution imposed can be very large in the lowest levels (1000 hPa) and for pressures lower than 400 hPa. In lines 13-15 of page 27115 you say: "Also due to higher day-to-day variability in the boundary layer and in the upper troposphere (high impact of stratospheric intrusions), the distributions are larger at these levels compared to the ones in the middle troposphere". In the case of 1000 hPa you could mention some of the potential processes responsible for the high variability close to surface (e.g. air masses close to surface are affected by fresh emissions,

subject to dry deposition of ozone, turbulence, ...). In the case of 400 hPa, I would avoid explicitly mention “high impact of stratospheric intrusions”, because you have not quantified that. Above Frankfurt, that level is still not so close to the tropopause and the impact of stratospheric intrusions will depend on the season (generally highest in winter-spring). You may just mention something like “the potential impact of stratospheric-tropospheric exchange”. And when you say that “the distributions are larger”, do you mean that “the distributions are broader”?]

We thank the referee for his suggestion. In the revised manuscript we mention some processes that could explain the higher variability in the lowest levels. Also we refer now to stratosphere troposphere exchange instead of stratospheric intrusion as suggested. Indeed we meant broader and not larger, this has been corrected in the revised manuscript.

Technical corrections (including typos) [. . .]

The technical corrections have been taken care of.

[- Understanding Figure 6 requires a big effort. Make sure it is larger in the ACP version of the paper. In addition, you are plotting the 1-sigma uncertainty of the slope against the slope of the linear trend (not the slope against the 1-sigma uncertainty as you say in the caption).]

We corrected the legend. A more detailed explanation of the figure in the text has been added to help the reader going through it.

[Finally, I am not convinced that the paper gives credit to the relevant literature. . . .]

We have taken into account the reviewer comments and give more references to previous work. The suggested literature has been added in the revised manuscript

---

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 27107, 2011.